Computing Expected Hitting Times for Imprecise Markov Chains

Thomas Krak[0000−0002−6182−7285]
ELIS – FLip, Ghent University, Belgium
Thomas.Krak@UGent.be

Abstract. We present a novel algorithm to solve a non-linear system of equations, whose solution can be interpreted as a tight lower bound on the vector of expected hitting times of a Markov chain whose transition probabilities are only partially specified. We also briefly sketch how this method can be modified to solve a conjugate system of equations that gives rise to the corresponding upper bound. We prove the correctness of our method, and show that it converges to the correct solution in a finite number of steps under mild conditions on the system. We compare the runtime complexity of our method to a previously published method from the literature, and identify conditions under which our novel method is more efficient.

Keywords: Imprecise Markov Chains · Expected Hitting Times · Imprecise Probabilities · Non-Linear Systems · Computational Methods

1 Introduction

We study the non-negative solution of, and in particular a numerical method to solve non-negatively, the (possibly) non-linear system

\[ \hat{h} = \mathbb{I}_{A^c} + \mathbb{I}_{A^c} \cdot T h, \]

where \( A \) is a non-empty strict subset of a non-empty finite set \( \mathcal{X} \), \( A^c := \mathcal{X} \setminus A \), \( \mathbb{R}^\mathcal{X} \) is the set of all maps from \( \mathcal{X} \) to \( \mathbb{R} \), \( \mathbb{I}_{A^c} \in \mathbb{R}^\mathcal{X} \) is the indicator of \( A^c \), defined as \( \mathbb{I}_{A^c}(x) := 1 \) if \( x \in A^c \) and \( \mathbb{I}_{A^c}(x) := 0 \) otherwise, \( \hat{h} \in \mathbb{R}^\mathcal{X} \) is the vector that solves the system, and \( T : \mathbb{R}^\mathcal{X} \rightarrow \mathbb{R}^\mathcal{X} \) is a (possibly) non-linear map that satisfies the coherence conditions [8]

C1. \( T(\alpha f) = \alpha T f \) for all \( f \in \mathbb{R}^\mathcal{X} \) and \( \alpha \in \mathbb{R}_{\geq 0} \); \hspace{1cm} (non-negative homogeneity)
C2. \( T f + T g \leq T(f + g) \) for all \( f, g \in \mathbb{R}^\mathcal{X} \); \hspace{1cm} (super-additivity)
C3. \( \min_{x \in \mathcal{X}} f(x) \leq T f \) for all \( f \in \mathbb{R}^\mathcal{X} \). \hspace{1cm} (lower bounds)

Note that in Equation (1), the operation \( \cdot \) denotes element-wise multiplication, here applied to the functions \( \mathbb{I}_{A^c} \) and \( T h \).

The motivation for this problem comes from the study of Markov chains using the theory of imprecise probabilities [1, 22]. In this setting, \( \mathcal{X} \) is interpreted as

* Supported by H2020-MSCA-ITN-2016 UTOPIAE, grant agreement 722734.
the set of possible states that some dynamical system of interest can be in, and $T$ is the lower transition operator, which represents the (imprecise) probability model for switching from any state $x$ to any state $y$ in a single time step, e.g.

$$[T I_y](x) = P(X_{n+1} = y | X_n = x), \quad (2)$$

where $P(X_{n+1} = y | X_n = x)$ denotes the lower probability, i.e. a tight lower bound on the probability, that the system will be in state $y$ at time $n + 1$, if it is in state $x$ at time $n$. This lower bound is understood to be taken with respect to the set of stochastic processes induced by $T$, and it is this set that is called the corresponding imprecise Markov chain. We refer to [3, 7–9, 12, 13, 15, 18, 20] for further general information on Markov chains with imprecise probabilities.

In this setting, the minimal non-negative solution $h$ to Equation (1) is then such that $h(x)$ can be interpreted as a tight lower bound on the expected number of steps it will take the system to move from the state $x$, to any state contained in $A$. This vector $h$ is called the (lower) expected hitting time of $A$, for the imprecise Markov chain characterised by $T$. A version of this characterisation of the expected hitting times for imprecise Markov chains was first published by De Cooman et al. [6], and was later generalised to Markov chains under various imprecise probabilistic interpretations by Krak et al. [16].

In the special case that $T$ is a linear map, i.e. when $C^2$ is satisfied with equality for all $f, g \in R^X$, then $T$ can be interpreted as a $|X| \times |X|$ matrix $T$, which is called the transition matrix in the context of (classical) Markov chains. This $T$ is row-stochastic, meaning that $\sum_{y \in X} T(x, y) = 1$ for all $x \in X$, and $T(x, y) \geq 0$ for all $x, y \in X$, and encodes the probability to move between states in one step, i.e. $T(x, y) = P(X_{n+1} = y | X_n = x)$, in analogy to Equation (2).

In this case, the minimal non-negative solution $h$ to Equation (1) is well-known to represent the expected hitting time of $A$ for the homogenous Markov chain identified by $T$; see e.g. [19] for details.

Before moving on, let us first consider why the problem is somewhat ill-posed.

**Proposition 1 ([16]).** Let $T : R^X \rightarrow R^X$ be a map that satisfies the coherence conditions $C^2$. Then Equation (1) has a (unique) minimal non-negative solution $h$ in $(R \cup \{+\infty\})^X$, where minimality means that $h(x) \leq h(x)$ for all $x \in X$ and all non-negative $h \in (R \cup \{+\infty\})^X$ that solve Equation (1).

So, under our present assumptions the solution is not necessarily unique—although there is a unique minimal solution—and the solution can be infinite-valued. It can be shown that in particular $h(x) = +\infty$ if and only if moving from $x$ to $A$ in a finite number of steps has upper probability zero. Conjugate to the notion of lower probabilities, this upper probability is a tight upper bound on the probability that some event will happen. To exclude this case from our analysis, we will in the sequel assume that $T$ also satisfies the reachability condition\[1\]

\[R1. \quad \text{for all } x \in A^c, \text{ there is some } n_x \in \mathbb{N} \text{ such that } [T^{n_x} 1_A](x) > 0.\]

\[1\] Here and in what follows, we take $\mathbb{N}$ to be the set of natural numbers without zero.
Verifying whether a given map $T$ satisfies R1 can be done e.g. using the algorithm for checking lower reachability described in [5], simply replacing the map $Q$ from that work with the map $T$. As that author notes, this algorithm takes $O(|X|^3)$ time to verify whether $T$ satisfies R1. We refer to Section 4 below for details on how $T$ may be evaluated in practice.

In any case, we are now ready to state our first main result.

**Proposition 2.** Let $T : \mathbb{R}^X \rightarrow \mathbb{R}^X$ be a map that satisfies the coherence conditions C1-C3 and the reachability condition R1. Then Equation (1) has a unique solution $h$ in $\mathbb{R}^X$, and this solution is non-negative.

The proof of this result requires some setup, and is given in Section 2. For now, we note that the unique solution $h$ in $\mathbb{R}^X$ that Proposition 2 talks about, is also the (unique) minimal non-negative solution in $((\mathbb{R} \cup \{+\infty\})^X).$ Hence in particular, we can apply the interpretation of $h$ as the vector of lower expected hitting times for an imprecise Markov chain, as discussed above. The following formalizes this.

**Corollary 1.** Let $T : \mathbb{R}^X \rightarrow \mathbb{R}^X$ be a map that satisfies the coherence conditions C1-C3. If Equation (1) has a unique solution $h$ in $\mathbb{R}^X$, and if $h$ is non-negative, then $h$ is the minimal non-negative solution of Equation (1) in $((\mathbb{R} \cup \{+\infty\})^X).$

**Proof.** Let $g \in (\mathbb{R} \cup \{+\infty\})^X$ be the (unique) minimal non-negative solution of Equation (1), whose existence is guaranteed by Proposition 1. Then $g(x) \leq h(x)$ for all $x \in X$, because $g$ is minimal and because $h$ is a non-negative solution to Equation (1). Because $h \in \mathbb{R}^X$, this implies that $g(x) \in \mathbb{R}$ for all $x \in X$, whence $g \in \mathbb{R}^X$, and therefore it follows that $g = h$ because $h$ is the unique solution of Equation (1) in $\mathbb{R}^X$. Hence, and because $g$ is the minimal non-negative solution of Equation (1) in $((\mathbb{R} \cup \{+\infty\})^X)$, so is $h$. \hfill \Box

Having established the existence of a unique real-valued solution to the system, our second main result is a numerical method for computing this solution. For comparison, in [16] the authors present an iterative method that can be directly applied as a computational tool—see Proposition 3 in Section 2—but which is only asymptotically exact, and whose runtime scales with $\|h\|_{\infty}$, making it impractical when (some of) the expected hitting times are numerically large. In contrast, the novel method that we present in Section 3 is independent of $\|h\|_{\infty}$, and we show that it converges to the correct solution in a finite number of steps under practically realistic assumptions on $T$ (but ignoring numerical issues with finite-precision implementations).

## 2 Existence of Solutions

The space $\mathbb{R}^X$ is endowed with the supremum norm, i.e. $\|f\| := \|f\|_\infty := \max_{x \in X} |f(x)|$ for all $f \in \mathbb{R}^X$. Mappings $M : f \mapsto Mf$ from $\mathbb{R}^X$ to $\mathbb{R}^X$ receive the induced operator norm $\|M\| := \sup\{\|Mf\| : f \in \mathbb{R}^X, \|f\| \leq 1\}$. Such
a map $M$ is called \textit{bounded} if it maps (norm-)bounded sets to (norm-)bounded sets; if the map is non-negatively homogeneous (i.e. if it satisfies Property \textit{C1}) then it is bounded if and only if $\|M\| < +\infty$. Note that this includes, as a special case, that \textit{linear} maps are bounded if and only if their norm is bounded.

An element $f \in \mathbb{R}^X$ that is identically equal to some $\mu \in \mathbb{R}$, i.e. $f(x) = \mu$ for all $x \in \mathcal{X}$, is simply written as $\mu \in \mathbb{R}^X$. For any $f, g \in \mathbb{R}^X$, we take $f \leq g$ to mean that $f(x) \leq g(x)$ for all $x \in \mathcal{X}$. For any $B \subseteq \mathcal{X}$ we define the indicator $1_B \in \mathbb{R}^X$ of $B$, for all $x \in \mathcal{X}$, as $1_B(x) := 1$ if $x \in B$, and $1_B(x) := 0$, otherwise.

For a given map $T$ that satisfies \textit{C1-C3} we introduce the \textit{conjugate} map $\overline{T}: \mathbb{R}^X \to \mathbb{R}^X$ that is defined, for all $f \in \mathbb{R}^X$, as $\overline{T}f := -T(-f)$. It is easily verified that this map satisfies the (conjugate) coherence conditions

\begin{enumerate}[label=UC\arabic*.,ref=UC\arabic*.,leftmargin=1cm]
\item $\overline{T}(\alpha f) = \alpha \overline{T}f$ for all $f \in \mathbb{R}^X$ and $\alpha \in \mathbb{R}_{\geq 0}$: \textit{(non-negative homogeneity)}
\item $\overline{T}(f + g) \leq \overline{T}f + \overline{T}g$ for all $f, g \in \mathbb{R}^X$: \textit{(sub-additivity)}
\item $\overline{T}f \leq \max_{x \in \mathcal{X}} f(x)$ for all $f \in \mathbb{R}^X$: \textit{(upper bounds)}
\end{enumerate}

This map $\overline{T}$ is known as the \textit{upper} transition operator in the context of imprecise Markov chains, and yields, in analogy to Equation (2), a tight upper bound on the probability of moving between states. Note that the map $\overline{T}$ is a linear map if and only if $Tf = \overline{T}f$ for all $f \in \mathbb{R}^X$. $\overline{T}$ gives rise to the system

$$
\overline{h} = 1_{A^c} + 1_{A^c} \cdot \overline{T} \overline{h},
$$

whose minimal non-negative solution $\overline{h}$ can be interpreted as a tight upper bound on the expected hitting times of a Markov chain with imprecise probabilities. It follows that in particular $h \leq \overline{h}$; we refer to \cite{10} for further details.

The following properties are well-known to hold for maps that satisfy \textit{C1-C3}; we state them here for convenience. Reference \cite{17} happens to contain all of these, but most of the references on imprecise Markov chains state at least some of them. Here and in what follows, for any $n \in \mathbb{N}$, we use $M^n$ to denote the $n$-fold composition of a map $M: \mathbb{R}^X \to \mathbb{R}^X$ with itself.

\begin{lemma}
Let $T : \mathbb{R}^X \to \mathbb{R}^X$ be a map that satisfies \textit{C1-C3} and let $T$ be its conjugate map. Then, for all $f, g \in \mathbb{R}^X$ and all $n \in \mathbb{N}$, it holds that

\begin{enumerate}[label=T\arabic*. ,ref=T\arabic*,leftmargin=1cm]
\item $\min_{x \in \mathcal{X}} f(x) \leq T^n f \leq \max_{x \in \mathcal{X}} f(x)$; \textit{(bounds)}
\item $f \leq g \Rightarrow T^n f \leq T^n g$ \textit{(monotonicity)}
\item $T^n(f + \mu) = T^n f + \mu$ for all constant $\mu \in \mathbb{R}^X$ \textit{(constant additivity)}
\item $\|T^n f - T^n g\| \leq \|f - g\|$ \textit{(non-expansiveness)}
\end{enumerate}

\end{lemma}

\begin{corollary}
Let $T : \mathbb{R}^X \to \mathbb{R}^X$ be a map that satisfies \textit{C1-C3}. Then $T$ is bounded.

\end{corollary}

\begin{proof}
Let $f \in \mathbb{R}^X$ be such that $\|f\| \leq 1$, and fix any $x \in \mathcal{X}$. It then follows from Property \textit{T11} and the definition of $\|f\|$ that $-\|f\| \leq T f(x) \leq \|f\|$; and hence $\|T f(x)\| \leq \|f\| \leq 1$. Because this is true for any $x \in \mathcal{X}$, it follows that $\|T f\| \leq 1$. Because this is true for any $f \in \mathbb{R}^X$ with $\|f\| \leq 1$, it follows from the definition of the operator norm that $\|T\| \leq 1$, whence $T$ is bounded. \hfill \Box
\end{proof}
The following result provides an iterative method to find the minimal non-negative solution(s) to the, in general, non-linear systems (1) and (3).

**Proposition 3 (16).** Let \( T : \mathbb{R}^X \to \mathbb{R}^X \) be a map that satisfies \( C1 \) \( C3 \), and let \( \overline{T} \) be its conjugate map. Define \( \overline{h}_0 := \overline{h}_0 := \mathbb{1}_{A^c} \) and, for all \( n \in \mathbb{N} \), let \( \overline{h}_n := \mathbb{1}_{A^c} + \mathbb{1}_{A^c} \cdot \overline{T} \overline{h}_{n-1} \) and \( \overline{h}_n := \mathbb{1}_{A^c} + \mathbb{1}_{A^c} \cdot \overline{T} \overline{h}_{n-1} \). Then \( \overline{h}_n := \lim_{n \to +\infty} \overline{h}_n \) and \( \overline{h}_n := \lim_{n \to +\infty} \overline{h}_n \) exist in \((\mathbb{R} \cup \{+\infty\})^X\). Moreover, \( \overline{h}_n \) is the minimal non-negative solution to Equation (1), and \( \overline{h}_n \) is the minimal non-negative solution to Equation (4).

**Proof.** This follows from [16, Proposition 10, Theorem 12, and Corollary 13]. \( \square \)

The scheme proposed in this result can also be directly applied as a computational method, the complexity of which we will analyse and compare to our proposed method, in Section 4.

We need the following, somewhat abstract, definition, that is often encountered in the context of Markov chains with imprecise probabilities.

**Definition 1.** Let \( \mathcal{T} \) be a set of maps from \( \mathbb{R}^X \) to \( \mathbb{R}^X \). For any map \( T : \mathbb{R}^X \to \mathbb{R}^X \) and any \( x \in X \), let \( T_x : \mathbb{R}^X \to \mathbb{R} : f \mapsto [Tf](x) \). Then we say that \( \mathcal{T} \) has separately specified rows, if \( \mathcal{T} = \{ T : \mathbb{R}^X \to \mathbb{R}^X \mid \forall x \in X : T_x \in \mathcal{T}_x \} \) where, for all \( x \in X \), \( \mathcal{T}_x := \{ T_x \mid T \in \mathcal{T} \} \).

This definition is needed to obtain the following well-known dual representation of \( \mathcal{T} \), of which we will make extensive use throughout the remainder of this work. Because the result is well-known (see e.g. [3, 13]), we omit the proof here.

**Proposition 4.** Let \( T : \mathbb{R}^X \to \mathbb{R}^X \) be a map that satisfies \( C1 \) \( C3 \). Then there is a unique non-empty, closed, and convex set \( \mathcal{T} \) that has separately specified rows, such that each \( T \in \mathcal{T} \) is a linear map from \( \mathbb{R}^X \) to \( \mathbb{R}^X \) that satisfies \( C1 \) \( C3 \) and such that, for all \( f \in \mathbb{R}^X \), \( \mathcal{T} f = \inf_{T \in \mathcal{T}} Tf \). Moreover, for all \( f \in \mathbb{R}^X \) there is some \( T \in \mathcal{T} \) such that \( \mathcal{T} f = Tf \).

Note that the elements \( T \in \mathcal{T} \) of the dual representation \( \mathcal{T} \) of \( \mathcal{T} \) are linear maps from \( \mathbb{R}^X \) to \( \mathbb{R}^X \) that all satisfy \( C1 \) \( C3 \). Therefore, as noted in Section 1, these maps can be represented as \( |X| \times |X| \) matrices that are row-stochastic.

The dual representation \( \mathcal{T} \) of \( \mathcal{T} \) is also conveniently connected to the conjugate map \( \mathcal{T} \), as follows. We omit the proof, which is straightforward.

**Corollary 3.** Let \( T : \mathbb{R}^X \to \mathbb{R}^X \) be a map that satisfies \( C1 \) \( C3 \), let \( \mathcal{T} \) be its conjugate map, and let \( \mathcal{T} \) be its dual representation. Then for all \( f \in \mathbb{R}^X \) it holds that \( \overline{T} f = \sup_{T \in \mathcal{T}} Tf \) and \( \overline{T} f = Tf \) for some \( T \in \mathcal{T} \).

We note that, because the elements \( T \in \mathcal{T} \) of the dual representation \( \mathcal{T} \) of \( \mathcal{T} \) are linear maps from \( \mathbb{R}^X \) to \( \mathbb{R}^X \) that satisfy \( C1 \) \( C3 \), the minimal non-negative solutions \( h_T \) of the linear system \( h_T = \mathbb{1}_{A^c} + \mathbb{1}_{A^c} \cdot \overline{T} h_T \) exist due to Proposition 3. As established in Section 1, these \( h_T \) can be interpreted as the vectors of expected hitting times of Markov chains parameterised by \( T \in \mathcal{T} \).

The next result is established in [16], and formalises the interpretation of the solutions to the systems (1) and (3) as providing bounds on the expected hitting times for the set of Markov chains induced by the elements \( T \) of \( \mathcal{T} \).
Proposition 5 ([16]). Let $T : \mathbb{R}^X \to \mathbb{R}^X$ be a map that satisfies [C1, C3] let $\bar{T}$ be its conjugate map, and let $\mathcal{T}$ be its dual representation. For any $T \in \mathcal{T}$, let $h_T$ be the minimal non-negative solution of the linear system $h_T = I_{A^c} + I_{A^c} \cdot T h_T$. Let $\underline{h}$ and $\overline{h}$ be the minimal non-negative solutions of the systems [1] and [3], respectively. Then
\[
\underline{h} = \inf_{T \in \mathcal{T}} h_T \quad \text{and} \quad \overline{h} = \sup_{T \in \mathcal{T}} h_T.
\]
Moreover, there is some $T \in \mathcal{T}$ such that $\underline{h} = h_T$, and there is some (possibly different) $T \in \mathcal{T}$ such that $\overline{h} = h_T$.

Proof. This follows from [16, Lemma 8, Theorem 12 and Corollary 13].

We are now ready to begin the analysis that shows that the reachability condition $\mathcal{R}_1$ is a sufficient assumption to establish Proposition 2. We start with the following result, which says that if $\bar{T}$ satisfies the reachability condition $\mathcal{R}_1$, then so do the elements of its dual representation $\mathcal{T}$.

Lemma 2. Let $T : \mathbb{R}^X \to \mathbb{R}^X$ be a map that satisfies [C1, C3] and the reachability condition $\mathcal{R}_1$ and let $\bar{T}$ be its dual representation. Then for all $T \in \mathcal{T}$ and all $x \in A^c$, there is some $n_x \in \mathbb{N}$ such that $[T^{n_x} I_A](x) > 0$; hence $T$ then also satisfies $\mathcal{R}_1$.

Proof. Fix any $T \in \mathcal{T}$ and any $x \in A^c$. Due to $\mathcal{R}_1$ there is some $n_x \in \mathbb{N}$ such that $[T^{n_x} I_A](x) > 0$. Because $T \in \mathcal{T}$, it follows from Proposition 4 that
\[
[T^{n_x + 1} I_A](x) = [T T^{n_x} I_A](x) \geq [T^{n_x + 1} I_A](x) \geq \cdots \geq [T I_A](x) > 0,
\]
where we repeatedly used the monotonicity property $\mathcal{R}_2$. Let us investigate this reachability property for linear maps more fully. We need the following two results.

Lemma 3. Let $T : \mathbb{R}^X \to \mathbb{R}^X$ be a linear map that satisfies [C1, C3] and the reachability condition $\mathcal{R}_1$. Then there is some $x \in A^c$ such that $[T I_A](x) > 0$.

Proof. Suppose ex absurdo that this is false. Since it follows from Property $\mathcal{R}_1$ that $[T^n I_A](x) \geq \min_{y \in A} I_A(y) = 0$, we must have $[T I_A](x) = 0$ for all $x \in A^c$. This provides the induction base for $n = 1$ in the following induction argument: suppose that for some $n \in \mathbb{N}$, $[T^n I_A](x) = 0$ for all $x \in A^c$; we will show that then also $[T^{n+1} I_A](x) = 0$ for all $x \in A^c$.

First, for any $x \in A$ we have $[T^n I_A](x) \leq \max_{y \in X} I_A(y) = 1 = I_A(x)$ due to $\mathcal{R}_1$. Moreover, for any $x \in A^c$ we have $[T^n I_A](x) = 0 = I_A(x)$ by the induction hypothesis. Hence we have $T^n I_A \leq I_A$. It follows that, for any $x \in A^c$,
\[
[T^{n+1} I_A](x) = [T T^n I_A](x) \leq [T I_A](x) = 0,
\]
using the monotonicity of $T$ (Property $\mathcal{R}_2$) for the inequality, and the argument at the beginning of this proof for the final equality. Because it follows from Property $\mathcal{R}_1$ that $[T^{n+1} I_A](x) \geq \min_{y \in A} I_A(y) = 0$, this implies that $[T^{n+1} I_A](x) = 0$ for all $x \in A^c$. This concludes the proof of the induction step.
Hence we have established that, for all \( x \in A^c \), \( [T^n1_A](x) = 0 \) for all \( n \in \mathbb{N} \) which, because \( A^c \) is non-empty, contradicts the assumption that \( T \) satisfies \( R1 \). Hence, our assumption must be false, and there must be some \( x \in A^c \) such that \( [T^n1_A](x) > 0 \).

**Lemma 4.** Let \( T : \mathbb{R}^X \rightarrow \mathbb{R}^X \) be a linear map that satisfies \( \text{CI} \) and the reachability condition \( \text{R1} \). Fix any \( x \in A^c \) and \( n \in \mathbb{N} \) with \( n > 1 \), and suppose that \( [T^n1_A](x) > 0 \) and \( [T^m1_A](x) = 0 \) for all \( m \in \mathbb{N} \) with \( m < n \). Then there is some \( y \in A^c \) such that \( [TI_(y)](x) > 0 \) and \( [T^n-11_A](y) > 0 \).

**Proof.** Because \( 1_A = \sum z \in A 1_{\{z\}} \) and using the linear character of \( T \)—and therefore of \( T^n \)—we have

\[
0 < [T^n1_A](x) = \sum z \in A [T^n1_{\{z\}}](x),
\]

and hence there must be some \( z \in A \) such that \( [T^n1_{\{z\}}](x) > 0 \). Next, we note that for any \( f \in \mathbb{R}^X \) it holds that \( f = \sum y \in X f(y)1_{\{y\}} \). Hence, expanding the product \( T^n \) and using the linearity of \( T \)—and therefore of \( T^n-1 \)—yields

\[
[T^n1_{\{z\}}](x) = [T(T^n-11_{\{z\}})](x) = \sum y \in X [T(T^n-11_{\{z\}})(y)1_{\{y\}}](x).
\]

Because \( [T^n1_{\{z\}}](x) > 0 \), there must be some \( y \in X \) such that

\[
0 < [T(T^n-11_{\{z\}})(y)1_{\{y\}}](x) = [T^n-11_{\{z\}}](y)[TI_(y)](x),
\]

using the linearity of \( T \) for the equality. Since both factors in this expression are clearly non-negative due to Property \( \text{R1} \), this implies that \( [TI_(y)](x) > 0 \) and \( [T^n-11_{\{z\}}](y) > 0 \). Because \( z \in A \) we have \( 1_{\{z\}} \leq 1_A \), and hence this last inequality together with Property \( \text{R2} \) implies that

\[
0 < [T^n-11_{\{z\}}](y) \leq [T^n-11_A](y),
\]

so we see that it only remains to show that \( y \in A^c \). Suppose \textit{ex absurdo} that \( y \in A \). Then \( 1_{\{y\}} \leq 1_A \), and because \( [TI_(y)](x) > 0 \), it follows from Property \( \text{R2} \) that

\[
0 < [TI_(y)](x) \leq [TI1_A](x),
\]

which contradicts the assumption that \( [T^m1_A](x) = 0 \) for all \( m < n \). Hence \( y \in A^c \), which concludes the proof. \( \square \)

In the sequel, we will be interested in some results about mappings on subspaces of \( \mathbb{R}^X \). The following definition introduces the required notation.

**Definition 2.** For any \( f \in \mathbb{R}^X \), we denote by \( f|_{A^c} \) the restriction of \( f \) to \( A^c \), i.e., the mapping \( f|_{A^c} : A^c \rightarrow \mathbb{R} : x \mapsto f(x) \), which is an element of \( \mathbb{R}^{A^c} \).

Moreover, let \( M : \mathbb{R}^X \rightarrow \mathbb{R}^X \) be a map. Then we define its restriction \( M|_{A^c} \) to the subspace \( \mathbb{R}^{A^c} \) of \( \mathbb{R}^X \), for all \( f \in \mathbb{R}^{A^c} \), as

\[
M|_{A^c}(f) := (M(g \cdot 1_{A^c}))(g|_{A^c} = f).
\]
The space $\mathbb{R}^{A^c}$ again receives the supremum norm, and maps from $\mathbb{R}^{A^c}$ to $\mathbb{R}^{A^c}$ the induced operator norm. Note that if the map $M$ in the previous definition is a linear map, then also its restriction $M|_{A^c}$ is a linear map. Hence in particular, in that case $M$ can be interpreted as a matrix, and the restriction $M|_{A^c}$ can then be interpreted as the $|A^c| \times |A^c|$ submatrix of $M$ on the coordinates $A^c$. Moreover, we note the following simple property:

**Lemma 5.** Let $M : \mathbb{R}^X \to \mathbb{R}^X$ be a bounded map. Then its restriction $M|_{A^c}$ to $\mathbb{R}^{A^c}$ is also bounded.

**Proof.** Choose any $f \in \mathbb{R}^{A^c}$ with $\|f\| \leq 1$, and let $g \in \mathbb{R}^X$ be such that $g|_{A^c} = f$ and $g(x) = 0$ for all $x \in A$. Then it holds that $g \cdot \mathbb{I}_{A^c} = g$ and $\|g\| = \|f\|$. Hence it follows that

$$\|M|_{A^c} f\| = \|(M \cdot \mathbb{I}_{A^c})|_{A^c}\| = \|(Mg)|_{A^c}\| \leq \|Mg\| \leq \|M\|,$$

where the last inequality used that $\|g\| = \|f\| \leq 1$. \(\square\)

We next need some elementary results from the spectral theory of bounded linear maps, in particular for linear maps from $\mathbb{R}^{A^c}$ to $\mathbb{R}^{A^c}$. The following definition aggregates concepts from \[10\] Chapter 7, Definitions 1.2, 3.1, 3.5).

**Definition 3 (\[10\]).** Let $M : \mathbb{R}^{A^c} \to \mathbb{R}^{A^c}$ be a bounded linear map. The spectrum of $M$ is the set $\sigma(M) := \{\lambda \in \mathbb{C} : (\lambda \mathbb{I} - M) \text{ is not one-to-one}\}$, where $I$ denotes the identity map on $\mathbb{R}^{A^c}$. The spectral radius of $M$ is $\rho(M) := \sup_{\lambda \in \sigma(M)} |\lambda|$. The resolvent $R(\lambda, M)$ of $M$ is defined for all $\lambda \in \mathbb{C} \setminus \sigma(M)$ as $R(\lambda, M) := (\lambda \mathbb{I} - M)^{-1}$.

Because $A^c$ is finite (since $X \supseteq A^c$ is finite), the spectrum $\sigma(M)$ of any bounded linear map from $\mathbb{R}^{A^c}$ to $\mathbb{R}^{A^c}$ is a non-empty and finite set \[10\] Chapter 7, Corollary 1.4]. We will need the following properties of these objects.

**Lemma 6 (\[10\], Chapter 7, Lemma 3.4]).** Let $M : \mathbb{R}^{A^c} \to \mathbb{R}^{A^c}$ be a bounded linear map. Then for all $\lambda \in \mathbb{C}$ with $|\lambda| > \rho(M)$, the series $\sum_{k=0}^{+\infty} M^k$ converges in norm to $R(\lambda, M)$.

The next property is well-known, but we had some trouble finding an easy-to-use reference; hence we prove it explicitly below.

**Corollary 4.** Let $M : \mathbb{R}^{A^c} \to \mathbb{R}^{A^c}$ be a bounded linear map, and suppose that $\rho(M) < 1$. Then $\lim_{n \to +\infty} M^n = 0$.

**Proof.** Because $\rho(M) < 1$, we have $\lim_{n \to +\infty} \|\sum_{k=0}^{n} M^k - R(1, M)\| = 0$ by Lemma 6. Now for any $n \in \mathbb{N}$, it holds that

$$\|M^n\| = \left\| \sum_{k=0}^{n} M^k - \sum_{k=0}^{n-1} M^k \right\| = \left\| \sum_{k=0}^{n} M^k - R(1, M) + R(1, M) - \sum_{k=0}^{n-1} M^k \right\|$$

$$\leq \left\| \sum_{k=0}^{n} M^k - R(1, M) \right\| + \left\| R(1, M) - \sum_{k=0}^{n-1} M^k \right\|.$$

Because both summands on the right-hand side vanish as we take $n$ to $+\infty$, it follows that $\lim_{n \to +\infty} \|M^n\| = 0$, or in other words, that $\lim_{n \to +\infty} M^n = 0$. \(\square\)
Lemma 7 ([10, Chapter 7, Lemma 3.4]). Let $M : \mathbb{R}^{A^c} \to \mathbb{R}^{A^c}$ be a bounded linear map. Then $\rho(M) \leq \|M^n\|^{\frac{1}{n}}$ for all $n \in \mathbb{N}$.

The crucial observation is now that, for a linear map $T$ that satisfies $\textbf{C1}$ and the reachability condition $\textbf{R1}$, its restriction $T|_{A^c}$ to $\mathbb{R}^{A^c}$ is a bounded linear map (due to Corollary 2 and Lemma 3) that for large enough $n \in \mathbb{N}$, as we will see in Corollary 5, satisfies $\|T|_{A^c}^n\| < 1$. We can therefore use Lemma 7 to establish that the spectral radius of $T|_{A^c}$ is less than 1, and we can then apply Lemma 8 and Corollary 4.

Let us establish that these claims are indeed true. We start with the following result, which gives some basic properties of these restrictions.

Lemma 8. Let $T : \mathbb{R}^X \to \mathbb{R}^X$ be a linear map that satisfies $\textbf{C1} \textbf{C2} \textbf{C3}$ and let $T|_{A^c}$ be its restriction to $\mathbb{R}^{A^c}$, as in Definition 2. Then $(T|_{A^c})^nf \leq (T|_{A^c})^ng$ for all $n \in \mathbb{N}$ and all $f, g \in \mathbb{R}^{A^c}$ with $f \leq g$. Moreover, it holds that $0 \leq (T|_{A^c})^n1 \leq 1$ for all $n \in \mathbb{N}$.

Proof. For the first claim, we proceed by induction on $n$. For the induction base, let $f', g' \in \mathbb{R}^X$ be such that $f'(x) := 0$ and $g'(x) := 0$ for all $x \in A$, and $f'(x) := f(x)$ and $g'(x) := g(x)$ for all $x \in A^c$. Then, clearly, $f'|_{A^c} = f$ and $g'|_{A^c} = g$, and $f' \leq g'$ because $f \leq g$. Moreover, it holds that $f' \cdot 1_{A^c} = f'$ and $g' \cdot 1_{A^c} = g'$. It follows from Definition 2 that, for any $x \in A^c$, since $f' \leq g'$, it holds that

$$[T|_{A^c}f](x) = [T(f' \cdot 1_{A^c})](x) = [Tf'](x) \leq [Tg'](x) = [T(g' \cdot 1_{A^c})](x) = [T|_{A^c}g](x),$$

where we used that $T$ satisfies Property $\textbf{I2}$ for the inequality. Because this is true for any $x \in \mathcal{X}$, it follows that $T|_{A^c}f \leq T|_{A^c}g$.

Now suppose that $(T|_{A^c})^nf \leq (T|_{A^c})^ng$ for some $n \in \mathbb{N}$. Then also

$$(T|_{A^c})^{n+1}f = T|_{A^c}((T|_{A^c})^nf) \leq T|_{A^c}((T|_{A^c})^ng) = (T|_{A^c})^{n+1}g,$$

using the argument for the induction base for the inequality.

For the second claim, start by noting that $1_{A^c}|_{A^c} = 1, 1_{A^c} \cdot 1_{A^c} = 1_{A^c}$, and $0 \leq 1_{A^c} \leq 1$. Hence it follows from Definition 2 that, for all $x \in A^c$,

$$0 = [T0](x) \leq [T1_{A^c}](x) = [T|_{A^c}1](x) = [T|_{A^c}1](x) \leq [T1](x) = 1,$$

where we used the linearity of $T$ for the first equality, Property $\textbf{I2}$ for the first inequality, Definition 2 for the second and third equalities, Property $\textbf{I2}$ for the second inequality, and Property $\textbf{I1}$ for the last equality.

Hence we have $0 \leq T|_{A^c}1 \leq 1$. By repeatedly using the monotonicity property established in the first part of this proof, it follows that also for all $n > 1$ we have that

$$(T|_{A^c})^n1 = (T|_{A^c})^{n-1}(T|_{A^c}1) \leq (T|_{A^c})^{n-1}(T|_{A^c}1) \leq \cdots \leq (T|_{A^c}1) \leq 1.$$
To establish that $0 \leq (T|_{A^c})^n 1$ for all $n \in \mathbb{N}$, we note that the case for $n = 1$ was already given above. Now, for any $n > 1$ we have

$$(T|_{A^c})^n 1 = (T|_{A^c})^{n-1}(T|_{A^c}) 1 \geq (T|_{A^c})^{n-1} 0 = 0,$$

where for the inequality we used the property $0 \leq T|_{A^c}$ together with the monotonicity property established in the first part of this proof; and where we used the linearity of $(T|_{A^c})^{n-1}$ for the final equality.

**Lemma 9.** Let $T : \mathbb{R}^X \to \mathbb{R}^X$ be a linear map that satisfies $[C_1, C_3]$ and the reachability condition $[R]$. Then there is some $n \in \mathbb{N}$ such that for all $k \in \mathbb{N}$ with $k \geq n$, it holds that $[(T|_{A^c})^m 1](x) < 1$ for all $x \in A^c$.

**Proof.** Because $T$ satisfies $[C_1]$ for all $x \in A^c$ there is some unique minimal $n_x \in \mathbb{N}$ such that $[T^{n_x \cdot 1}](x) > 0$. Let $n := \max_{x \in A} n_x$. We will show that $[(T|_{A^c})^m 1](x) < 1$ for all $x \in A^c$, for all $k \in \mathbb{N}$ with $k \geq n$.

We proceed by induction, as follows. Fix any $x \in A^c$ with $n_x = 1$; then $[T^{n_x \cdot 1}](x) > 0$. Because $1 = I_A + I_{A^c}$, and since $[T 1](x) = 1$ due to Property $[I]$ it therefore follows from the fact that $T$ is a linear map that

$$1 = [T 1](x) = [T^{n_x \cdot 1}](x) + [T I_{A^c}](x) > [T I_{A^c}](x).$$

With $g := I_{A^c}$ we have $g|_{A^c} = 1$ and $g \cdot I_{A^c} = I_{A^c}$, so it follows from the above together with Definition 2 that

$$[T|_{A^c} 1](x) = [T^{n_x \cdot 1}](x) < 1.$$  

Hence for any $x \in A^c$ with $n_x = 1$ it holds that $[T|_{A^c} 1](x) < 1$. This enables the induction base for $m = 1$ in the following induction argument: suppose that for some $m \in \mathbb{N}$ it holds that $[(T|_{A^c})^m 1](x) < 1$ for all $x \in A^c$ with $n_x \leq m$; we will show that then also $[(T|_{A^c})^{m+1} 1](x) < 1$ for all $x \in A^c$ with $n_x \leq m + 1$.

First consider any $x \in A^c$ with $n_x \leq m$. By Lemma 8 we have that $(T|_{A^c}) 1 \leq 1$ and therefore, that

$$[(T|_{A^c})^{m+1} 1](x) = [(T|_{A^c})^m (T|_{A^c} 1)](x) \leq [(T|_{A^c})^m 1](x) < 1,$$

where we used the monotonicity of $(T|_{A^c})^m$ established in Lemma 8 for the first inequality, and the induction hypothesis that $[(T|_{A^c})^m 1](x) < 1$ since $n_x \leq m$ for the final inequality. Hence we have $[(T|_{A^c})^{m+1} 1](x) \leq [(T|_{A^c})^m 1](x) < 1$ for all $x \in A^c$ with $n_x \leq m$.

Now consider any $x \in A^c$ with $n_x = m + 1$. Because $n_x$ was chosen to be minimal, this means that $[T^{m+1 \cdot 1}](x) > 0$ and $[T^{k \cdot 1}](x) = 0$ for all $k \in \mathbb{N}$ with $k \leq m$. Due to Lemma 11 and since $m + 1 > 1$, this implies that there is some $y \in A^c$ such that $[T^1 y](x) > 0$ and $[T^{m \cdot 1} y](x) > 0$. This last inequality implies that $n_y \leq m$, since $n_y$ was chosen to be minimal. By the induction hypothesis, this implies that $[(T|_{A^c})^m 1](y) < 1$.

Let $g \in \mathbb{R}^X$ be such that $g|_{A^c} = (T|_{A^c})^m 1$, and $g(z) := 0$ for all $z \in A$. Then $g \cdot I_{A^c} = g$, whence it follows from Definition 2 that

$$[(T|_{A^c})^{m+1} 1](x) = [T|_{A^c} ((T|_{A^c})^m 1)](x) = [T(g \cdot I_{A^c})](x) = [T g](x).$$
Next, we note that $g = \sum_{z \in X} g(z) I_{\{z\}}$, and so, using the linear character of $T$, it holds that $[Tg](x) = \sum_{z \in X} g(z) [TI_{\{z\}}](x)$. Focussing on the summand for $z = y$ we find that

$$\sum_{z \in X} g(z) [TI_{\{z\}}](x) = g(y) [TI_{\{y\}}](x) + \sum_{z \in X \setminus \{y\}} g(z) [TI_{\{z\}}](x)$$

$$< [TI_{\{y\}}](x) + \sum_{z \in X \setminus \{y\}} g(z) [TI_{\{z\}}](x)$$

$$\leq [TI_{\{y\}}](x) + \sum_{z \in X \setminus \{y\}} [TI_{\{z\}}](x) = \sum_{z \in X} [TI_{\{z\}}](x) = 1,$$

where the strict inequality holds because, as established above, $[TI_{\{y\}}](x) > 0$ and $g(y) = [(T|_{A^c})^m]_1(y) < 1$; where the second inequality follows because $g(z) = 0$ for all $z \in A$, because $g(z) \leq 1$ for all $z \in A^c$ due to Lemma 8 and because $[TI_{\{z\}}](x) \geq 0$ for all $z \in X$ because $T$ satisfies $[T]$ and where the final equality used that $\sum_{z \in X} I_{\{z\}} = 1$ together with the fact that $T$ is a linear map that satisfies $[T]$.

Hence we find that $[(T|_{A^c})^{n+1}]_1(x) < 1$ for all $x \in A^c$ with $n_x = m + 1$. Because we already established that $[(T|_{A^c})^{m+1}]_1(x) < 1$ for all $x \in A^c$ with $n_x \leq m$, we conclude that, indeed, $[(T|_{A^c})^{m+1}](x) < 1$ for all $x \in A^c$ with $n_x \leq m + 1$. This concludes the proof of the induction step, and hence our induction argument implies that, for any $n \in \mathbb{N}$, $[(T|_{A^c})^n](x) < 1$ for all $x \in A^c$ with $n_x \leq m$. Because $n = \max_{x \in A} n_x$ satisfies $n \geq n_x$ for all $x \in A^c$, it follows that, as claimed, for all $k \geq n$ we have $n_x \leq n \leq k$ and hence $[(T|_{A^c})^k](x) < 1$ for all $x \in A^c$.

**Corollary 5.** Let $T : \mathbb{R}^X \to \mathbb{R}^X$ be a linear map that satisfies $[C1],[C5]$ and the reachability condition $[R1]$. Then there is some $n \in \mathbb{N}$ such that $\| (T|_{A^c})^n \| < 1$.

**Proof.** By Lemma 9 there is some $n \in \mathbb{N}$ such that $[(T|_{A^c})^n](x) < 1$ for all $x \in A^c$. Since $0 \leq (T|_{A^c})^n 1$ due to Lemma 8 this implies that $\| (T|_{A^c})^n \| < 1$. It therefore suffices to show that $\| (T|_{A^c})^n \| \leq \| (T|_{A^c})^{n-1} \|$. Consider any $f \in \mathbb{R}^{A^c}$ such that $\| f \| \leq 1$. Then $-1 \leq f \leq 1$ due to the definition of $\| f \|$. Using the linear character of $T|_{A^c}$ (and therefore of $(T|_{A^c})^n$) it follows that

$$- (T|_{A^c})^n 1 = (T|_{A^c})^n (-1) \leq (T|_{A^c})^n f \leq (T|_{A^c})^n 1,$$

where the inequalities use the monotonicity properties of $(T|_{A^c})^n$ established in Lemma 8. This implies that $\| (T|_{A^c})^n f \| \leq \| (T|_{A^c})^{n-1} \|$. Because this is true for every $f \in \mathbb{R}^{A^c}$ such that $\| f \| \leq 1$, it follows that

$$\| (T|_{A^c})^n \| = \sup \{ \| (T|_{A^c})^n f \| : f \in \mathbb{R}^{A^c}, \| f \| \leq 1 \} \leq \| (T|_{A^c})^{n-1} \|,$$

which concludes the proof. \( \square \)

The next result summarizes the results established above, and will be crucial for the remainder of this paper.
Lemma 10. Let $T : \mathbb{R}^X \to \mathbb{R}^X$ be a map that satisfies $\mathbf{C1}$, $\mathbf{C3}$ and the reachability condition $\mathbf{R1}$ and let $T$ be its dual representation. For any $T \in \mathcal{T}$, it holds that $\lim_{k \to +\infty} (T|_{A^c})^k = 0$ and, moreover, that $(I - T|_{A^c})^{-1} = \sum_{k=0}^{+\infty} (T|_{A^c})^k$, where $I$ is the identity map on $\mathbb{R}^{A^c}$.

Proof. Fix any $T \in \mathcal{T}$. Then it follows from Proposition 3 and Lemma 2 that $T$ is a linear map from $\mathbb{R}^X$ to $\mathbb{R}^X$ that satisfies $\mathbf{C1}$, $\mathbf{C3}$ and $\mathbf{R1}$. By Corollary 5 this implies that there is some $n \in \mathbb{N}$ such that $\|(T|_{A^c})^n\| < 1$, which implies that also $\|(T|_{A^c})^n\|^{\frac{1}{n}} < 1$. Because $T|_{A^c}$ is a linear map (due to Definition 2) that is bounded by Corollary 2 and Lemma 5, Lemma 7 therefore implies that $\rho(T|_{A^c}) \leq \|(T|_{A^c})^n\|^{\frac{1}{n}} < 1$.

By Corollary 4, this implies that $\lim_{k \to +\infty} (T|_{A^c})^k = 0$ and, by Lemma 6, that $\sum_{k=0}^{+\infty} (T|_{A^c})^k = R(1, T|_{A^c}) = (I - T|_{A^c})^{-1}$. \(\square\)

Finally, we need the following result, which identifies the solution of the linear version of the system in Equation (1).

Proposition 6. Let $T : \mathbb{R}^X \to \mathbb{R}^X$ be a map that satisfies $\mathbf{C1}$, $\mathbf{C3}$ and the reachability condition $\mathbf{R1}$ and let $T$ be its dual representation. Then, for all $T \in \mathcal{T}$, there exists a unique solution $h_T \in \mathbb{R}^X$ of the linear system $h_T = I_{A^c} + I_{A^c} \cdot Th_T$. This $h_T$ is non-negative, and in particular, $h_T|_{A} = 0$ and $h_T|_{A^c} = (I - T|_{A^c})^{-1}$.

Proof. Let $h_T$ be a vector such that $h_T|_{A} := 0$ and $h_T|_{A^c} := (I - T|_{A^c})^{-1}$, where we note that $(I - T|_{A^c})^{-1}$ exists by Lemma 10. Then $h_T \in \mathbb{R}^X$; specifically, $h_T(x)$ is finite for all $x \in X$, because $(I - T|_{A^c})^{-1}$ is the inverse of a bounded linear map and thus bounded itself.

It is easily verified that $h_T$ satisfies the linear system of interest; for any $x \in A$ we trivially have

$$h_T(x) = 0 = I_{A^c}(x) + I_{A^c}(x) \cdot [Th_T](x).$$

Conversely, on $A^c$ it holds that $h_T|_{A^c} = (I - T|_{A^c})^{-1}$, so, multiplying both sides with $(I - T|_{A^c})$, we obtain $h_T|_{A^c} = 1 + T|_{A^c}(h_T|_{A^c})$. We note that, because $h_T|_{A} = 0$ it holds that $h_T \cdot I_{A^c} = h_T$, and hence

$$T|_{A^c}(h_T|_{A^c}) = (T(h_T \cdot I_{A^c}))|_{A^c} = (Th_T)|_{A^c},$$

using Definition 2, and so

$$h_T|_{A^c} = 1 + (Th_T)|_{A^c} = I_{A^c} \cdot A^c + I_{A^c} \cdot (Th_T)|_{A^c}.$$

This establishes that $h_T$ solves the linear system of interest.

Let us prove that $h_T$ is non-negative. Let $c := \min_{y \in X} h_T(y)$, and let $y \in X$ be such that $h_T(y) = c$. Then $c \in \mathbb{R}$ because $h_T \in \mathbb{R}^X$. Now suppose ex absurdo that $y \in A^c$. Then, since $h_T - c \geq 0$, and because, as established above, $h_T|_{A^c} = 1 + (Th_T)|_{A^c}$, it follows that

$$h_T|_{A^c} - c = 1 + (Th_T)|_{A^c} - c = 1 + (T(h_T - c))|_{A^c} \geq 1 + (T(0))|_{A^c} = 1,$$
where the second equality used Property 13 and where the inequality used Property 12, which we can do because $T$ satisfies $[C1,C3]$ due to Proposition 4 and where the final equality used the linearity of $T$. Because $y \in A^c$, this implies in particular that $0 = h_T(y) - c \geq 1$. From this contradiction, it follows that $y \in A$.

Since $h_T(x) = 0$ for all $x \in A$, this implies that $c = h_T(y) = 0$, which means that $h_T$ is non-negative.

It remains to establish that $h_T$ is the unique solution in $\mathbb{R}^X$. To this end, let $g \in \mathbb{R}^X$ be any real-valued solution of this system, i.e. suppose that also $g = \mathbb{I}_{A^c} + \mathbb{I}_{A^c} \cdot Tg$. Then some algebra analogous to the above yields that $g|_{A^c} = 0$ and $g|_{A^c} = (I - T|_{A^c})^{-1}1$, thus $g = h_T$, whence the solution is unique. □

We can now prove our first main result, which was stated in Section 1.

Proof of Proposition 2. Let $h$ be the minimal non-negative solution to (1); this solution exists by Proposition 11 although it could be in $(\mathbb{R} \cup \{+\infty\})^X$. Let us first establish that in fact $h \in \mathbb{R}^X$. To this end, we note that by Proposition 3 there is some $T \in \mathcal{T}$ such that $h = h_T$, where $h_T$ is the minimal non-negative solution of the linear system $h_T = \mathbb{I}_{A^c} + \mathbb{I}_{A^c} \cdot Th_T$.

Because $T$ satisfies $[C1,C3]$ and $[R1]$ it follows from Proposition 4 that the linear system $h_T = \mathbb{I}_{A^c} + \mathbb{I}_{A^c} \cdot Th_T$ has a unique solution in $\mathbb{R}^X$ that is non-negative, which equals the minimal non-negative solution $h_T$ due to Corollary 1.

Hence $h_T \in \mathbb{R}^X$, and because $h = h_T$, it follows that also $h \in \mathbb{R}^X$.

To establish that the solution of (1) is unique in $\mathbb{R}^X$, let $h_T$ and $h_S$ be any two solutions of (1) in $\mathbb{R}^X$; without further assumptions, we will show that $h_T \leq h_S$.

First, for any $x \in A$, it holds that $h_T(x) = \mathbb{I}_{A^c}(x) + \mathbb{I}_{A^c}(x) \cdot [T h_T](x) = 0$, and, by the same argument, $h_S(x) = 0$, so $h_T$ and $h_S$ agree on $A$.

Next, due to Proposition 11 there are $T,S \in \mathcal{T}$ such that $Th_T = Th_T$ and $T h_S = Sh_S$. It also follows from Proposition 3 that $Th_T = Th_T \leq Sh_T$, because $S \in \mathcal{T}$. In turn, and using that $h_T(x) = 0$ for all $x \in A$, and that therefore $h_T \cdot \mathbb{I}_{A^c} = h_T$, this implies that

$$T|_{A^c}(h_T|_{A^c}) = (Th_T)|_{A^c} \leq (Sh_T)|_{A^c} = S|_{A^c}(h_T|_{A^c}).$$

Substituting this inequality into the system (1) restricted to $A^c$ yields

$$h_T|_{A^c} = 1 + T|_{A^c}(h_T|_{A^c}) \leq 1 + S|_{A^c}(h_T|_{A^c}).$$

Now note that $S|_{A^c}$ is a monotone map due to Proposition 3 and Lemma 8. Hence, and using the fact that $S|_{A^c}$ is a linear map, expanding this inequality into the right-hand side gives, after $n \in \mathbb{N}$ expansions,

$$h_T|_{A^c} \leq (S|_{A^c})^n(h_T|_{A^c}) + \sum_{k=0}^{n-1} (S|_{A^c})^k1.$$

Taking limits in $n$, and using that $\lim_{n \to +\infty}(S|_{A^c})^n = 0$ and $\sum_{k=0}^{+\infty}(S|_{A^c})^k = (I - S|_{A^c})^{-1}$ due to Lemma 10 it follows that

$$h_T|_{A^c} \leq (I - S|_{A^c})^{-1}1 = h_S|_{A^c},$$
where we used Proposition 6 for the equality, which we can do because, since \( h_S \) satisfies Equation (1) by assumption and since \( \mathcal{T} h_S = S h_S \) by the selection of \( S \in \mathcal{T} \), it holds that \( h_S = \mathbb{I}_{A^c} + \mathbb{I}_{A^c} \cdot S h_S \). Hence we have found that \( h_T|_{A^c} \leq h_S|_{A^c} \). However, we can now repeat the above argument, mutatis mutandis, to establish that also \( h_S|_{A^c} \leq h_T|_{A^c} \). Hence we conclude that \( h_T \) and \( h_S \) agree on \( A^c \). Because we already established that they agree on \( A \), we conclude that \( h_T = h_S \), and that therefore Equation (1) has a unique solution in \( \mathbb{R}^X \). Because \( \underline{h} \) is a solution to Equation (1) in \( \mathbb{R}^X \), we conclude that \( \underline{h} \) must be the unique solution. Because \( \underline{h} \) is non-negative, it follows that the unique solution in \( \mathbb{R}^X \) is non-negative.

\[ \square \]

3 A Computational Method

The computational method that we propose in Proposition 7 works by iterating over specific choices of the extreme points of the dual representation \( \mathcal{T} \) of \( \mathcal{T} \). The next result establishes that these extreme points exist and that \( \mathcal{T}(\cdot) \) obtains its value in an extreme point of \( \mathcal{T} \); this is well-known, so we omit the proof.

**Lemma 11.** Let \( \mathcal{T} : \mathbb{R}^X \to \mathbb{R}^X \) be a map that satisfies \( C1 \) \( C3 \) and let \( \mathcal{T} \) be its dual representation. Then \( \mathcal{T} \) has a non-empty set of extreme points. Moreover, for all \( f \in \mathbb{R}^X \), there is an extreme point \( T \) of \( \mathcal{T} \) such that \( \mathcal{T} f = T f \).

We can now state the second main result of this work.

**Proposition 7.** Let \( \mathcal{T} : \mathbb{R}^X \to \mathbb{R}^X \) be a map that satisfies \( C1 \) \( C3 \) and the reachability condition \( RA \) and let \( \mathcal{T} \) be its dual representation. Let \( T_1 \in \mathcal{T} \) be any extreme point of \( \mathcal{T} \) and, for all \( n \in \mathbb{N} \), let \( h_n \) be the unique solution in \( \mathbb{R}^X \) of the linear system \( h_n = \mathbb{I}_{A^c} + \mathbb{I}_{A^c} \cdot T_n h_n \), and let \( T_{n+1} \) be an extreme point of \( \mathcal{T} \) such that \( \mathcal{T} h_n = T_{n+1} h_n \). Then the sequence \( \{h_n\}_{n \in \mathbb{N}} \) is non-increasing, and its limit \( h_* := \lim_{n \to +\infty} h_n \) is the unique solution of (1) in \( \mathbb{R}^X \).

**Proof.** First, for any \( n \in \mathbb{N} \), the solution \( h_n \in \mathbb{R}^X \) of the linear system \( h_n = \mathbb{I}_{A^c} + \mathbb{I}_{A^c} \cdot T_n h_n \) exists and is non-negative due to Proposition 6 and there is an extreme point \( T_{n+1} \) of \( \mathcal{T} \) that satisfies \( \mathcal{T} h_n = T_{n+1} h_n \) due to Lemma 11. Hence the sequence \( \{h_n\}_{n \in \mathbb{N}} \) is well-defined and bounded below by 0. Due to Proposition 6, it holds that \( h_n(x) = 0 \) for all \( x \in A \) and all \( n \in \mathbb{N} \). Thus the sequence \( \{h_n\}_{n \in \mathbb{N}} \) is trivially non-increasing and convergent on \( A \). We will next establish the same on \( A^c \).

First note that \( h_n \cdot \mathbb{I}_{A^c} = h_n \) for all \( n \in \mathbb{N} \), because \( h_n(x) = 0 \) for all \( x \in A \). Now fix any \( n \in \mathbb{N} \). Then \( T_{n+1} h_n = \mathcal{T} h_n \leq T_n h_n \) because \( T_n \in \mathcal{T} \), and therefore in particular it holds that

\[
T_{n+1}|_{A^c}(h_n|_{A^c}) = (T_{n+1} h_n)|_{A^c} \leq (T_n h_n)|_{A^c} = T_n|_{A^c}(h_n|_{A^c}).
\]

Therefore, and because \( h_n = \mathbb{I}_{A^c} + \mathbb{I}_{A^c} \cdot T_n h_n \), it holds that

\[
h_n|_{A^c} = 1 + T_n|_{A^c}(h_n|_{A^c}) \geq 1 + T_{n+1}|_{A^c}(h_n|_{A^c}).
\]
Now note that $T_{n+1}|A^c$ is a monotone map due to Proposition 4 and Lemma 8. Hence, and using the fact that $T_{n+1}|A^c$ is a linear map, repeatedly expanding this inequality into the right-hand side gives, after $m \in \mathbb{N}$ expansions,

$$h_n|A^c \geq (T_{n+1}|A^c)^m(h_n|A^c) + \sum_{k=0}^{m-1} (T_{n+1}|A^c)^k 1.$$ 

Taking limits in $m$, and using $\lim_{m \to +\infty}(T_{n+1}|A^c)^m = 0$ and $\sum_{k=0}^{+\infty}(T_{n+1}|A^c)^k = (I-T_{n+1}|A^c)^{-1}$ due to Lemma 10 it follows that

$$h_n|A^c \geq (I-T_{n+1}|A^c)^{-1}1 = h_{n+1}|A^c,$$

where we used Proposition 5 for the equality, which we can do because, by construction, $h_{n+1} \in \mathbb{R}^X$ and $h_{n+1} = I_{A^c} + I_{A^c} \cdot T_{n+1}h_{n+1}$. Thus the sequence $\{h_n\}_{n \in \mathbb{N}}$ is non-increasing. Because we know that $h_n \geq 0$ for all $n \in \mathbb{N}$ due to Proposition 6 it follows that the limit $h_* := \lim_{n \to +\infty} h_n$ exists.

Let us now show that $h_*$ solves (1). To this end, fix $n \in \mathbb{N}$ and note that

$$\|h_* - I_{A^c} - I_{A^c} \cdot T h_*\| \leq \|h_* - h_{n+1}\| + \|h_{n+1} - I_{A^c} - I_{A^c} \cdot T h_{n+1}\|$$

$$+ \|I_{A^c} \cdot T h_{n+1} - I_{A^c} \cdot T h_*\|$$

$$= \|h_* - h_{n+1}\| + \|I_{A^c} \cdot T h_{n+1} - I_{A^c} \cdot T h_*\|$$

$$\leq \|h_* - h_{n+1}\| + ||T_{n+1}h_{n+1} - T h_*\|$$

$$\leq \|h_* - h_{n+1}\| + ||T_{n+1}h_{n+1} - T h_*\|$$

$$\leq \|h_* - h_{n+1}\| + ||T_{n+1}h_{n+1} - T h_*\|$$

$$= \|h_* - h_{n+1}\| + ||T_{n+1}h_{n+1} - T h_*\|$$

$$\leq ||T_{n+1}h_{n+1} - T h_*\|,$$

where we used property $13$ for the final two inequalities. Taking limits in $n$, all summands on the right-hand side vanish, from which we conclude that $h_* = I_{A^c} + I_{A^c} \cdot T h_*$. In other words, $h_*$ is a solution of (1). Because $T$ satisfies $11, 12$ and $13$ by Proposition 2 Equation (1) has a unique solution $h_*$, and hence we have that $h = h_*$. 

We remark without proof that if, in the statement of Proposition 7 we instead take each $T_{n+1}$ to be an extreme point of $T$ such that $Th_n = T_{n+1}h_n$, then the sequence instead becomes non-decreasing, and converges to a limit that is the unique non-negative solution of (3) in $\mathbb{R}^N$. So, we can use a completely similar method to compute upper expected hitting times. That said, in the remainder of this work, we will focus on the version for lower expected hitting times.

## 4 Complexity Analysis

We will say that a sequence $\{h_n\}_{n \in \mathbb{N}}$ of vectors is strictly decreasing if $h_{n+1} \leq h_n$ and $h_{n+1} \neq h_n$ for all $n \in \mathbb{N}$. 
Corollary 6. Let $T : \mathbb{R}^X \to \mathbb{R}^X$ be a map that satisfies C1-C3 and H7, and let $T$ be its dual representation. Then any sequence $\{h_n\}_{n \in \mathbb{N}}$ in $\mathbb{R}^X$ constructed as in Proposition 4 is either strictly decreasing everywhere, or, for some $m \in \mathbb{N}$, is strictly decreasing for all $n < m$ and satisfies $h_k = h_\star$ for all $k \geq m$, with $h_\star = \lim_{n \to +\infty} h_n$.

Proof. Because we know from Proposition 7 that the sequence $\{h_n\}_{n \in \mathbb{N}}$ is non-increasing, the converse of this statement is that there is some $m \in \mathbb{N}$ such that $h_m = h_{m+1}$ with $h_m \neq h_\star$. Suppose ex absurdo that this is the case.

Consider the co-sequence $\{T_n\}_{n \in \mathbb{N}}$ of extreme points of $T$ that was used to construct the sequence $\{h_n\}_{n \in \mathbb{N}}$. Let $T := T_{m+1}$ and $h := h_m = h_{m+1}$. Then it holds that $Th = Th_n$, and $h$ is the unique solution in $\mathbb{R}^X$ of the linear system $h = I_{A^c} + I_{A^c} \cdot Th$. Following the conditions of Proposition 7, we now construct a sequence $\{T'_n\}_{n \in \mathbb{N}}$, and a co-sequence $\{h'_n\}_{n \in \mathbb{N}}$ in $\mathbb{R}^X$, such that $h'_n = I_{A^c} + I_{A^c} \cdot T'_n h'_n$ for all $n \in \mathbb{N}$. To this end, first set $T_1' := T$. This yields $h'_1 = h$, and because $T \in T$ is an extreme point of $T$ that satisfies $T h' = T h = T h_\star$, we can take $T_2' := T$, which yields $h'_2 = h$. Proceeding in this fashion, we obtain $T_n' = T$ and $h'_n = h$ for all $n \in \mathbb{N}$. This sequence $\{h'_n\}_{n \in \mathbb{N}}$ satisfies the conditions of Proposition 7 but $\lim_{n \to +\infty} h'_n = h = h_m \neq h_\star$, a contradiction.

Corollary 7. Let $T : \mathbb{R}^X \to \mathbb{R}^X$ be a map that satisfies C1-C3 and H7, and let $T$ be its dual representation. Suppose that $T$ has a finite number $m \in \mathbb{N}$ of extreme points. Then any sequence $\{h_n\}_{n \in \mathbb{N}}$ in $\mathbb{R}^X$ constructed as in Proposition 7 satisfies $h_k = h_\star = \lim_{t \to +\infty} h_t$ for all $k \geq n$, for some $n \leq m$.

Proof. Because $T$ has $m \in \mathbb{N}$ extreme points, it follows from Proposition 7 that the elements of the sequence $\{h_n\}_{n \in \mathbb{N}}$ can take at most $m$ distinct values; each corresponding to an extreme point of $T$. Hence there is some $n \leq m$ such that $h_n = h_{n+1}$, that is, $h_{m+1}$ must have a value that was already obtained earlier in the sequence. Now apply Corollary 6.

This result suggests that the numerical scheme proposed in Proposition 7 takes at most $m$ iterations, where $m$ is the number of extreme points of $T$. In fact, it is possible to show that this bound is tight, i.e. for any $m \in \mathbb{N}$ there is a map $T$ whose dual representation $T$ has $m$ extreme points, for which the method suggested in Proposition 7 takes exactly $m$ iterations when $T_1 \in T$ is chosen carefully. Unfortunately, the construction that shows the tightness is fairly involved, and we must omit it here for reasons of brevity.

Let us now remark on the other computational aspects of the algorithm. At each step $n \in \mathbb{N}$ of the method, we need to solve two problems. First, given $T_n \in T$, we need to compute $h_n$. This is equivalent to computing the expected hitting time of $A$ for the precise Markov chain identified by $T_n$, so this step can be solved by any method available in the literature for the latter problem. However, Proposition 6 tells us that $h_n$ can be obtained by solving a linear system. Hence, ignoring issues of numerical precision, the complexity of this step is at most
\( \mathcal{O}(|\mathcal{X}|^\omega) \), which is the complexity of multiplying two \(|\mathcal{X}| \times |\mathcal{X}|\) matrices; the current best estimate for this exponent is \( \omega \approx 2.38 \) \[2\].

Secondly, we need to find an extreme point \( T_{n+1} \in \mathcal{T} \) such that \( T_{n+1}h_n = \inf_{T \in \mathcal{T}} Th_n \). The complexity of this step will depend strongly on the way \( \mathcal{T} \) is encoded. In many practical applications, however, it will be derived from a given set \( \mathcal{T} \) of linear maps satisfying \[3\] that is, the dual representation \( \mathcal{T} \) is often given explicitly and used to describe \( \mathcal{T} \) rather than the other way around. Since \( \mathcal{T} \) has separately specified rows, to identify \( \mathcal{T} \) one only has to specify the sets \( \mathcal{T}_x \). In turn, these sets \( \mathcal{T}_x \), can, in practice, often be described by a finite number of linear inequalities; in the context of Markov chains, these inequalities represent given bounds on the transition probabilities for moving from the state \( x \) to other states. Under these conditions, each \( \mathcal{T}_x \) will be a set that is non-empty (assuming feasibility of the specified constraints), closed, and convex, with a finite number of extreme points.

In this case, computing each \([Tf](x)\) (and in particular \([Th_n](x)\)) reduces to solving a linear programming problem in \( N = |\mathcal{X}| + c \) variables\[2\] where \( c \) is the number of constraints used to specify \( \mathcal{T}_x \). For instance, the simplex algorithm can be used to solve this problem, with the added benefit that the returned optimal solution is an extreme point of \( \mathcal{T}_x \). Unfortunately, the complexity of the simplex algorithm depends on the pivot rule that is used, and the existence of a polynomial-time pivot rule is still an open research question \[3\].

Nevertheless, it typically performs very well in practice; roughly speaking, probabilistic analyses typically yield an expected number of iterations on the order of \( \mathcal{O}(N^{2+\alpha}) \), with \( \alpha > 0 \) depending on the specifics of the analysis. For example, \[4\] gives a bound for the expected number of iterations on the order of \( \mathcal{O}(|\mathcal{X}|^2\sqrt{\log c\sigma^{-2}} + |\mathcal{X}|^3 \log^4 c) \), where \( \sigma \) is a parameter of the probabilistic analysis. Crucially, however, we are not aware of a result that shows an (expected) runtime of the simplex algorithm that is sub-quadratic in \(|\mathcal{X}|\). Moreover, a very recent result \[2\] gives a deterministic interior-point algorithm that solves this problem (approximately), and which runs in \( \mathcal{O}(N^{\omega}) \)\[3\] where, as before, \( \mathcal{O}(N^{\omega}) \) is the complexity of multiplying two \( N \times N \) matrices. As that author notes, this complexity can be regarded as essentially optimal. However, because this is an interior-point method, the resulting solution is not guaranteed to be an extreme point of \( \mathcal{T}_x \), which we require for our algorithm, so we cannot use the result here other than for illustrating the complexity of the minimisation problem.

In any case, the above considerations suggest that finding \( T_x \in \mathcal{T}_x \) such that \( T_xf = [Tf](x) \), for a generic \( T_x \) that is specified using a finite number of linear constraints, will typically take \textit{at least} \( \mathcal{O}(|\mathcal{X}|^2) \) time. This then needs to be repeated \( |\mathcal{X}| \) times, once for each \( x \in \mathcal{X} \). Hence, finding an extreme point \( T_{n+1} \in \mathcal{T} \) such that \( Th_n = T_{n+1}h_n \) can be done in polynomial expected time,

\[2\] It is worth noting that this analysis is somewhat pessimistic, because it assumes no real structure on the underlying dynamical system; in practice it will often be unlikely that a given state \( x \) can move to every other state \( y \) in a single step, whence \( \mathcal{T}_x \) will live in a subspace of dimension (much) less than \(|\mathcal{X}|\).

\[3\] Following \[2\], \( \mathcal{O}(N) \) hides \( \text{polylog}(N) \) factors.
but requires at least $O(|\mathcal{X}|^3)$ time (in expectation). In conclusion, the above analysis tells us that a single step of our proposed method has an expected runtime complexity of at least $O(|\mathcal{X}|^3)$, where the complexity is dominated by finding $T_{n+1} \in \mathcal{T}$ such that $T_h n = T_{n+1} h_n$, rather than by computing $h_n$ from the previous $T_n$.

Let us finally compare this to the only other method in the literature of which we are aware, which is described in Proposition 3. Starting with $h_0 := I_A c$, at each step of this method, we need to compute $T h n - 1$. Following the discussion above, and depending on the algorithm and analysis used, we can expect this to have a complexity of at least $O(|\mathcal{X}|^3)$. Since this was also the dominating factor in the complexity of our new algorithm, we conclude that the two methods have comparable complexity (order) per iteration. What remains is a comparison of the number of iterations they require. With $h_n$ as in Proposition 3 it follows from [17, Lemma 42] that $h_n \leq n + 1$ for all $n \in \mathbb{N}$. Hence, for any $\epsilon > 0$, this method will take at least $(\|h\| - \epsilon) - 1$ iterations, before $\|h_n - h\| \leq \epsilon$. The new method from Proposition 7 is thus more efficient than the method from Proposition 3 when the number $m$ of extreme points of $\mathcal{T}$ is less than $\|h\|$. Of course, since we do not know $h$ to begin with, this does not provide a practically useful way to determine which method to use in practice.

Moreover, because the number $m$ of extreme points of $\mathcal{T}$ is potentially extremely large, one may wonder whether the condition $m < \|h\|$ is ever really satisfied in practice. However, a first empirical analysis suggests that the bound $m$ on the number of iterations, although tight, is not representative of the average case complexity of our algorithm.

In particular, we performed experiments where we randomly created sets $\mathcal{T}$, and applied the method from Proposition 7 to the resulting systems. In these experiments, we varied the size of $\mathcal{X}$ between $10^2$ and $10^3$, while keeping $\mathcal{A}$ a singleton throughout. For each setting, we generated sets $\mathcal{T}_x$ by sampling 50 points uniformly at random in the space of probability mass functions on $\mathcal{X}$, and taking $\mathcal{T}_x$ to be the convex hull of these points. Thus each $\mathcal{T}_x$ has 50 extreme points (almost surely with respect to the uniform sampling), and the set $\mathcal{T}$ constructed from them has $50|\mathcal{X}|$ extreme points. We then ran the method from Proposition 7 where we selected $T_1$ such that $T h A = T_1 I_A$, and the number of iterations until convergence was recorded. Using the notation from Proposition 7 we say that it converged after $n > 1$ iterations, when we observe $h_n = h_{n-1}$; this is valid by Corollary 6. This was repeated 50 times for each value of $|\mathcal{X}|$.

The results of these experiments are shown in Figure 1. We see that in the majority of cases, the method converges to the correct solution in three iterations, although sometimes it takes four iterations. There were no instances where the method took more iterations to converge, although in one instance (not shown) with $|\mathcal{X}| = 10^3$, the method converged in only two iterations.

Although still preliminary, we believe that these results are indicative of an average runtime complexity that is vastly more efficient than the worst case suggested by Corollary 2. In future work, we hope to examine this average case performance more thoroughly. Finally, in [10] the authors already noted the close
connection between the system \cite{1}, and the equations of optimality that one encounters in the theory of Markov decision processes (MDPs) \cite{11}. Moreover, the method that we propose in Proposition \ref{prop:policy} although discovered independently, is reminiscent of the policy iteration algorithm for MDPs. Interestingly, the method in Proposition \ref{prop:value} bears a similar resemblance to the value iteration algorithm for MDPs. We hope to explore these connections more fully in future work.

**Acknowledgements.** The main results of this work were previously presented, without proof, as an abstract-and-poster contribution at ISIPTA 2019. The author wishes to express his sincere gratitude to Jasper De Bock for stimulating discussions and insightful comments during the preparation of this manuscript. He would also like to thank an anonymous reviewer for their helpful suggestions.

**References**

\[1\] Augustin, T., Coolen, F.P., Cooman, G.D., Troffaes, M.C.: Introduction to Imprecise Probabilities. Wiley (2014)

\[2\] Brand, J.v.d.: A deterministic linear program solver in current matrix multiplication time. In: Symposium on Discrete Algorithms, pp. 259–278 (2020)

\[3\] Campos, M.A., Dimuro, G.P., da Rocha Costa, A.C., Kreinovich, V.: Computing 2-step predictions for interval-valued finite stationary Markov chains (2003)

\[4\] Dadush, D., Huiiberts, S.: A friendly smoothed analysis of the simplex method. Proceedings of STOC 2018 pp. 390–403 (2018)

\[5\] De Bock, J.: The limit behaviour of continuous-time imprecise Markov chains. Journal of Nonlinear Science 27(1), 159–196 (2017)

\[6\] De Cooman, G., De Bock, J., Lopatatzidis, S.: Imprecise stochastic processes in discrete time: global models, imprecise Markov chains, and ergodic theorems. International Journal of Approximate Reasoning 76, 18–46 (2016)

\[7\] De Cooman, G., Hermans, F.: Imprecise probability trees: Bridging two theories of imprecise probability. Artificial Intelligence 172(11), 1400–1427 (2008)
[8] De Cooman, G., Hermans, F., Antonucci, A., Zaffalon, M.: Epistemic irrelevance in credal nets: the case of imprecise Markov trees. International Journal of Approximate Reasoning 51(9), 1029–1052 (2010)
[9] De Cooman, G., Hermans, F., Quaeghebeur, E.: Imprecise Markov chains and their limit behavior. Probability in the Engineering and Informational Sciences 23(4), 597–635 (2009)
[10] Dunford, N., Schwartz, J.T.: Linear operators, part 1: general theory, vol. 10. John Wiley & Sons (1988)
[11] Feinberg, E.A., Shwartz, A.: Handbook of Markov decision processes: methods and applications, vol. 40. Springer Science & Business Media (2012)
[12] Hartfiel, D.J.: Markov set-chains. Springer (2006)
[13] Hermans, F., Škulj, D.: Stochastic processes. In: Augustin, T., Coolen, F.P., Cooman, G.D., Troffaes, M.C. (eds.) Introduction to Imprecise Probabilities, chap. 11. Wiley (2014)
[14] Ibarra, O.H., Moran, S., Hui, R.: A generalization of the fast LUP matrix decomposition algorithm and applications. Journal of Algorithms 3(1), 45–56 (1982)
[15] Kozine, I.O., Utkin, L.V.: Interval-valued finite Markov chains. Reliable computing 8(2), 97–113 (2002)
[16] Krak, T., T’Joens, N., De Bock, J.: Hitting times and probabilities for imprecise Markov chains. Proceedings of ISIPTA 2019 pp. 265–275 (2019)
[17] Krak, T., T’Joens, N., De Bock, J.: Hitting times and probabilities for imprecise Markov chains. https://arxiv.org/abs/1905.08781 (2019)
[18] Lopatatzidis, S.: Robust Modelling and Optimisation in Stochastic Processes Using Imprecise Probabilities, with an Application to Queueing Theory. Ph.D. thesis, Ghent University (2016)
[19] Norris, J.: Markov Chains. Cambridge University Press (1997)
[20] Škulj, D.: Finite discrete time Markov chains with interval probabilities. In: Soft Methods for Integrated Uncertainty Modelling, pp. 299–306. Springer (2006)
[21] Todd, M.: The many facets of linear programming. Mathematical Programming 91 (2002). https://doi.org/10.1007/s101070100261
[22] Walley, P.: Statistical reasoning with imprecise probabilities (1991)