INCREASING CONVEX MONOTONE MARKOV CHAINS: THEORY, ALGORITHM AND APPLICATIONS*

MOUAD BEN MAMOUN[†‡], ANA BUŠIĆ[†], JEAN-MICHEL FOURNEAU[†], AND NIHAL PEKERGIN[†§]

Abstract. We develop theoretical and algorithmic aspects of discrete-time Markov chain comparison with the increasing convex order. This order is based on the variability of the process and it is expected that one can get more accurate bounds with such an order although the monotonicity property is more complex. We give a characterization for finite state space to obtain an algebraic description which is suitable for an algorithmic framework. We develop an algorithm and we introduce some applications related to the worst case stochastic analysis when some high level information is known, but not the complete structure of the chain.

Key words. Markov chains, stochastic bounds

AMS subject classifications. 60E15, 60J22, 15A51

1. Introduction. Comparison techniques have gained an increasing popularity in the study of stochastic processes [23]. These techniques may be related to various mathematical theory (stochastic ordering, polyhedral theory, Markov chain decision process, stochastic recurrence equation).

In the context of numerical analysis of Markov chains, the first idea was to analyze systems which are too difficult for numerical analysis. One can compare a chain of the model with another one which is simpler to solve. A recent survey [16] presents several solutions that we can group into two key ideas: reduction of the state space or using an ad hoc structure, the numerical analysis of which is simpler. The first approach was shown to be very efficient [25]. The stochastic approach was developed using projection or functions of Markov chains [13] and an algorithmic derivation of smaller chains based on strong or weak lumpability [28, 21, 15] was proposed. Other algorithms to obtain upper Hessenberg or single input macro state chains were also proposed in [16, 8]. A tool providing all these algorithms was also demonstrated [14].

However all these approaches are based on the strong stochastic ordering ($\preceq_{st}$-ordering) among random variables. This order is quite natural because it is associated with sample-paths and coupling. Nevertheless, many other stochastic orders have been studied. For instance, the variability orders ($\preceq_{icx}$ and $\preceq_{icv}$) have been used to compare different queueing models when we change the variability of arrival or service distributions. To the best of our knowledge, very little was done to construct bounding Markov chains with these orders. Vincent's pioneering work [1] was the only reference we could find. Note that the lack of algorithms for the increasing convex order ($\preceq_{icx}$) has precluded to compare the accuracy of these orderings when we bound Markov chains. However, $\preceq_{icx}$-ordering provides more accurate bounds when we compare random variables, as $\preceq_{st}$-comparison implies $\preceq_{icx}$-comparison.

Another completely different application of bounds was recently proposed by P. Buchholz [7]. The main assumption is that the modelers do not know the real transi-

*This work was supported by project Sure-Paths from ACI Sécurité and EURO-NGI network of excellence.

[†]PRiSM, Université de Versailles Saint-Quentin-en-Yvelines, 45, Av. des Etats-Unis, 78035 Versailles, France. Email: {mobe, abusic, jmf, nih}@prism.uvsq.fr

[‡]Université Mohammed V, B.P 1014, Rabat, Maroc

[§]Centre Marin Mersenne, Université Paris 1, 75013 Paris, France

tion probabilities. Thus, one wants to model a system by a family of Markov chains where the transition probabilities belong to an interval of probabilities. One has to derive the worst case (or the best case) for all the matrices in the set. The theoretical arguments rely on Courtois's polyhedral approach. The algorithms are very accurate as the bounds can be reached by a matrix in the set. Unfortunately the complexity is quite high. Very recently a similar problem was solved independently by Haddad and Moreaux [18]. Again, one has to find a bound for a set of matrices. However, Haddad and Moreaux's approach is quite different and relies on stochastic comparison with $\preceq_{st}$ order. The set is given through componentwise extremal matrices. Note that these matrices are useless for a direct computation as they are not stochastic. The authors derive an algorithm to find an upper bounding monotone matrix for all elements in the set according to the $\preceq_{st}$ order. This algorithm is very simple and its complexity is relatively small (less than quadratic). However the method seems to be less accurate than Buchholz's method. Note that to the best of our knowledge there is no comparison between these two new methods.

The techniques we use in this paper are quite different. Here we improve the theory of $\preceq_{icx}$-ordering for finite discrete time Markov chains. This order is known for a long time but very little was known in the context of Markov chains. Unlike the $\preceq_{st}$ order, this stochastic order imposes difficult constraints for the monotonicity property and, until recently, it was an open problem to build $\preceq_{icx}$-monotone bounding matrices. In [4] an algorithm has been designed to construct $\preceq_{icx}$-monotone bounding matrices that belong to a class of matrices denoted as class $\mathcal{C}$. However the algorithm relies on the specific monotonicity characterization of class $\mathcal{C}$ Markov chains. Here we develop a general algorithm to obtain an $\preceq_{icx}$-monotone upper bound for a given stochastic matrix. We also show that the $\preceq_{icx}$ order is more accurate than the $\preceq_{st}$ order when we derive some worst case stochastic process for which only the expected value and a pattern of nonzero transitions are known. This problem is somehow related to the problem considered by Buchholz, Haddad and Moreaux. We do not assume that the set of Markov chains we must bound is defined by intervals on the elements, but we know the pattern of nonzero transitions and the average, and we build the processes which provides extremal distributions.

The remaining of the paper is organized as follows. In Section 2 we present a brief overview of the comparison of random variables and Markov chains under some stochastic order when the state space is endowed with a total order. We also stress that the stochastic comparison approach is much more versatile than the polyhedral technique developed by Courtois [10, 11] even if it is often less accurate. Then we develop in Section 3 the theory to compare finite discrete time Markov chains (DTMCs) with the $\preceq_{icx}$ order. Section 4 is devoted to an algorithm to build an $\preceq_{icx}$-monotone upper bound of a Markov chain. Finally in Section 5 we present two applications of our approach. The first application consists in deriving the worst case of a family of Markov chains where the transitions are defined by their expectation and a pattern for nonzero transitions. The second one is related to the absorption time of a DTMC with one absorbing state and can be used to bound a phase type (PH) distribution modeling a general service time. By doing so, the complexity in two level modeling formalisms can be significantly reduced.

2. A brief presentation. Here we state some basic definitions and results on the stochastic comparison approach. We refer to [24, 26] for further details. First we give the definition of stochastic comparison of two random variables taking values on a totally ordered space $\mathcal{E}$. Let $\mathcal{F}_{st}$ denote the class of all increasing real functions on

$\mathcal{E}$ and $\mathcal{F}_{icx}$ the class of all increasing and convex real functions on $\mathcal{E}$. We denote by $\preceq_{\mathcal{F}}$ the stochastic order relation, where $\mathcal{F}$ can be replaced by st, icx to be associated respectively with the class of functions $\mathcal{F}_{st}$, $\mathcal{F}_{icx}$. Throughout the paper, $\leq$ denotes the componentwise comparison when comparing two vectors or matrices.

DEFINITION 2.1. *Let X and Y be two random variables taking values on a totally ordered space $\mathcal{E}$,*

$$X \preceq_{\mathcal{F}} Y \iff Ef(X) \leq Ef(Y), \quad \forall f \in \mathcal{F}$$

whenever the expectations exist.

Remark that, since $\mathcal{F}_{icx} \subset \mathcal{F}_{st}$, the $\preceq_{st}$-comparison is stronger than the $\preceq_{icx}$-comparison, i.e.

$$X \preceq_{st} Y \quad \Longrightarrow \quad X \preceq_{icx} Y.$$

Notice that for discrete random variables X and Y with probability vectors p and q, the notations $p \preceq_{\mathcal{F}} q$ and $X \preceq_{\mathcal{F}} Y$ are used interchangeably. Stochastic comparison of discrete random variables according to $\preceq_{st}$ and $\preceq_{icx}$ orders can also be defined through matrices (see [20, 22]). We assume here that $\mathcal{E} = \{1, \ldots, n\}$, but the following statements may be extended to the infinite case. We denote by $\boldsymbol{K}_{\mathcal{F}}$ the matrix related to the $\preceq_{\mathcal{F}}$ order, $\mathcal{F} \in \{st, icx\}$:

$$\boldsymbol{K}_{st} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 1 & 1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \ldots & 1 \end{pmatrix}, \quad \boldsymbol{K}_{icx} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 2 & 1 & 0 & \ldots & 0 \\ 3 & 2 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n & n-1 & n-2 & \ldots & 1 \end{pmatrix}. \tag{2.1}$$

PROPOSITION 2.2. *Let X and Y be two random variables with probability vectors $p = (p_i)_{i=1}^n$ and $q = (q_i)_{i=1}^n$ ($p_i = P(X = i)$ and $q_i = P(Y = i)$, $1 \leq i \leq n$). Then*

$$X \preceq_{\mathcal{F}} Y \iff p\boldsymbol{K}_{\mathcal{F}} \leq q\boldsymbol{K}_{\mathcal{F}}.$$

For the $\preceq_{st}$ and $\preceq_{icx}$ orders, this can be given as follows:

$$X \preceq_{st} Y \iff \sum_{k=i}^{n} p_k \leq \sum_{k=i}^{n} q_k, \quad \forall i \in \{1, \ldots, n\}, \tag{2.2}$$

$$X \preceq_{icx} Y \iff \sum_{k=i}^{n} (k-i+1)\ p_k \leq \sum_{k=i}^{n} (k-i+1)\ q_k, \quad \forall i \in \{1, \ldots, n\}. \tag{2.3}$$

In the following we only compare discrete time Markov chains (DTMCs). Continuous time models can be considered after uniformization.

It is shown in Theorem 5.2.11 of [24, p. 186] that monotonicity and comparability of the probability transition matrices of time-homogeneous Markov chains yield sufficient conditions to compare stochastically the underlying chains. We first define the monotonicity and comparability of stochastic matrices and then state this fundamental theorem.

DEFINITION 2.3. *Let $\boldsymbol{P}$ be a stochastic matrix. $\boldsymbol{P}$ is said to be stochastically $\preceq_{\mathcal{F}}$-monotone if for any probability vectors p and q,*

$$p \leq_{\mathcal{F}} q \implies p\boldsymbol{P} \leq_{\mathcal{F}} q\boldsymbol{P}.$$

DEFINITION 2.4. *Let $\boldsymbol{P}$ and $\boldsymbol{Q}$ be two stochastic matrices. $\boldsymbol{Q}$ is said to be an upper bounding matrix of $\boldsymbol{P}$ in the sense of the $\preceq_{\mathcal{F}}$ order ($\boldsymbol{P} \preceq_{\mathcal{F}} \boldsymbol{Q}$) if*

$$\boldsymbol{P}\ \boldsymbol{K}_{\mathcal{F}} \leq \boldsymbol{Q}\ \boldsymbol{K}_{\mathcal{F}}.$$

Let us remark that this is equivalent to saying that $\boldsymbol{P} \preceq_{\mathcal{F}} \boldsymbol{Q}$, if

$$P_{i,*} \preceq_{\mathcal{F}} Q_{i,*}, \quad \forall i \in \{1, \ldots, n\}$$

where $P_{i,*}$ denotes the i^{th} row of matrix $\boldsymbol{P}$.

DEFINITION 2.5. *Let $\{X_k\}_{k\geq 0}$ and $\{Y_k\}_{k\geq 0}$ be two homogeneous DTMCs. We say that the chain $\{X_k\}$ is smaller than the chain $\{Y_k\}$ in the sense of $\preceq_{\mathcal{F}}$ order,*

$$\{X_k\} \preceq_{\mathcal{F}} \{Y_k\}$$

if

$$X_k \preceq_{\mathcal{F}} Y_k \text{ for all } k \geq 0.$$

THEOREM 2.6. *Two homogeneous Markov chains $\{X_k\}_{k\geq 0}$ and $\{Y_k\}_{k\geq 0}$ with the transition matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ satisfy $\{X_k\} \preceq_{\mathcal{F}} \{Y_k\}$ if*

(i) *$X_0 \preceq_{\mathcal{F}} Y_0$,*

(ii) *there exists an $\preceq_{\mathcal{F}}$-monotone transition matrix $\boldsymbol{R}$ such that*

$$\boldsymbol{P} \preceq_{\mathcal{F}} \boldsymbol{R} \preceq_{\mathcal{F}} \boldsymbol{Q}.$$

A special case of Theorem 2.6 is the comparison of two chains when at least one of two transition matrices $\boldsymbol{P}$ or $\boldsymbol{Q}$ is $\preceq_{\mathcal{F}}$-monotone.

COROLLARY 2.7. *Let $\{X_k\}_{k\geq 0}$ and $\{Y_k\}_{k\geq 0}$ be two homogeneous DTMCs with the transition matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$. If*

(i) *$X_0 \preceq_{\mathcal{F}} Y_0$,*

(ii) *$\boldsymbol{P} \preceq_{\mathcal{F}} \boldsymbol{Q}$,*

(iii) *at least one of the transition matrices $\boldsymbol{P}$ or $\boldsymbol{Q}$ is $\preceq_{\mathcal{F}}$-monotone,*

then

$$\{X_k\} \preceq_{\mathcal{F}} \{Y_k\}.$$

COROLLARY 2.8. *Let $\boldsymbol{P}$ and $\boldsymbol{Q}$ be two transition matrices such that there exists an $\preceq_{\mathcal{F}}$-monotone matrix $\boldsymbol{R}$ satisfying $\boldsymbol{P} \preceq_{\mathcal{F}} \boldsymbol{R} \preceq_{\mathcal{F}} \boldsymbol{Q}$. If the steady-state distributions ($\pi_{\boldsymbol{P}}$, $\pi_{\boldsymbol{Q}}$) exist, then*

$$\pi_{\boldsymbol{P}} \preceq_{\mathcal{F}} \pi_{\boldsymbol{Q}}.$$

Let us now consider the comparison of two absorbing chains. The comparison of two chains in the sense of $\preceq_{\mathcal{F}}$ order (we remind that $\preceq_{\mathcal{F}}$ stands for $\preceq_{st}$ or $\preceq_{icx}$ order) provides $\preceq_{st}$-comparison of their absorption times.

PROPOSITION 2.9. *Let $\{X_k\}_{k\geq 0}$ and $\{Y_k\}_{k\geq 0}$ be two homogeneous Markov chains with an absorbing state n (the last one), and let $T_a(X)$ and $T_a(Y)$ denote respectively the absorption times into n for the two chains. If*

$$\{X_k\} \preceq_{\mathcal{F}} \{Y_k\},$$

then

$$T_a(Y) \preceq_{st} T_a(X).$$

Notice that the $\preceq_{st}$-comparison of absorption times is now on random variables T_a defined on $\mathbb{N}_0$ (dates), and not on states.

Proof. We have $X_k \preceq_{\mathcal{F}} Y_k$, for all $k \geq 0$. Particularly, for both $\preceq_{st}$ and $\preceq_{icx}$ order, this implies that

$$P(X_k = n) \leq P(Y_k = n), \text{ for all } k \geq 0. \tag{2.4}$$

Thus, $P(T_a(X) \leq k) = P(X_k = n) \leq P(Y_k = n) = P(T_a(Y) \leq k)$, which gives $T_a(Y) \preceq_{st} T_a(X)$. ∎

Remark that we obtain the $\preceq_{st}$-comparison of absorption times even if we compare the DTMCs in the $\preceq_{icx}$-ordering sense. In fact, the ordering relation needs only to satisfy (2.4).

Like the polyhedral approach, the stochastic comparison approach enables the comparison of steady-state distributions. The Courtois's approach is often more accurate because the polyhedral constraints are usually weaker than the monotonicity constraints. But the stochastic comparison approach is much more versatile. It can provide bounds on the distribution at any time time t, it can give a stochastic bound for the absorption time and for several measures of interest in performance evaluation, reliability modeling, stochastic model checking. For instance, it has been shown that path properties which are studied by stochastic model checking can be simplified using the stochastic comparison approach [5].

3. Increasing convex ordering of finite DTMCs. The monotonicity is an important property for comparison of Markov chains (Theorem 2.6). However, it is obvious that checking the monotonicity of a transition matrix using Definition 2.3 is not tractable since we must check the implication for all comparable probability vectors. Thus, an algorithmic characterization of monotonicity is mandatory.

For the usual stochastic order $\preceq_{st}$, a matrix characterization of monotonicity has been established [20, 24]: $\boldsymbol{P}$ is $\preceq_{st}$-monotone if and only if all the entries of the matrix $\boldsymbol{K}_{st}^{-1}\boldsymbol{P}\boldsymbol{K}_{st}$ are non negative. This is valid for both finite and infinite state space cases. In [22] an equivalent characterization for the $\preceq_{icx}$-monotonicity has been provided when the state space is infinite (for the chains taking values in $\mathbb{Z}$): $\boldsymbol{P}$ is $\preceq_{icx}$-monotone if and only if all the entries of the matrix $\boldsymbol{K}_{icx}^{-1}\boldsymbol{P}\boldsymbol{K}_{icx}$ are non negative. The finite case has been first studied by Vincent in his pioneering work and the above condition was assumed to be sufficient [1]. Note that if the condition is necessary and sufficient for infinite state space, it is only sufficient for finite state space.

In this section we complete Vincent's work and we obtain a complete matrix characterization of the $\preceq_{icx}$-monotonicity. First, we state in the following proposition a necessary and sufficient condition for the $\preceq_{icx}$-monotonicity in terms of transition probabilities $P_{i,j}$. For a transition matrix $\boldsymbol{P}$ on the state space $\mathcal{E} = \{1, \ldots, n\}$, for all i and j, we will denote by $f_{i,j}(\boldsymbol{P})$,

$$f_{i,j}(\boldsymbol{P}) = \sum_{k=j}^{n} (k - j + 1) P_{i,k}. \tag{3.1}$$

Notice that

$$f_{i,j}(\boldsymbol{P}) = (\boldsymbol{P}\boldsymbol{K}_{icx})_{i,j}, \quad \forall i, j \in \{1, \ldots, n\}. \tag{3.2}$$

REMARK 3.1. *For a vector $x \in \mathbb{R}^n$ we will write $x \in \mathcal{F}$ if and only if vector x, seen as a real function on $\mathcal{E} = \{1, \dots, n\}$, is in $\mathcal{F}$:*

(i) *$x \in \mathcal{F}_{st}$ if and only if vector x is increasing, i.e.*

$$x \in \mathcal{F}_{st} \Longleftrightarrow x_i \leq x_{i+1}, \forall i \in \{1, \dots, n-1\}.$$

(ii) *$x \in \mathcal{F}_{icx}$ if and only if vector x is increasing and convex, i.e.*

$$x \in \mathcal{F}_{icx} \Longleftrightarrow x_1 \leq x_2 \text{ and } 2x_i \leq x_{i-1} + x_{i+1}, \forall i \in \{2, \dots, n-1\}.$$

Notice that $x_1 \leq x_2$ and $2x_i \leq x_{i-1} + x_{i+1}, 1 < i < n-1$ imply $x_i \leq x_{i+1}$, $\forall i < n$.

PROPOSITION 3.2. *A stochastic matrix $\boldsymbol{P}$ is $\preceq_{icx}$-monotone if and only if, for all $j \in \{2, \dots, n\}$, the vectors*

$$f_{*,j}(\boldsymbol{P}) = (f_{i,j}(\boldsymbol{P}))_{i=1}^n,$$

defined by (3.1) are increasing and convex, i.e.

$$f_{1,j}(\boldsymbol{P}) \leq f_{2,j}(\boldsymbol{P}) \text{ and } 2f_{i,j}(\boldsymbol{P}) \leq f_{i-1,j}(\boldsymbol{P}) + f_{i+1,j}(\boldsymbol{P}),\ \forall i \in \{2, \dots, n-1\}.$$

Proof. We first show the following relation for two given probability vectors p and q

$$p\boldsymbol{P} \preceq_{icx} q\boldsymbol{P} \Longleftrightarrow \sum_{k=1}^n p_k f_{k,j}(\boldsymbol{P}) \leq \sum_{k=1}^n q_k f_{k,j}(\boldsymbol{P}), \quad \forall j. \tag{3.3}$$

Using Proposition 2.2, characterizing the $\preceq_{icx}$-comparison of two probability vectors, and (3.2) we have:

$$\begin{aligned} p\boldsymbol{P} \preceq_{icx} q\boldsymbol{P} &\Longleftrightarrow (p\boldsymbol{P})\boldsymbol{K}_{icx} \leq (q\boldsymbol{P})\boldsymbol{K}_{icx} \Longleftrightarrow p(\boldsymbol{P}\boldsymbol{K}_{icx}) \leq q(\boldsymbol{P}\boldsymbol{K}_{icx}) \\ &\Longleftrightarrow \sum_{k=1}^n p_k f_{k,j}(\boldsymbol{P}) \leq \sum_{k=1}^n q_k f_{k,j}(\boldsymbol{P}), \quad \forall j. \end{aligned}$$

$\Leftarrow$ *Necessary condition:* suppose that vectors $f_{*,j}(\boldsymbol{P})$ are increasing and convex and show that $\boldsymbol{P}$ is $\preceq_{icx}$-monotone. For that let us consider two probability vectors p and q such that $p \preceq_{icx} q$ and prove that $p\boldsymbol{P} \preceq_{icx} q\boldsymbol{P}$. According to Definition 2.1, $p \preceq_{icx} q$ implies that $\sum_{k=1}^n p_k h(k) \leq \sum_{k=1}^n q_k h(k)$, $\forall h \in \mathcal{F}_{icx}$. For each $j \in \mathcal{E}$, let us denote by g_j a function on $\mathcal{E}$ defined by

$$g_j(i) = f_{i,j}(\boldsymbol{P}),\ \forall i.$$

Since $f_{*,j}(\boldsymbol{P})$ are increasing and convex for all $j \geq 2$, functions g_j belong to $\mathcal{F}_{icx}$, $\forall j \geq 2$ by hypothesis of the proposition. As $f_{i,1}(\boldsymbol{P}) = 1 + f_{i,2}(\boldsymbol{P})$, $\forall i$, g_1 belongs also to $\mathcal{F}_{icx}$. Thus, $\sum_{k=1}^n p_k f_{k,j}(\boldsymbol{P}) \leq \sum_{k=1}^n q_k f_{k,j}(\boldsymbol{P})$, $\forall j \in \mathcal{E}$, so $p\boldsymbol{P} \preceq_{icx} q\boldsymbol{P}$ follows from (3.3).

$\Rightarrow$ *Sufficient condition:* Suppose that $\boldsymbol{P}$ is $\preceq_{icx}$-monotone, and show that vectors $f_{*,j}(\boldsymbol{P}) \in \mathcal{F}_{icx}$, for all $j \geq 2$.
Let us define the probability vectors $p^{(i)}$, $1 \leq i \leq n$:

$$p^{(i)} = (p_1^{(i)} = 0, \dots, p_i^{(i)} = 1, \dots, p_n^{(i)} = 0).$$

Using (2.3), it is obvious that $p^{(i)} \preceq_{icx} p^{(i+1)},\ 1 \le i \le n-1$. Matrix $\boldsymbol{P}$ is $\preceq_{icx}$-monotone, thus $p^{(i)}\boldsymbol{P} \preceq_{icx} p^{(i+1)}\boldsymbol{P}$. It follows from (3.3) that

$$f_{i,j}(\boldsymbol{P}) = \sum_{k=1}^{n} p_k^{(i)} f_{k,j}(\boldsymbol{P}) \le \sum_{k=1}^{n} p_k^{(i+1)} f_{k,j}(\boldsymbol{P}) = f_{i+1,j}(\boldsymbol{P}),\ \forall j \in \mathcal{E}.$$

Hence, vectors $f_{*,j}(\boldsymbol{P})$ are increasing.
Let us now define the probability vectors $q^{(i)},\ 2 \le i \le n-1$ as follows:

$$q^{(i)} = (q_1^{(i)} = 0,\ \ldots,\ q_{i-1}^{(i)} = \frac{1}{2},\ q_i^{(i)} = 0,\ q_{i+1}^{(i)} = \frac{1}{2},\ \ldots,\ q_n^{(i)} = 0).$$

It can be easily shown that $p^{(i)} \preceq_{icx} q^{(i)},\ 2 \le i \le n-1$. Thus,

$$f_{i,j}(\boldsymbol{P}) = \sum_{k=1}^{n} p_k^{(i)} f_{k,j}(\boldsymbol{P}) \le \sum_{k=1}^{n} q_k^{(i)} f_{k,j}(\boldsymbol{P}) = \frac{1}{2} f_{i-1,j}(\boldsymbol{P}) + \frac{1}{2} f_{i+1,j}(\boldsymbol{P}),\ \forall j.$$

Therefore, $2f_{i,j}(\boldsymbol{P}) \le f_{i-1,j}(\boldsymbol{P}) + f_{i+1,j}(\boldsymbol{P})$ and vectors $f_{*,j}(\boldsymbol{P})$ are convex for all $j \in \mathcal{E}$. ∎

The above proposition can also be proved from Stoyan's theorem 5.2.3 [24] which provides conditions ensuring monotonicity of transition matrices in the general case of an integral stochastic order $\preceq$, and an arbitrary state space. However, this requires to introduce some notions like maximal generators used in this theorem. For the sake of simplicity and to make the paper self contained, we preferred to give a direct proof based only on the definition of monotonicity.

In the following proposition we give the matrix characterization for the $\preceq_{icx}$-monotonicity.

PROPOSITION 3.3. *A stochastic matrix $\boldsymbol{P}$ is $\preceq_{icx}$-monotone if and only if*

$$\boldsymbol{Z}_{icx}\boldsymbol{P}\boldsymbol{K}_{icx} \ge \boldsymbol{0}, \tag{3.4}$$

where $\boldsymbol{K}_{icx}$ is the matrix given by (2.1) and

$$\boldsymbol{Z}_{icx} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & -2 & 1 \end{pmatrix} \tag{3.5}$$

Proof. Let us denote by $\boldsymbol{A} = (A_{i,j})_{i,j=1}^{n}$ the matrix $\boldsymbol{P}\boldsymbol{K}_{icx}$, then

$$\boldsymbol{Z}_{icx}\boldsymbol{A} \ge 0 \iff \forall\ j \in \{1,\ldots,n\},\ \begin{cases} A_{1,j} \ge 0 \\ -A_{1,j} + A_{2,j} \ge 0 \\ A_{i-2,j} - 2A_{i-1,j} + A_{i,j} \ge 0, \quad \forall\ i \ge 3 \end{cases}$$

Notice that the inequality $A_{1,j} \ge 0$ is always satisfied and that the conditions for $j = 1$ follow from the conditions for $j = 2$ ($A_{i,1} = A_{i,2} + 1$). Since $A_{i,j} = \sum_{k=j}^{n}(k - j+1)P_{i,k} = f_{i,j}(\boldsymbol{P})$, by Proposition 3.2 it follows that $\boldsymbol{P}$ is $\preceq_{icx}$-monotone if and only if (3.4) holds. ∎

Let us give some remarks concerning the matrix characterization of Proposition 3.3. First this characterization is not unique. Indeed, the values of the first row of $\boldsymbol{Z}_{icx}$ can be replaced by any non negative values and several set of values are possible for the first column of $\boldsymbol{K}_{icx}$. However, the values of the other rows of $\boldsymbol{Z}_{icx}$ and columns of $\boldsymbol{K}_{icx}$ are necessary. For instance, in the case where all values of the first column of $\boldsymbol{K}_{icx}$ are replaced by 1, we obtain the matrix $\boldsymbol{K}_{icx}(1)$:

$$\boldsymbol{K}_{icx}(1) = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 2 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & n-1 & n-2 & \cdots & 1 \end{pmatrix}$$

and we have the following characterization:

$$\boldsymbol{P} \text{ is } \preceq_{icx}\text{-monotone} \iff \boldsymbol{K}_{icx}(1)^{-1}\boldsymbol{P}\boldsymbol{K}_{icx}(1) \geq 0.$$

In fact it can be seen that $\boldsymbol{K}_{icx}(1)^{-1} = \boldsymbol{Z}_{icx}$.

Let us emphasize that the matrix $\boldsymbol{K}_{icx}(1)$ allows also to define the $\preceq_{icx}$ order. For two probability vectors p and q defined on $\mathcal{E} = \{1, \ldots, n\}$, it can easily be seen that

$$p\ \boldsymbol{K}_{icx} \leq q\ \boldsymbol{K}_{icx} \iff p\ \boldsymbol{K}_{icx}(1) \leq q\ \boldsymbol{K}_{icx}(1).$$

Indeed the sum of the first two columns of $\boldsymbol{K}_{icx}(1)$ gives the first column of $\boldsymbol{K}_{icx}$.

Contrary to the $\preceq_{st}$ order, in the case of the finite state space $\mathcal{E} = \{1, \ldots, n\}$ the condition $\boldsymbol{K}_{icx}^{-1}\boldsymbol{P}\boldsymbol{K}_{icx} \geq 0$ is not equivalent to the $\preceq_{icx}$-monotonicity of the matrix $\boldsymbol{P}$. We show this by a counter example but we give before $\boldsymbol{K}_{icx}^{-1}$ which is the inverse of the matrix $\boldsymbol{K}_{icx}$ given by (2.1).

$$\boldsymbol{K}_{icx}^{-1} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & -2 & 1 \end{pmatrix}$$

Note that the only difference between the matrices $\boldsymbol{Z}_{icx}$ and $\boldsymbol{K}_{icx}^{-1}$ is the value of the first element of the second row. Let us consider the matrix $\boldsymbol{P}$:

$$\boldsymbol{P} = \begin{pmatrix} 0.5 & 0.1 & 0.4 \\ 0.4 & 0.15 & 0.45 \\ 0.3 & 0.2 & 0.5 \end{pmatrix}$$

we have

$$\boldsymbol{K}_{icx}^{-1}\boldsymbol{P}\boldsymbol{K}_{icx} = \begin{pmatrix} 1.9 & 0.9 & 0.4 \\ -1.75 & -0.75 & -0.35 \\ 0 & 0 & 0 \end{pmatrix} \quad \boldsymbol{Z}_{icx}\boldsymbol{P}\boldsymbol{K}_{icx} = \begin{pmatrix} 1.9 & 0.9 & 0.4 \\ 0.15 & 0.15 & 0.15 \\ 0 & 0 & 0 \end{pmatrix}$$

Hence, the condition $\boldsymbol{K}_{icx}^{-1}\boldsymbol{P}\boldsymbol{K}_{icx} \geq 0$ is not satisfied while $\boldsymbol{P}$ is $\preceq_{icx}$-monotone as $\boldsymbol{Z}_{icx}\boldsymbol{P}\boldsymbol{K}_{icx} \geq 0$. In fact, the condition $\boldsymbol{K}_{icx}^{-1}\boldsymbol{P}\boldsymbol{K}_{icx} \geq 0$ is sufficient but not necessary for the $\preceq_{icx}$-monotonicity.

We want to emphasize here the importance of this counter-example and the matrix characterization of Proposition 3.3. Indeed, the condition $\boldsymbol{K}_{icx}^{-1}\boldsymbol{P}\boldsymbol{K}_{icx} \geq 0$ has very important consequences. It implies that the first and last states are absorbing ($P_{1,1} = P_{n,n} = 1$) as it has been shown in [1]. Fortunately, the condition $\boldsymbol{Z}_{icx}\boldsymbol{P}\boldsymbol{K}_{icx} \geq 0$

given in Proposition 3.3 is weaker than the condition $\boldsymbol{K}_{icx}^{-1}\boldsymbol{P}\boldsymbol{K}_{icx} \geq 0$ and it can be easily proven that $\boldsymbol{K}_{icx}^{-1}\boldsymbol{P}\boldsymbol{K}_{icx} \geq 0 \implies \boldsymbol{Z}_{icx}\boldsymbol{P}\boldsymbol{K}_{icx} \geq 0$. Moreover, as we will see in the following section, it is possible to construct $\preceq_{icx}$ bounding monotone matrices without absorbing states.

In [1], the authors studied the irreducibility of monotone bounding matrices. They state that if all sub-diagonal entries of the matrix K characterizing the considered order are positive such as for $\boldsymbol{K}_{st}$ and $\boldsymbol{K}_{icx}$, then the bounding matrix has only one recurrent class. Let us emphasize that this result is independent on the algorithm of construction of the bounding matrix. We recall this result in the case of the $\preceq_{icx}$ order which is of particular interest for us in this work.

PROPOSITION 3.4. *Let $\boldsymbol{P} = (P_{i,j})_{i,j=1}^n$ be an irreducible stochastic matrix. A stochastic matrix $\boldsymbol{Q} = (Q_{i,j})_{i,j=1}^n$ such that $\boldsymbol{P} \preceq_{icx} \boldsymbol{Q}$ and $\boldsymbol{Q}$ is $\preceq_{icx}$-monotone, has only one recurrent class. This class contains the state n.*

Proof. We show that $\forall i < n,\ \exists j > i$ such that $Q_{i,j} > 0$. Indeed, this implies that for each state $i < n$, there exist a path between state i and state n and consequently there is only one recurrent class which contains necessarily state n.

By contradiction, if we suppose that $\exists i < n$, such that $\forall j > i,\ Q_{i,j} = 0$, then $\sum_{j=i+1}^n (j-i)Q_{i,j} = 0$. Using the $\preceq_{icx}$-monotonicity of $\boldsymbol{Q}$ (Proposition 3.2), we have: $\forall k \leq i, \sum_{j=i+1}^n (j-i)Q_{k,j} = 0$. On the other hand, since $\boldsymbol{P} \preceq_{icx} \boldsymbol{Q}$, then $\forall k \leq i, \sum_{j=i+1}^n (j-i)P_{k,j} \leq \sum_{j=i+1}^n (j-i)Q_{k,j} = 0$. This implies that $\forall k \leq i,\ \forall j > i,\ P_{k,j} = 0$, which means that the set $\{1, \ldots, i\}$ is absorbing. This is impossible because $\boldsymbol{P}$ is irreducible. ∎

Recall that $\preceq_{st}$-comparison implies $\preceq_{icx}$-comparison. However, we cannot compare $\preceq_{st}$ and $\preceq_{icx}$-monotonicity property.

EXAMPLE 3.5. *Let us consider the following two matrices*

$$\boldsymbol{P} = \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.3 & 0.4 \\ 0.1 & 0.4 & 0.5 \end{pmatrix} \text{ and } \boldsymbol{Q} = \begin{pmatrix} 0.2 & 0.5 & 0.3 \\ 0.3 & 0.3 & 0.4 \\ 0.2 & 0.3 & 0.5 \end{pmatrix}.$$

Matrix $\boldsymbol{P}$ is an $\preceq_{st}$-monotone matrix that is not $\preceq_{icx}$-monotone. On the other hand, $\boldsymbol{Q}$ is an example of an $\preceq_{icx}$-monotone matrix that is not $\preceq_{st}$-monotone.

4. Algorithm for an $\preceq_{icx}$ bound. Let us suppose that we have a stochastic matrix $\boldsymbol{P}$. We would like to compute an $\preceq_{icx}$-monotone matrix $\boldsymbol{Q}$ such that $\boldsymbol{P} \preceq_{icx} \boldsymbol{Q}$. Then by Corollary 2.7 we have $\preceq_{icx}$-comparison of underlying chains.

Contrary to the case of $\preceq_{st}$ order (see [1]), in the case of $\preceq_{icx}$ order it is not generally possible to find an optimal $\preceq_{icx}$-monotone upper bound $\boldsymbol{Q}$, i.e. a matrix $\boldsymbol{Q}$ such that

(i) $\boldsymbol{P} \preceq_{icx} \boldsymbol{Q}$,
(ii) $\boldsymbol{Q}$ is $\preceq_{icx}$-monotone, and
(iii) for each $\preceq_{icx}$-monotone transition matrix $\boldsymbol{U}$, $\boldsymbol{P} \preceq_{icx} \boldsymbol{U} \Rightarrow \boldsymbol{Q} \preceq_{icx} \boldsymbol{U}$.

EXAMPLE 4.1. *We will consider the transition matrix*

$$\boldsymbol{P} = \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.3 & 0.4 \\ 0.1 & 0.4 & 0.5 \end{pmatrix}.$$

Let us suppose now that there is an optimal $\preceq_{icx}$-monotone upper bound $\boldsymbol{Q}$ for $\boldsymbol{P}$. Both

$$\bar{\boldsymbol{U}} = \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.3 & 0.4 \\ 0.1 & 0.2 & 0.7 \end{pmatrix} \text{ and } \hat{\boldsymbol{U}} = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.3 & 0.4 \\ 0.1 & 0.4 & 0.5 \end{pmatrix}$$

are $\preceq_{icx}$-monotone upper bounds for matrix $\boldsymbol{P}$. Thus, $\boldsymbol{Q}$ should satisfy $\boldsymbol{Q} \preceq_{icx} \bar{\boldsymbol{U}}$ and $\boldsymbol{Q} \preceq_{icx} \hat{\boldsymbol{U}}$. If we consider only the last column, this implies $Q_{*,3} = (0.1, 0.4, 0.5)$, which is not convex. This is in contradiction with the fact that $\boldsymbol{Q}$ is $\preceq_{icx}$-monotone (see Proposition 3.2).

However, for each matrix $\boldsymbol{P}$ we can find an $\preceq_{icx}$-monotone upper bound. A trivial one is given by

$$Q_{i,j} = \begin{cases} 1, j = n, \\ 0, j < n. \end{cases}$$

In this section we discuss the compatibility of $\preceq_{icx}$-monotonicity and comparison constraints, and we propose an algorithm to derive a non-trivial $\preceq_{icx}$-monotone upper bound for an arbitrary finite transition matrix $\boldsymbol{P}$.

Let us remind that, for a transition matrix $\boldsymbol{P}$ of size n, $f_{i,j}(\boldsymbol{P})$ denotes $f_{i,j}(\boldsymbol{P}) = \sum_{k=j}^{n}(k - j + 1)P_{i,j}, \forall i, j$ (see (3.1)). Similarly we will define $s_{i,j}(\boldsymbol{P})$ as

$$s_{i,j}(\boldsymbol{P}) = \sum_{k=j}^{n} P_{i,j}, \forall i, j.$$

It can be easily shown that

$$f_{i,j}(\boldsymbol{P}) = f_{i,j+1}(\boldsymbol{P}) + s_{i,j}(\boldsymbol{P}),\ \forall i,\ \forall j < n. \tag{4.1}$$

Notice that $Q_{i,j} \geq 0,\ \forall i, j$ if and only if

$$s_{i,n}(\boldsymbol{Q}) \geq 0,\ \forall i \text{ and } s_{i,j}(\boldsymbol{Q}) \geq s_{i,j+1}(\boldsymbol{Q}),\ \forall i, \forall j < n.$$

The constraints on $\boldsymbol{Q}$ can be then given in terms of $s_{i,j}(\boldsymbol{Q})$ and $f_{i,j}(\boldsymbol{Q})$ as follows:

1. Comparison ($\boldsymbol{P} \preceq_{icx} \boldsymbol{Q}$)

$$f_{i,j}(\boldsymbol{Q}) \geq f_{i,j}(\boldsymbol{P}),\ \forall i, \forall j \geq 2. \tag{4.2}$$

2. Monotonicity (see Proposition 3.2)

$$\text{vectors } f_{*,j}(\boldsymbol{Q}) \text{ are increasing and convex for all } j \geq 2. \tag{4.3}$$

3. $\boldsymbol{Q}$ is a stochastic matrix

$$\begin{array}{l} 0 \leq s_{i,n}(\boldsymbol{Q}) \leq 1,\ \forall i, \\ s_{i,j+1}(\boldsymbol{Q}) \leq s_{i,j}(\boldsymbol{Q}) \leq 1,\ \forall i, \forall j < n, \\ s_{i,1}(\boldsymbol{Q}) = 1, \forall i. \end{array} \tag{4.4}$$

We will compute the entries of the bounding matrix decreasingly by columns. Remark that we need to compute $n - 1$ columns as the first one is completely determined by the fact that the matrix $\boldsymbol{Q}$ is stochastic.

The constraints for the last column can be written as follows:

$$\begin{array}{l} Q_{i,n} \geq P_{i,n},\ \forall i \\ Q_{*,n} \text{ is increasing and convex} \\ Q_{n,n} \leq 1. \end{array}$$

Remark that $Q_{*,n} \in \mathcal{F}_{icx}$ and $Q_{n,n} \leq 1$ imply $Q_{i,n} \leq 1, \forall i$. Let $\mathbf{0}, \mathbf{1} \in \mathbb{R}^n$ denote the vectors $\mathbf{0} = (0, \ldots, 0)$ and $\mathbf{1} = (1, \ldots, 1)$. The last column of a bounding matrix $\boldsymbol{Q}$ is then a solution of Problem 4.2 for $a = P_{*,n}$ and $b = \mathbf{1}$. Moreover, we will show in §4.1 (Proposition 4.4) that the computation of each column of an $\preceq_{icx}$-monotone bounding matrix can be seen as Problem 4.2 with different vectors a and b that can be easily computed.

Problem 4.2. *Let a and b be two vectors such that $\mathbf{0} \leq a \leq b$. Find an increasing and convex vector $x \in \mathbb{R}^n$ such that $a \leq x \leq b$.*

Let us remark that Problem 4.2 does not always have a solution. Take for instance $a = b$ and $a_1 > a_2$. In our particular problem of last column computation, $b = \mathbf{1}$ is trivially increasing and convex, so there is at least one solution of Problem 4.2, $x = b$. This trivial solution is not satisfying in our case, as this would result by a bounding matrix $\boldsymbol{Q}$ with all the elements of the last column equal to 1. Intuitively, we are interested in finding a solution x of Problem 4.2 that is as closest as possible to the vector a. Notice that if vector a is increasing and convex, we can simply take $x = a$. If a is not increasing and convex it is generally not possible to find an optimal solution x of Problem 4.2 in the following sense:

$$\text{for each } y, \text{ a solution of Problem 4.2, } x \leq y.$$

Example 4.3. *Consider for example vectors*

$$a = (0.1, 0.4, 0.5) \text{ and } b = (1, 1, 1).$$

The vectors $\bar{y} = (0.1, 0.4, 0.7)$ and $\hat{y} = (0.3, 0.4, 0.5)$ are both solutions of Problem 4.2. Thus the optimal solution x of Problem 4.2 should satisfy $x \leq \bar{y}$ and $x \leq \hat{y}$ which implies $x = a$. This is not possible as vector a is not increasing and convex ($a_3 - a_2 = 0.1 < 0.3 = a_2 - a_1$).

Some heuristics for Problem 4.2 will be given in §4.2.

4.1. Basic algorithm. We will present here an algorithm to construct an $\preceq_{icx}$-monotone bounding matrix for a given transition matrix $\boldsymbol{P}$. If we suppose that we know the last $n - j$ columns of the bounding matrix $\boldsymbol{Q}$, then the sufficient and necessary conditions on column j are given by the following proposition.

Proposition 4.4. *Let us suppose that we have already computed the columns n to $j+1, 1 < j < n$ of matrix $\boldsymbol{Q}$, satisfying conditions (4.2), (4.3) and (4.4). If we fix the last $n - j$ already computed columns of $\boldsymbol{Q}$, then $f_{*,j}(\boldsymbol{Q})$ must be a solution of Problem 4.2 with vectors a and b such that*

$$a_i = \max(f_{i,j}(\boldsymbol{P}), f_{i,j+1}(\boldsymbol{Q}) + s_{i,j+1}(\boldsymbol{Q})),\ b_i = f_{i,j+1}(\boldsymbol{Q}) + 1, \forall i. \tag{4.5}$$

Problem 4.2 with the above vectors a and b always has a solution. Furthermore, each solution of Problem 4.2 with the above vectors a and b can be taken as $f_{*,j}(\boldsymbol{Q})$.

Proof. We will first show that Problem 4.2 with vectors a and b given by (4.5) always has a solution. Notice that vector b given by (4.5) is increasing and convex by the hypothesis of the proposition (condition (4.3)). Therefore, if $a \leq b$, Problem 4.2 always has at least one trivial solution b. By the hypothesis of the proposition we have $f_{i,j+1}(\boldsymbol{P}) \leq f_{i,j+1}(\boldsymbol{Q})$ (condition (4.2)) and $s_{i,j+1}(\boldsymbol{Q}) \leq 1$ (condition (4.4)), so $f_{i,j}(\boldsymbol{P}) = f_{i,j+1}(\boldsymbol{P}) + s_{i,j}(\boldsymbol{P}) \leq f_{i,j+1}(\boldsymbol{Q}) + 1 = b_i, \forall i$ and $f_{i,j+1}(\boldsymbol{Q}) + s_{i,j+1}(\boldsymbol{Q}) \leq b_i, \forall i$. Thus $a \leq b$.

Let x be now any solution of Problem 4.2 with vectors a and b given by (4.5), and let us take this solution as vector $f_{*,j}(\boldsymbol{Q})$. We will show that $f_{*,j}(\boldsymbol{Q})$ satisfies the conditions (4.2), (4.3) and (4.4). In other words, that any solution of Problem 4.2 with vectors a and b given by (4.5) can be taken as $f_{*,j}(\boldsymbol{Q})$. Condition (4.3) follows directly from the fact that any solution of Problem 4.2 is increasing and convex. Condition (4.2) follows from $f_{*,j}(\boldsymbol{Q}) \geq a$. It remains us to show that $s_{i,j+1}(\boldsymbol{Q}) \leq s_{i,j}(\boldsymbol{Q}) \leq 1, \forall i$. We know that $f_{*,j}(\boldsymbol{Q}) \geq a$, thus

$$s_{i,j}(\boldsymbol{Q}) = f_{i,j}(\boldsymbol{Q}) - f_{i,j+1}(\boldsymbol{Q}) \geq s_{i,j+1}(\boldsymbol{Q}), \forall i.$$

On the other hand, $f_{*,j}(\boldsymbol{Q}) \leq b$ implies $s_{i,j}(\boldsymbol{Q}) \leq 1, \forall i$. Therefore, any solution of Problem 4.2 with vectors a and b given by (4.5) can be taken as $f_{*,j}(\boldsymbol{Q})$.

Finally, we will show that this is also a necessary condition for $f_{*,j}(\boldsymbol{Q})$. Let us suppose that $f_{*,j}(\boldsymbol{Q})$ satisfies conditions (4.2), (4.3) and (4.4). Condition (4.2) implies $f_{i,j}(\boldsymbol{Q}) \geq f_{i,j}(\boldsymbol{P}), \forall i$, and condition (4.4) implies $s_{i,j}(\boldsymbol{Q}) \leq s_{i,j+1}(\boldsymbol{Q}), \forall i$. We have

$$f_{i,j}(\boldsymbol{Q}) = f_{i,j+1}(\boldsymbol{Q}) + s_{i,j}(\boldsymbol{Q}) \leq f_{i,j+1}(\boldsymbol{Q}) + s_{i,j+1}(\boldsymbol{Q}), \ \forall i.$$

Thus, $f_{*,j}(\boldsymbol{Q}) \geq a$. On the other hand, (4.4) implies $s_{i,j}(\boldsymbol{Q}) \leq 1, \forall i$ which gives $f_{i,j}(\boldsymbol{Q}) = f_{i,j+1}(\boldsymbol{Q}) + s_{i,j}(\boldsymbol{Q}) \leq b_i, \forall i$. Finally, (4.3) implies that $f_{*,j}(\boldsymbol{Q})$ is increasing and convex. Therefore, $f_{*,j}(\boldsymbol{Q})$ is a solution of Problem 4.2 with vectors a and b given by (4.5). ∎

Proposition 4.4 allows us to build the bounding matrix $\boldsymbol{Q}$ decreasingly by columns. As there is generally no optimal $\preceq_{icx}$-monotone upper bounding matrix (see Example 4.3), we prefer to give first the general algorithm. Algorithm 4.5 computes an $\preceq_{icx}$-monotone upper bounding matrix $\boldsymbol{Q}$ for an arbitrary finite transition matrix $\boldsymbol{P}$. We reduce the problem of computing an $\preceq_{icx}$-monotone upper bound to Problem 4.2. Then, in §4.2 we present some heuristics to solve Problem 4.2.

ALGORITHM 4.5. *Let $\boldsymbol{P}$ be an arbitrary transition matrix of size n. An $\preceq_{icx}$-monotone upper bound $\boldsymbol{Q}$ for matrix $\boldsymbol{P}$ can be obtained as follows:*

1. *Solve Problem 4.2 with $a = P_{*,n}$ and $b = \mathbf{1}$. Let x_n denote the obtained solution. Set $q_{*,n} = s_{*,n}(\boldsymbol{Q}) = f_{*,n}(\boldsymbol{Q}) = x_n$.*
2. *For each $j = n-1$ to 2:*

Solve Problem 4.2 with vectors a and b as in Proposition 4.4, i.e.

$$a_i = \max(f_{i,j}(\boldsymbol{P}), f_{i,j+1}(\boldsymbol{Q}) + s_{i,j+1}(\boldsymbol{Q})),\ b_i = f_{i,j+1}(\boldsymbol{Q}) + 1, \forall i.$$

Denote the solution by x_j.
Set $f_{i,j}(\boldsymbol{Q}) = x_j,\ s_{i,j}(\boldsymbol{Q}) = f_{i,j}(\boldsymbol{Q}) - f_{i,j+1}(\boldsymbol{Q}),\ Q_{i,j} = s_{i,j}(\boldsymbol{Q}) - s_{i,j+1}(\boldsymbol{Q})$.

3. $f_{i,1}(\boldsymbol{Q}) = f_{i,2}(\boldsymbol{Q}) + 1,\ s_{i,1}(\boldsymbol{Q}) = 1,\ q_{i,1} = 1 - s_{i,2}(\boldsymbol{Q})$.

THEOREM 4.6. *Matrix $\boldsymbol{Q}$ obtained by Algorithm 4.5 is an $\preceq_{icx}$-monotone matrix such that $\boldsymbol{P} \preceq_{icx} \boldsymbol{Q}$.*

Proof. Follows directly from definition of Problem 4.2 and from Proposition 4.4 by induction on j. ∎

Notice that Theorem 4.6 does not depend on how Problem 4.2 is solved in Algorithm 4.5.

4.2. Solving Problem 4.2. Let us remark first that in all the $n-1$ instances of Problem 4.2 in Algorithm 4.5, the vector b is increasing and convex. In this case, Problem 4.2 always has a trivial solution $x = b$. We remind that we are interested in finding an increasing and convex vector x that is as close as possible to the vector a. In terms of Algorithm 4.5, this corresponds to the local optimization for a current column. Recall that, generally, a global optimal $\preceq_{icx}$-monotone upper bounding matrix does not exist (Example 4.3). Additionally, we need solutions that can be easily computed. More precisely, we will consider here only the algorithmic constructions with complexity of $O(n)$ for Problem 4.2.

Vector a does not need to be increasing or convex. However, as we have $a \leq b$, and b is an increasing vector, we know that $r(a) \leq b$, where $r(a)$ denotes the vector of local maxima of vector a,

$$r(a)_i = \max_{k \leq i} a_k.$$

In the following, we will suppose that the vector a is increasing. If this is not the case, we can simply take $r(a)$ instead of a. Notice that $r(a)$ can be easily computed as

$$r(a)_1 = a_1,\ r(a)_i = \max(r(a)_{i-1}, a_i),\ i > 1.$$

Therefore, we will only consider Problem 4.2 with an increasing vector a, and an increasing and convex vector b.

Furthermore, we will distinguish a special case where the vector b is a constant vector. We will show that the general case, where vector b is an arbitrary increasing and convex vector, can be reduced to this special case.

PROPOSITION 4.7. *Let a and b be two increasing vectors such that $a \leq b$. Let the vector b be additionally convex. Denote by δ the maximal distance between a and b, and let d be the vector with all the entries equal to δ,*

$$\delta = \max_{1 \leq i \leq n} \{b_i - a_i\}, \quad d = (\delta, \ldots, \delta).$$

If y is a solution of Problem 4.2 with vectors $a + d - b$ and d, then $x = y + b - d$ is a solution of Problem 4.2 with vectors a and b.

Proof. Both b and y are increasing and convex vectors, thus x is increasing and convex. As $a + d - b \leq y \leq d$, we have $a \leq x \leq b$. ◻

In the following we present some heuristics for Problem 4.2, far from being exhaustive. We remind that we focus only on linear time complexity algorithms.

We have the following constraints:

1. $a_i \leq x_i \leq b_i$,
2. $x_2 \leq x_1$, $x_i \geq 2x_{i-1} - x_{i-2}, \forall i \geq 3$,

where a and b are increasing vectors and b is additionally convex.

Forward computation. The very simple idea is to order the above inequalities increasingly in row index and to take equalities instead of inequalities. We obtain

$$\begin{cases} x_1 = a_1, \\ x_2 = a_2, \\ x_i = \max\{2x_{i-1} - x_{i-2},\ a_i\},\ \forall i \geq 3. \end{cases}$$

However, this can yield $x \not\leq b$. For example, for $a = (0.1, 0.5, 0.5, 0.7)$ and $b = (1, 1, 1, 1)$, we obtain $x = (0.1, 0.5, 0.9, 1.3)$. We can notice that we cannot guarantee that vector x will be a solution of Problem 4.2 even in the special case when b is a constant vector.

Backward computation. If we use the same basic idea as above, but we compute the entries in decreasing order, in the case of a constant vector b, x is a solution of Problem 4.2.

$$\begin{cases} x_n = a_n, \\ x_{n-1} = a_{n-1}, \\ x_i = \max\{2x_{i+1} - x_{i+2},\ a_i\},\ \forall i \leq n-2. \end{cases} \tag{4.6}$$

PROPOSITION 4.8. *Let a be an increasing vector and let $b = (\beta, \dots, \beta)$ be a constant vector such that $a \leq b$. Then vector x computed by backward computation (4.6) is a solution of Problem 4.2, i.e. x is an increasing and convex vector such that $a \leq x \leq b$.*

Proof. $x \geq a$ and x is convex are trivial. It remains us to show that $x_i \leq x_{i+1}, \forall i < n$. Then we have $x_i \leq x_n = a_n \leq \beta$, so x is an increasing convex vector such that $a \leq x \leq b$. We will show that $x_i \leq x_{i+1}, \forall i < n$ by induction on i. For $i = n-1$, we have $x_n = a_n \geq a_{n-1} = x_{n-1}$. Let us suppose now that $x_k \leq x_{k+1}, \forall k, i < k < n$. Then, $x_{i+1} - x_i = x_{i+1} - \max\{2x_{i+1} - x_{i+2},\ a_i\} = \min\{x_{i+2} - x_{i+1},\ x_{i+1} - a_i\} \geq 0$, since $x_{i+1} \geq a_{i+1} \geq a_i$. Thus, $x_i \leq x_{i+1}, \forall i < n$. ∎

Notice that, if there are i and j such that $i < j$ and $a_i = a_j$, then backward computation yields $a_k = a_j$ for all $k \leq j$.

EXAMPLE 4.9. *For vectors $a = (0, 0, 0.2, 0.2, 0.45, 0.6, 0.6, 0.9)$ and $b = \mathbf{1}$, the solution of Problem 4.2 obtained by backward computation is $x = (0.6, 0.6, 0.6, 0.6, 0.6, 0.6, 0.6, 0.9)$.*

To avoid this problem, we propose the following heuristic for Problem 4.2.

Modified backward computation. We suppose that a is not convex (otherwise, take simply $x = a$). Particularly, a is not a constant vector. Let us denote by $\{i_1 > 1, \dots, i_s\}$ indices for which vector a strictly increases, i.e.

$$\text{for all } j < s, a_{i_j} < a_{i_{j+1}} \text{ and } a_k = a_{i_j},\ i_j \leq k < i_{j+1},$$
$$a_k = a_{i_s},\ k \geq i_s.$$

ALGORITHM 4.10 (Modified backward computation).

1. *If $i_s = n$, then $x_n = a_n,\ d_s = 1$,*

else,

$$\begin{cases} x_n = b_n - \delta,\ \text{where } 0 \leq \delta \leq b_n - a_{i_s}. \\ d_s = \frac{x_n - a_{i_s}}{n - i_s}, \\ x_k = x_n - (n-k)d_s,\ i_s \leq k < n. \end{cases}$$

2. *For $t = s-1$ to 1,*

$$\begin{cases} d_t = \min\left\{d_{t+1}, \frac{x_{i_{t+1}} - a_{i_t}}{i_{t+1} - i_t}\right\}, \\ x_k = x_{i_{t+1}} - (i_{t+1} - k)d_t,\ i_t \leq k < i_{t+1}. \end{cases}$$

3. *Computation of entries $x_k,\ 1 \leq k < i_1$:*

$$d \leftarrow d_1,$$
$$\text{For } k = i_1 - 1, \dots, 1, \begin{cases} x_k = \max\{x_{k+1} - d, a_1\}, \\ d \leftarrow x_{k+1} - x_k. \end{cases}$$

PROPOSITION 4.11. *Let a be an increasing vector and let $b = (\beta, \dots, \beta)$ be a constant vector such that $a \leq b$. Then vector x computed by Algorithm 4.10 is a solution of Problem 4.2, i.e. x is an increasing and convex vector such that $a \leq x \leq b$.*

Proof. We will first show that $a \leq x \leq b$. For x_n, either $i_s = n$ and $x_n = a_{i_s} = a_n$ or $x_n = \beta - \delta$, where $0 \leq \delta \leq \beta - a_{i_s} = \beta - a_n$. Therefore, $a_n \leq x_n \leq \beta$. If $i_s < n$, then for $x_k,\ i_s \leq k < n$, we have $x_k = x_n - (n-k)d_s \leq x_n$, as $d_s \geq 0$. On the other hand,

$$x_k = x_n - \frac{n-k}{n-i_s}(x_n - a_{i_s}) \geq a_{i_s} = a_k.$$

Thus, $a_k \le x_k \le x_n \le \beta,\ i_s \le k < n$. Similarly, by induction on t, we can show that $d_t \ge 0$, for each t such that $1 \le t \le s-1$. Thus, $x_k \le x_{i_{t+1}},\ i_t \le k < i_{t+1}$. By induction on t, $x_k \le \beta,\ k \ge i_1$. On the other hand,

$$x_k = x_{i_{t+1}} - (i_{t+1} - k)d_t \ge x_{i_{t+1}} - \frac{i_{t+1} - k}{i_{t+1} - i_t}(x_{i_{t+1}} - a_{i_t}) \ge a_{i_t} \ge a_k,\ i_t \le k < i_{t+1}.$$

Thus, $a_k \le x_k \le \beta,\ k \ge i_1$. For $k < i_1$, using the same arguments as in the proof of Proposition 4.8, it can be easily shown that $a_1 = a_k \le x_k \le x_{i_1} \le \beta$.

It remains us to show that vector x is increasing and convex. It can be easily seen that $0 \le d_1 \le \ldots \le d_s$. We have $x_{k+1} - x_k = d_t,\ i_t \le k < i_{t+1}$. Thus, $x_{k+1} - x_k \le x_{k+2} - x_{k+1},\ i_1 \le k \le n$. Using the same arguments as in the proof of Proposition 4.8, it can be shown that $x_{k+1} - x_k \le x_{k+2} - x_{k+1}, k < i_1$, and that $x_2 \ge x_1$. Therefore, vector x is increasing and convex. ∎

EXAMPLE 4.12. *For vectors a and b from Example 4.9, the solution of Problem 4.2, obtained by modified backward computation is vector x given below:*

$$x = (0, 0.075, 0.2, 0.325, 0.45, 0.6, 0.75, 0.9).$$

Remark that heuristics we presented here do not use the fact that b is a convex vector. Therefore, it is not surprising that we can actually guarantee that they always yield a solution for Problem 4.2 only in the case of a constant vector b. Heuristics that guarantee the solution in the case when b is an arbitrary increasing and convex vector should exploit the fact that we can start with an initial solution. This can be either b, or a solution obtained by means of simple heuristics described above and Proposition 4.7. Describing those heuristics is not in the scope of this paper. However, notice that, if we know a solution x of Problem 4.2, then it can be locally improved in the following way. Suppose $\hat{x}_i = x_i - \epsilon$, $\hat{x}_j = x_j,\ j \ne i$. Then,

$$\epsilon \le \min\{x_i - a_i, x_i - x_{i-1}, x_i - 2x_{i-1} + x_{i-2}, x_i - 2x_{i+1} + x_{i+2}\}.$$

For $i = n-1, n$, we have $\epsilon \le \min\{x_i - a_i, x_i - x_{i-1}, x_i - 2x_{i-1} + x_{i-2},\}$, for $i = 2$, $\epsilon \le \min\{x_2 - a_2, x_2 - x_1, x_2 - 2x_3 + x_4\}$, and for $i = 1$, $\epsilon \le \min\{x_1 - a_1, x_1 - 2x_2 + x_3\}$. This local improvement is interesting only for the entries of a solution x where the slope changes, i.e. where

$$x_{i-1} - x_{i-2} < x_i - x_{i-1} \text{ and } x_{i+1} - x_i < x_{i+2} - x_{i+1}.$$

However, this idea can be generalized to take into account the intervals of constant slope.

4.3. Example. We will illustrate Algorithm 4.5 on a small matrix $\boldsymbol{P}$.

$$\boldsymbol{P} = \begin{pmatrix} 0.2 & 0 & 0.4 & 0.4 & 0 \\ 0.1 & 0.5 & 0.2 & 0.1 & 0.1 \\ 0.25 & 0.25 & 0 & 0.3 & 0.2 \\ 0.2 & 0.1 & 0 & 0.3 & 0.4 \\ 0.1 & 0 & 0.35 & 0 & 0.55 \end{pmatrix}$$

We will denote by x^t the transposed vector of x. If we use the backward computation algorithm for the last column, we obtain $(Q_{*,5})^t = (0, 0.1, 0.25, 0.4, 0.55)$. Let us consider now column 4. We have

$$\begin{array}{rcl} a &=& (\max(f_{*,4}(\boldsymbol{P}), f_{*,5}(\boldsymbol{Q}) + s_{*,5}(\boldsymbol{Q})))^t &=& (0.4, 0.3, 0.7, 1.1, 1.1), \\ b &=& (f_{*,5}(\boldsymbol{Q}) + \mathbf{1})^t &=& (1.0, 1.1, 1.25, 1.4, 1.55). \end{array}$$

Thus, in order to compute column 4, we have to solve Problem 4.2 with vectors $r(a) = (0.4, 0.4, 0.7, 1.1, 1.1)$ and b. Note that our vector b is not a constant vector, so Proposition 4.8 does not apply. Indeed, the backward computation yields $x = (1.1, 1.1, 1.1, 1.1, 1.1) \not\leq b$ so x is not a solution. By means of Proposition 4.7, we have $\delta = \max_{1\leq i\leq n}\{b_i - a_i\} = 0.7,\ d = (0.7, \ldots, 0.7)$, and $a+d-b = (0.1, 0, 0.15, 0.4, 0.25)$. Thus, using backward computation for Problem 4.2 with vectors $a + d - b$ and d, we have $y = (0.4, 0.4, 0.4, 0.4, 0.4)$, and

$$(f_{*,4}(\boldsymbol{Q}))^t = y + b - d = (0.7, 0.8, 0.95, 1.1, 1.25).$$

For column 3 we have

$$\begin{array}{lclcl} a & = & (\max(f_{*,3}(\boldsymbol{P}), f_{*,4}(\boldsymbol{Q}) + s_{*,4}(\boldsymbol{Q})))^t & = & (1.4, 1.5, 1.65, 1.8, 2.0), \\ b & = & (f_{*,5}(\boldsymbol{Q}) + \mathbf{1})^t & = & (1.7, 1.8, 1.95, 2.1, 2.25). \end{array}$$

We can notice that a is increasing and convex. Thus, we can take

$$(f_{*,3}(\boldsymbol{Q}))^t = a = (1.4, 1.5, 1.65, 1.8, 2.0).$$

Finally, for column 2 we have

$$a = (\max(f_{*,2}(\boldsymbol{P}), f_{*,3}(\boldsymbol{Q}) + s_{*,3}(\boldsymbol{Q})))^t = (2.1, 2.2, 2.35, 2.6, 2.9),$$

which is increasing and convex. Thus, $f_{*,2}(\boldsymbol{Q})) = a^t$. The bounding matrix $\boldsymbol{Q}$ is given below. The matrix $\boldsymbol{Q'}$ is obtained by Algorithm 4.5 and modified backward computation.

$$\boldsymbol{Q} = \begin{pmatrix} 0.3 & 0 & 0 & 0.7 & 0 \\ 0.3 & 0 & 0 & 0.6 & 0.1 \\ 0.3 & 0 & 0 & 0.45 & 0.25 \\ 0.2 & 0.1 & 0 & 0.3 & 0.4 \\ 0.1 & 0.15 & 0.05 & 0.15 & 0.55 \end{pmatrix} \quad \boldsymbol{Q'} = \begin{pmatrix} 0.2 & 0.1 & 0.2 & 0.5 & 0 \\ 0.3 & 0.1 & 0 & 0.5 & 0.1 \\ 0.25 & 0.1 & 0 & 0.4 & 0.25 \\ 0.2 & 0.1 & 0 & 0.3 & 0.4 \\ 0.15 & 0.1 & 0 & 0.2 & 0.55 \end{pmatrix}$$

The steady-state distributions are respectively

$$\begin{array}{ll} \pi_{\boldsymbol{P}} & = (0.1663, 0.1390, 0.1982, 0.1998, 0.2966), \\ \pi_{\boldsymbol{Q}} & = (0.1955, 0.0872, 0.0172, 0.3553, 0.3447), \\ \pi_{\boldsymbol{Q'}} & = (0.1951, 0.1000, 0.0390, 0.3293, 0.3366). \end{array}$$

with expectations $E(\pi_{\boldsymbol{P}}) = 3.3213$, $E(\pi_{\boldsymbol{Q}}) = 3.5665$, and $E(\pi_{\boldsymbol{Q'}}) = 3.5122$. Notice that $\preceq_{st}$-monotone upper bound obtained by Vincent's algorithm [1] is given by

$$\boldsymbol{R} = \begin{pmatrix} 0.2 & 0 & 0.4 & 0.4 & 0 \\ 0.1 & 0.1 & 0.4 & 0.3 & 0.1 \\ 0.1 & 0.1 & 0.3 & 0.3 & 0.2 \\ 0.1 & 0.1 & 0.1 & 0.3 & 0.4 \\ 0.1 & 0 & 0.2 & 0.15 & 0.55 \end{pmatrix}$$

with $\pi_{\boldsymbol{R}} = (0.1111, 0.0544, 0.2302, 0.2594, 0.3449)$ and $E(\pi_{\boldsymbol{R}}) = 3.6726$. We can notice that, for this example, $E(\pi_{\boldsymbol{Q'}}) < E(\pi_{\boldsymbol{Q}}) < E(\pi_{\boldsymbol{R}})$. Numerical experimentations have shown that this is not always the case.

5. Applications. We will not develop here a complete algorithm to reduce the state space or the complexity of numerical resolution such as the algorithms presented in [15] for the $\preceq_{st}$ order. Indeed, the $\preceq_{st}$-ordering constraints are consistent with ordinary lumpability [28] and with some matrix structures which allow a simpler resolution technique [16]. Clearly, it is more complex to build a monotone upper bound matrix for the $\preceq_{icx}$ order. Thus, it is difficult to generalize the various algorithms presented in [16] which design in only one step a monotone bound simpler to solve.

Instead, we advocate a two step approach to design $\preceq_{icx}$ lumpable bounds. In the first step we obtain using one of the algorithms described in Section 4 an $\preceq_{icx}$-monotone upper bound $\boldsymbol{B}$ of matrix $\boldsymbol{A}$. Then we use a simpler algorithm to design a lumpable $\preceq_{icx}$ bound (say $\boldsymbol{C}$) of $\boldsymbol{B}$. $\boldsymbol{C}$ is not monotone (it may be but it is not enforced by the method). This is a direct consequence of Theorem 2.6. We do not detail here how we can build a lumpable upper bound. Instead, we present two applications. The first one consists in the worst case analysis of models which are not completely specified, while the second one is related to a more traditional use of bounds to reduce the complexity of numerical computation. Formally, in both cases we are interested in finding $\preceq_{st}$ and $\preceq_{icx}$ bounds in a family of distributions.

5.1. Worst case arrivals in a Batch/D/1/N queue. We consider a queue with a single server, finite capacity N, batch arrivals and deterministic service. Let $A = (a_0, \ldots, a_K)$ denote the distribution of the batch arrivals. We assume that we only know the average batch size $\alpha = E(A)$. Note that the average batch size is closely related to the load which is quite simple to measure. Assume that $N >> K$ and $\alpha < 1$. The exact values of a_i $(0 \le i \le K)$ are unknown. A natural question when we analyze the average queue size of such a system is to find the worst batch distribution. We also describe the model by the set of possible nonzero transitions. Obviously, we know the abstract transition matrix of the Markov chain:

$$\boldsymbol{P} = \begin{pmatrix} a_0 & a_1 & \cdots & a_K & 0 & \cdots & 0 \\ a_0 & a_1 & \cdots & a_K & 0 & \cdots & 0 \\ 0 & a_0 & a_1 & \cdots & a_K & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & a_0 & a_1 & \cdots & a_K \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & a_0 & \sum_{i=1}^{K} a_i \end{pmatrix}$$

Let $\mathcal{F}_\alpha$ be the family of all distributions on the space $\mathcal{E} = \{0, \cdots, n\}$ having the same mean α. In the two following properties, we study the existence of a maximal element in the set $\mathcal{F}_\alpha$.

PROPERTY 5.1. *(See [26, Theorem 2.A.9]) The worst case distribution in the sense of the $\preceq_{icx}$ order is given by $q = (\frac{n-\alpha}{n}, 0, \ldots, 0, \frac{\alpha}{n})$, i.e.*

$$q \in \mathcal{F}_\alpha \quad \textit{and} \quad p \preceq_{icx} q,\ \forall p \in \mathcal{F}_\alpha.$$

The non existence of the $\preceq_{st}$ bound in the same set (next property) is simply related to the fact that if $Y \preceq_{st} X$ and $E(X) = E(Y)$ then X and Y have the same distribution.

PROPERTY 5.2. *For the $\preceq_{st}$ order, there is no distribution r satisfying:*

$$r \in \mathcal{F}_\alpha \quad \textit{and} \quad p \preceq_{st} r,\ \forall p \in \mathcal{F}_\alpha$$

We proceed in the following way to obtain matrix $\boldsymbol{B}$ which is an $\preceq_{icx}$-monotone upper bounding matrix for $\boldsymbol{P}$:

(i) First, we obtain an upper bound $\boldsymbol{Q}$ which is not $\preceq_{icx}$-monotone. We apply Proposition 5.1 at each row to construct matrix $\boldsymbol{Q}$. For the sake of brevity, we cannot detail the whole process here (see [9] for a complete derivation). Note that we must adapt Property 5.1 because the pattern of nonzero transitions changes at every row. For rows $i \leq N-K+1$, the probability mass is concentrated in transitions to states $i-1$ and $i-1+K$. Then the bounding distribution is $q = (\bar{b}, 0, \ldots, 0, b)$ where $b = \frac{m}{K}$ and $\bar{b} = 1-b$. For the other rows the probability mass is concentrated in transitions to states $i-1$ and N. When $i \geq N-K+2$, $Q_{i,i-1} = 1-b_i$ and $Q_{i,N} = b_i$ with $b_i = \frac{m}{N-i+1}$.
Matrix $\boldsymbol{Q}$ is not $\preceq_{icx}$-monotone. Indeed the last row is not convex. We can see that $Q_{N-K,N} = 0$, $Q_{N-K+1,N} = \frac{m}{K}$, and $Q_{N-K+2,N} = \frac{m}{K-1}$. Moreover, it is not sufficient to make the last row convex.

(ii) We now apply the polynomial transform $t_\delta(\boldsymbol{Q}) = \delta\boldsymbol{Q} + (1-\delta)\boldsymbol{Id}$, where $\boldsymbol{Id}$ is the identity matrix. This transform does not change the steady-state distribution and it is known to increase the accuracy of the $\preceq_{st}$ bounds when it is used as a preprocessing [12]. Here it allows to move some probability mass to the diagonal elements.

(iii) Then we apply the forward algorithm to the last row of the transformed matrix, $t_\delta(\boldsymbol{Q})$.

(iv) Finally we change some diagonal and sub-diagonal elements to make the matrix $\preceq_{icx}$-monotone and we obtain matrix $\boldsymbol{B}$:

$$\boldsymbol{B} = \begin{cases} B_{0,0} = 1-\delta b & B_{0,K} = \delta b & \\ i = 1, \cdots, N-K+1 : & & \\ B_{i,i-1} = \delta(1-b) & B_{i,i} = 1-\delta & B_{i,i+K-1} = \delta b \\ i = N-K+2, \cdots, N-1 : & & \\ B_{i,i-1} = f & B_{i,i} = e & B_{i,N} = \delta b(i-N+K) \\ B_{N,N-1} = \delta(1-m) & B_{N,N} = 1-\delta+\delta m & \end{cases}$$

where

$$e = 1-\delta+\delta m - \delta m\frac{(i-N+K)(N-i+1)}{K} \text{ and } f = 1-e-\delta b(i-N+K).$$

Let U be the maximum value of $\frac{(i-N+K)(N-i+1)}{K}$. We have proved in [9] the following property:

PROPERTY 5.3. *If $\delta \leq \frac{1}{1+mU}$ then matrix $\boldsymbol{B}$ is irreducible and stochastic. Furthermore, $\boldsymbol{B}$ is $\preceq_{icx}$-monotone and provides an upper bound of the steady state distribution of $\boldsymbol{P}$.*

Matrix $\boldsymbol{B}$ has the same structure as $t_\delta(\boldsymbol{Q})$ and its stationary distribution can be easily computed using an elimination algorithm. This matrix provides an upper bound for the distribution of the population in the queue.

5.2. Absorption time. Several high level modeling approaches combine a hierarchy of submodels. For instance, PEPA nets [17] are based on Petri nets and the descriptions of places and transitions use PEPA, a Stochastic Process Algebra. This is an explicit two level model. For Stochastic Automata Network the hierarchy is implicit [27]: the automata describe local transitions and the interaction between them is modeled by synchronized transitions and functions which are carried by the labeled transitions. This hierarchy is appealing to model complex structures but it is not always useful to solve the model because the classical technique considers the global state space. Note that even if we represent the transition matrix by a tensor representation [27], we consider the global state space during the resolution process. So we want to develop new methods which can use the submodels during the resolution and combine them in an efficient way. Here we present an application of $\preceq_{icx}$ bounds when the model of the low level is an absorbing Markov chain while the first level of the hierarchy exhibits some structural properties.

We just introduce the approach which will be developed in a sequel paper. We assume here that the high level formalism is a precedence graph but the approach can be easily generalized to other formalisms. There are n nodes representing tasks to complete according to the synchronization constraints defined by the arcs of the precedence graph. Each task is modeled by a DTMC with one absorbing state and the service time (holding time) of task i, d_i is the the absorption time of this Markov chain. We assume that the precedence graph has an unique input node (1) and an unique end node (n). The overall completion time is defined as the duration between the beginning of node 1 and the termination of node n. We are interested in computing the distribution of the completion time for the underlying graph. Note that if a return arc from the end node to the beginning node is added, similar techniques can be used to compute the distribution of the cycle time or the throughput of the system.

The service time of each node follows indeed a discrete PHase type (PH) distribution. A discrete PH distribution is defined by the initial distribution (say τ) and the transition probability matrix Q. Let X be the absorption time of the corresponding chain when the initial distribution is τ. Without loss of generality we assume that there exists only one absorbing state which is the last one. We also assume that the initial distribution τ is $(1, 0, \ldots, 0)$. Indeed, a general distribution can be considered by adding an extra state at the beginning. The global description of the model is Markovian but the state space is huge. Typically we must consider a subset of the product of all the PH distributions for task service times. The main idea here consists in the reduction of the complexity of a PH distribution taking into account the properties of both levels in the model.

It is known that precedence graphs exhibit (max,+) linear equations for their sequence of daters (i.e. the triggering instant of a transition) [2]. Let t_i be the completion time of node i, and $P(i)$ be the set of predecessors of i in the precedence graph. Since node i is trigged (executed) as soon as all its predecessors have been completed, its completion time is defined as follows:

$$t_i = d_i + \max_{j \in P(i)} t_j$$

Thus we obtain linear equations with two operators: the addition and the maximum. Remark that in the case $t_1 = d_1$, the overall completion time is t_n. Such linear (max,+) equations have been extensively studied as they allow new types of analytical or numerical methods which are not based on exponential delays or embedded Markov chains. They also allow an important reduction of complexity for Markovian models.

A fundamental property of increasing convex ordering is the compatibility with these operators. Thus if we can build upper bounds on service times of nodes, we obtain upper bounds on the node completion times, so an upper bound on the completion time of the global system. Note that service times of nodes are supposed to be independent.

Property 5.4. *If for some i, $d_i \preceq_{st} m$ (res. $d_i \preceq_{icx} m$), then $t_n \preceq_{st} \tilde{t}_n$ (res. $t_n \preceq_{icx} \tilde{t}_n$), where $\tilde{t}_n$ denotes the completion time of node n in the system where d_i has been replaced by m.*

In our application example, we aim to bound PH type service times of nodes. Let us first define a family of random variables related to a well known set in reliability modeling [3].

Definition 5.5 (Discrete New Better than Used (in Expectation) (DNBU(E))). *Let X_t be an integer valued random variable modeling the residual time of X, given that $X > t$, i.e. $X_t = [X - t | X > t]$. X is said to be [19]:*

(i) *DNBUE if $E(X_t) \le E(X)$ for all t integer,*
(ii) *DNBU if $X_t \preceq_{st} X$ for all t integer.*

Notice that DNBU $\Rightarrow$ DNBUE.

The main result we use is the $\preceq_{icx}$-comparison for any DNBUE random variable with a geometric one.

Property 5.6. *[19] If X is DNBUE of mean m, then X is smaller in the $\preceq_{icx}$ sense than a geometric distributed random variable of mean m.*

Since we only need a one state model to generate a geometric distribution, if we can bound the low level model (PH distributions for service times of nodes) by a DNBUE distribution, the global state space would be largely reduced.

Of course a PH distribution is not in general DNBUE. However it is simple to bound some set of PH distributions by DNBUE ones. We show in the sequel how DNBUE bounds can be constructed for acyclic PH distributions. These distributions have received considerable attention as they are sufficient to approximate general distributions (see [6] for the theory and a fitting algorithm). The following property is a direct consequence of the memoryless property of the geometric distribution.

Property 5.7. *A sum of independent but not necessarily identically distributed geometric distributions is DNBUE. Remark that this distribution is the uniformization of a hypoexponential distribution.*

Property 5.8. *An arbitrary acyclic discrete PH distribution is upper bounded in the $\preceq_{st}$ sense by a PH distribution defined as a sum of independent but not necessarily identically distributed geometric distributions.*

Recall that $\preceq_{st}$-comparison implies $\preceq_{icx}$-comparison. Thus, this property can be combined with Properties 5.6 and 5.7 to derive an upper bound for an acyclic PH distribution. Let us now illustrate how we can algorithmically construct this upper bound. Suppose that the acyclic PH distribution X is given by its upper triangular transition matrix denoted by $\boldsymbol{P}$. We only need to find a lower bounding (in the $\preceq_{st}$ or $\preceq_{icx}$ sense) matrix $\boldsymbol{Q}$ which has nonzero entries only on the diagonal and first upper diagonal. Note that a matrix of this form is always $\preceq_{st}$ and $\preceq_{icx}$-monotone. Let us denote the corresponding PH distribution by Y. Then $X \preceq_{st} Y$ follows directly from Proposition 2.9.

It can be easily shown that the greatest lower bound of the above form for matrix $\boldsymbol{P}$ in the $\preceq_{st}$ (res. $\preceq_{icx}$) sense is the matrix $\boldsymbol{A}$ (res. $\boldsymbol{B}$), where

$$A_{i,i+1} = s_{i,i+1}(\boldsymbol{P}) = \sum_{j=i+1}^{n} P_{i,j}, \quad B_{i,i+1} = \min(f_{i,i+1}(\boldsymbol{P}), 1), \ \forall i. \tag{5.1}$$

Note that for each matrix $\boldsymbol{P}$ we obtain $\boldsymbol{A} \preceq_{st} \boldsymbol{B}$, since $s_{i,i+1}(\boldsymbol{P}) \le \min(f_{i,i+1}(\boldsymbol{P}), 1)$ (see (4.1)). Thus, in this simple case the $\preceq_{icx}$ order provides better bounds.

Let us now illustrate this on a small numerical example. Let $\boldsymbol{P}$ be the matrix of an acyclic PH distribution X,

$$\boldsymbol{P} = \begin{pmatrix} 0.3 & 0.5 & 0.2 \\ 0 & 0.3 & 0.7 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ given by (5.1) are as follows:

$$\boldsymbol{A} = \begin{pmatrix} 0.3 & 0.7 & 0 \\ 0 & 0.3 & 0.7 \\ 0 & 0 & 1 \end{pmatrix}, \ \boldsymbol{B} = \begin{pmatrix} 0.1 & 0.9 & 0 \\ 0 & 0.3 & 0.7 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let Y and Z denote PH distributions (starting at state 1) corresponding to A and B. We have $E(X) = 2.449$, $E(Y) = 2.857$, and $E(Z) = 2.540$. Clearly, the $\preceq_{st}$ bound ($\boldsymbol{A}$) is less accurate than the $\preceq_{icx}$ bound ($\boldsymbol{B}$). Furthermore, from Property 5.6 it follows that $X \preceq_{icx} Geom(p)$, where $1/p = E(Z) = 2.540$.

We can also easily check if an arbitrary PH distribution (starting at state 1) is DNBUE. The following property provides only a sufficient condition.

PROPERTY 5.9. *Let $a_i, i = 1, \ldots, n-1$ denote the mean residual time X_t of PH distribution X (with transition matrix $\boldsymbol{P}$), given that $X_t = i$. Suppose that $P_{i,i} < 1, \forall i$. Then a_i can be found by solving the linear system:*

$$a_i = \frac{1}{1 - P_{i,i}}(1 + \sum_{j \neq i} P_{i,j} a_j).$$

If $a_1 \geq \max_{i>1} a_i$, then X (starting at 1) is DNBUE.

Finally, we can use the following property to build an $\preceq_{st}$-lower bound that is DNBUE for an arbitrary PH distribution and combine this result with Property 5.6 in order to reduce the state space of our model.

PROPERTY 5.10. *An arbitrary PH distribution starting at state 1 with the transition matrix that is $\preceq_{st}$ or $\preceq_{icx}$-monotone is DNBU (and, consequently, DNBUE).* The proof of this property uses similar arguments as the proof of Proposition 2.9.

6. Conclusion. The stochastic comparison has been largely applied in different areas of applied probability. Recently, algorithmic stochastic comparison in the sense of the $\preceq_{st}$ order has been developed for Markovian analysis. In this paper, we have considered theoretical and algorithmic issues of the $\preceq_{icx}$ order. We hope that these will open new horizons for Markov chain analysis approaches. Our aim was not to compare $\preceq_{icx}$ bounds with $\preceq_{st}$ bounds. We found that it is generally not possible to stochastically compare the distributions obtained by both methods. We will see in the future which approach, if any, is the more accurate and we will develop the complexity issues about these new algorithms, especially for sparse matrices. We presented some applications for the worst case analysis which is an important issue in performance evaluation when the complete specification of the underlying model is unknown.

REFERENCES

[1] O. ABU-AMSHA AND J.-M. VINCENT, *An algorithm to bound functionals of Markov chains with large state space*, in 4th INFORMS Conference on Telecommunications, Boca Raton, FL, 1998.

[2] F. BACCELLI, G. COHEN, G. J. OLSDER, AND J.-P. QUADRAT, *Synchronization and Linearity: An Algebra for Discrete Event Systems*, Willey, New York, 1992.

[3] R. E. Barlow and F. Proschan, *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart and Winston, New York, 1975.
[4] M. Ben Mamoun and N. Pekergin, *Closed-form stochastic bounds on the stationary distribution of Markov chains*, Probability in the Engineering and Informational Sciences, 16 (2002), pp. 403–426.
[5] M. Ben Mamoun, N. Pekergin, and Sana Younes, *Model checking of continuous-time Markov chains by closed-form bounding distributions*. Submitted.
[6] A. Bobbio, A. Horvath, M. Scarpa, and M. Telek, *Acyclic discrete phase type distributions: properties and a parameter estimation algorithm*, Performance Evaluation, 54 (2003), pp. 1–32.
[7] P. Buchholz, *An improved method for bounding stationary measures of finite Markov processes*, Performance Evaluation, 62 (2005), pp. 349–365.
[8] A. Busic and J.-M. Fourneau, *A matrix pattern compliant strong stochastic bound*, in 2005 IEEE/IPSJ International Symposium on Applications and the Internet Workshops (SAINT 2005 Workshops), Trento, Italy, IEEE Computer Society, 2005, pp. 256–259.
[9] A. Busic, J.-M. Fourneau, and N. Pekergin, *Worst case analysis of batch arrivals with the icx ordering*. Submitted.
[10] P.-J. Courtois and P. Semal, *Bounds for the positive eigenvectors of nonnegative matrices and for their approximations by decomposition*, Journal of the ACM, 31 (1984), pp. 804–825.
[11] ———, *Computable bounds for conditional steady-state probabilities in large Markov chains and queueing models*, IEEE Journal on Selected Areas in Communications, 4 (1986), pp. 926–937.
[12] T. Dayar, J.-M. Fourneau, and N. Pekergin, *Transforming stochastic matrices for stochastic comparison with the st-order*, RAIRO Operations Research, 37 (2003), pp. 85–97.
[13] M. Doisy, *Comparaison de Processus Markoviens*, PhD thesis, Université de Pau et des Pays de l'Adour, 1992.
[14] J.-M. Fourneau, M. Le Coz, N. Pekergin, and F. Quessette, *An open tool to compute stochastic bounds on steady-state distributions and rewards*, in 11th International Workshop on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS 2003), Orlando, FL, IEEE Computer Society, 2003, p. 219.
[15] J.-M. Fourneau, M. Le Coz, and F. Quessette, *Algorithms for an irreducible and lumpable strong stochastic bound*, Special Issue on the Conference on the Numerical Solution of Markov Chains 2003, Linear Algebra and its Applications, 386 (2004), pp. 167–185.
[16] J.-M. Fourneau and N. Pekergin, *An algorithmic approach to stochastic bounds*, in Performance Evaluation of Complex Systems: Techniques and Tools, Performance 2002, Tutorial Lectures, M. Calzarossa and S. Tucci, eds., vol. 2459 of Lecture Notes in Computer Science, Springer, 2002, pp. 64–88.
[17] S. Gilmore, J. Hillston, L. Kloul, and M. Ribaudo, *PEPA nets: a structured performance modelling formalism*, Performance Evaluation, 54 (2003), pp. 79–104.
[18] S. Haddad and P. Moreaux, *Sub-stochastic matrix analysis for bounds computation: Theoretical results*. To appear in European Journal of Operational Research.
[19] T. Hu, M. Ma, and A. K. Nanda, *Moment inequalities for discrete ageing families*, Communications in Statistics, 32 (2003), pp. 61–90.
[20] J. Keilson and A. Kester, *Monotone matrices and monotone Markov processes*, Stochastic Processes and their Applications, 5 (1977), pp. 231–241.
[21] J. Ledoux and L. Truffet, *Markovian bounds on functions of finite Markov chains*, Advances in Applied Probability, 33 (2001), pp. 505–519.
[22] H. Li and M. Shaked, *Stochastic convexity and concavity of Markov processes*, Mathematics of Operations Research, 19 (1994), pp. 477–493.
[23] K. Mosler and M. Scarsini, *Stochastic Orders and Applications: A Classified Bibliography*, vol. 401 of Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, 1993.
[24] A. Muller and D. Stoyan, *Comparison Methods for Stochastic Models and Risks*, Wiley, New York, NY, 2002.
[25] N. Pekergin, *Stochastic performance bounds by state space reduction*, Performance Evaluation, 36-37 (1999), pp. 1–17.
[26] M. Shaked and J. G. Shantikumar, *Stochastic Orders and their Applications*, Academic Press, San Diego, CA, 1994.
[27] W. J. Stewart, *Introduction to the Numerical Solution of Markov Chains*, Princeton University Press, New Jersey, NJ, 1994.
[28] L. Truffet, *Reduction technique for discrete time Markov chains on totally ordered state space using stochastic comparisons*, Journal of Applied Probability, 37 (2000), pp. 795–806.