ERGODICITY CONDITIONS FOR UPPER TRANSITION OPERATORS

FILIP HERMANS AND GERT DE COOMAN

ABSTRACT. We study ergodicity of bounded, sub-additive and non-negatively homogeneous maps on finite dimensional spaces which we call upper transition operators. We show that ergodicity coincides with the necessary and sufficient condition for a generalised Perron-Frobenius theorem for upper transition operators. We show that ergodicity is equivalent with regular absorningness of the upper transition operator: there has to be a top class that is regular and absorbing. Using this conditions, it can be shown that top class regularity can be checked by solving a linear eigenvalue problem where the stochastic matrix is build with the values of the upper transition operator for every atom. To check top class absorption it is shown that less than n evaluations of the upper transition operator have to be done.

1. INTRODUCTION

Throughout the paper, $\mathscr{X}$ denotes a finite non-empty set of elements that we also refer to as *states*, and $\mathscr{L}(\mathscr{X})$ is the set of all real-valued maps on $\mathscr{X}$. We provide the finite-dimensional linear space $\mathscr{L}(\mathscr{X})$ with the supremum norm $\|\cdot\|_\infty$, or with the topology of uniform convergence, so the result is a Banach space. Observe that uniform convergence and point-wise convergence coincide on this finite-dimensional space.

Definition 1. An *upper transition operator* on $\mathscr{L}(\mathscr{X})$ is a transformation $\mathrm{T}\colon \mathscr{L}(\mathscr{X}) \to \mathscr{L}(\mathscr{X})$ that has the following properties:

(1) $\min f \leq \mathrm{T}f \leq \max f$ — T is *bounded*;
(2) $\mathrm{T}(f+g) \leq \mathrm{T}f + \mathrm{T}g$ — T is *sub-additive*;
(3) $\mathrm{T}(\lambda f) = \lambda \mathrm{T}g$ — T is *non-negatively homogeneous*;

for arbitrary f, g in $\mathscr{L}(\mathscr{X})$ and real $\lambda \geq 0$.

In this paper we will often refer to the work done on finite and discrete Markov chains. At any timepoint k, these Markov chains can be described by a transition matrix $M^{(k)}$ where the i-th row $M^{(k)}_{i,\cdot}$ gives the probability distribution over the states at timepoint $k+1$ conditionally on the chain being in state x_i at time k. If we now assume that these conditional transition distributions can be picked from a set of probability distributions $\mathscr{M}_{x_i}$ depending on the state x_i but with the restriction that each set has to be time-invariant and convex, then any transition operator has to belong to

$$\mathscr{T} := \left\{ M \in \mathbb{R}^{|\mathscr{X}|\times|\mathscr{X}|} \colon (\forall x_i \in \mathscr{X})(M_{i,\cdot} \in \mathscr{M}_{x_i}) \right\}.$$

It can be shown [1] that there corresponds exactly one upper transition operator T with $\mathscr{T}$ such that the upper envelope of the expectation of a vector $f \in \mathscr{L}(\mathscr{X})$ after k steps can be calculated as

$$\max \left\{ M^{(1)}_{i,\cdot} M^{(2)} \dots M^{(k)} f \colon M^{(j)} \in \mathscr{T} \right\} = \mathrm{T}^k f(x_i).$$

As a result, $\mathrm{T}\mathbb{I}_A(x)$ can be interpreted as the upper transition probability and $1 - \mathrm{T}\mathbb{I}_{A^c}(x)$ the lower transition probability to go from state x in one step to a set of states A. In general, an upper transition operator can be seen as the summarization of a set of non-stationary Markov chains.

2000 *Mathematics Subject Classification.* 60J10, 93B35, 47H09, 47N30.
Key words and phrases.

Any upper transition operator T automatically also satisfies the following interesting properties:

(T4) $\mathrm{T}(f+\mu) = \mathrm{T}f + \mu$ — T is *constant-additive*;
(T5) if $f \leq g$ then $\mathrm{T}f \leq \mathrm{T}g$ — T is *order-preserving*;
(T6) if $f_n \to f$ then $\mathrm{T}f_n \to \mathrm{T}f$ — T is *continuous*;
(T7) $\mathrm{T}f + \mathrm{T}(-f) \geq 0$ — T is *upper–lower consistent*;

for arbitrary f, g, f_n in $\mathscr{L}(\mathscr{X})$ and real μ. Clearly, for any n in the set of natural numbers (with zero) $\mathbb{N}_0$, T^n is an upper transition operator as well.

Properties (T4) and (T5) define a topical map [4]. It is easy to see [4] that every topical map is also non-expansive under the supremum norm:

(T8) $\|\mathrm{T}g - \mathrm{T}g\|_\infty \leq \|f - g\|_\infty$ — T is *non-expansive*;

for every f and g in $\mathscr{L}(\mathscr{X})$.

A very useful result for non-expansive maps by Sine [6] states that for every element f of the finite-dimensional domain of a non-expansive transformation T, there is some natural number p such that the sequence $\mathrm{T}^{np}f$ converges. More importantly, Sine proves that we can find a 'period' p *common* to all maps f in the domain $\mathscr{L}(\mathscr{X})$. This means that for any f, the set $\omega_\mathrm{T}(f)$ of limit points of the set of iterates $\{\mathrm{T}^n f \colon n \in \mathbb{N}\}$ has a number of elements $|\omega_\mathrm{T}(f)|$ that divides p.[1] T is cyclic on $\omega_\mathrm{T}(f)$, with period $|\omega_\mathrm{T}(f)|$ (and therefore also with period p). Lemmens and Scheutzow [4] managed to prove that an upper bound for the common periods of all topical functions $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^n$ is $\binom{n}{\lfloor n/2 \rfloor}$. This upper bound is tight in the sense that there is always at least one topical function that has this bound as its smallest common period.

In Sec. 2 we introduce ergodicity and explain the link with Perron-Frobenius conditions. The relationship between ergodicity and eigenvalues of the upper transition operator is also exposed. In Sec. 3 we attempt to develop an algorithm that allows us the check for regularity in practise. We start doing so by orderering the statespace by means of an accessibility relation. From the classes that are induced by the ordering we will be able to formulate two alternative conditions for ergodicity: top class regularity and top class absorption. In the remainder of the section we will work out a test for every condition separately. Each of these test will turn out to have a very nice complexity. In Sec. 4 we will explain the pros and contras of the coefficient of ergodicity that has been defined by other authors.

2. Perron-Frobenius condition for upper transition operators

In this section we introduce the notion of ergodicity for upper transition operators. We allow ourselves to be inspired by corresponding notions for non-stationary Markov chains [5, p. 136] and Markov set chains [2], which leads us to the following definition.

Definition 2. An upper transition operator T on $\mathscr{L}(\mathscr{X})$ is called *ergodic* if for all $f \in \mathscr{L}(\mathscr{X})$, $\lim_{n\to\infty} \mathrm{T}^n f$ exists and is a constant function.

Consider any $f \in \mathscr{L}(\mathscr{X})$. Ergodicity of an upper transition operator T not only means that the sequence T^n converges, so $\omega_\mathrm{T}(f)$ is a singleton $\{\xi_f\}$, but also that this limit ξ_f is a constant function. Observe that by (T6), ξ_f is a fixed point for all T^k: $T^k\xi_f = \xi_f$ and therefore $\xi_{T^k f} = \xi_f$ for all $k \in \mathbb{N}$. If we denote the constant value of ξ_f by $\overline{E}(f)$, then this defines a real functional $\overline{E}$ on $\mathscr{L}(\mathscr{X})$. This functional is an *upper expectation*: it is bounded, sub-additive and non-negatively homogeneous [compare with (T1)–(T3)]. This upper expectation is T-invariant in the sense that $\overline{E} \circ \mathrm{T} = \overline{E}$, and it is the only such upper expectation.

[1] $|A|$ denotes the cardinality of a set A and $\mathbb{N}$ is the set of natural numbers (without zero).

Definition 3. An upper transition operator T on $\mathscr{L}(\mathscr{X})$ is called *Perron–Frobenius-like* if there is some real functional $\overline{E}_\infty$ on $\mathscr{L}(\mathscr{X})$ such that for all upper expectations $\overline{E}$ on $\mathscr{L}(\mathscr{X})$ and all $f \in \mathscr{L}(\mathscr{X})$:

$$\lim_{n\to\infty} \overline{E}(\mathrm{T}^n f) = \overline{E}_\infty(f).$$

or in other words, if the sequence of upper expectations $\overline{E}\circ\mathrm{T}^n$ converges to some limit that does not depend on the initial value $\overline{E}$.

As an immediate result, conditions for ergodicity of upper transition operators are conditions for a Perron–Frobenius-like theorem for such transformations to hold.

Theorem 1 (Perron–Frobenius)**.** *An upper transition operator* T *is Perron–Frobenius-like if and only if it is ergodic, and in that case* $\overline{E}_\infty = \overline{E}_\mathrm{T}$.

Proof. Sufficiency. Suppose T is ergodic. Then using the notations established above, $\mathrm{T}^n f \to \xi_f$ and therefore $\overline{E}(\mathrm{T}^n f) \to \overline{E}(\xi_f)$ because any upper expectation $\overline{E}$ is continuous [compare with (T6)]. Observe that, since any upper expectation $\overline{E}$ is constant-additive [compare with (T4) and (T1)], $\overline{E}(\xi_f) = \overline{E}_\mathrm{T}(f)$. Hence $\overline{E}\circ\mathrm{T}^n \to \overline{E}_\mathrm{T}$, and therefore T is Perron–Frobenius-like, with $\overline{E}_\infty = \overline{E}_\mathrm{T}$.

Necessity. Suppose that T is Perron–Frobenius-like, with limit upper expectation $\overline{E}_\infty$. Fix any $x \in \mathscr{X}$, and consider the upper expectation $\overline{E}_x$ defined by $\overline{E}_x(f) := f(x)$ for all $f \in \mathscr{L}(\mathscr{X})$. Then by assumption $T^n f(x) = \overline{E}_x(\mathrm{T}^n f) \to \overline{E}_\infty(f)$. Since this holds for all $x \in \mathscr{X}$, we see that T is ergodic with $\overline{E}_\mathrm{T} = \overline{E}_\infty$. □

It follows from the discussion in Sec. 1 that $\bigcup_{f\in\mathscr{L}(\mathscr{X})} \omega_\mathrm{T}(f)$ is the set of all periodic points of T—*a periodic point* being an element $f \in \mathscr{L}(\mathscr{X})$ for which there is some $n \in \mathbb{N}$ for which $\mathrm{T}^n f = f$. Because of (T4), this set containts all constant maps. We now see that for T to be ergodic, this set cannot contain any other maps.

Proposition 2. *An upper transition operator* T *is ergodic if and only if all of its periodic points are constant maps.*

3. Ergodicity in practise

We now turn to the issue of determining in actual practice whether an upper transition operator is ergodic.

In the case of finite-state, discrete-time Markov chains, a nice approach to deciding upon ergodicity is given by Kemeny and Snell [3, Sec. 1.4]. It is based on the notion of an accessibility relation. This is a binary (weak order) relation on set of states $\mathscr{X}$ that captures whether it is possible to go from one state to another in a finite number of steps. In this section, we show that it is possible to associate an accessibility relation with an upper transition operator, and that this relation provides us with an intuitive interpretation of the notion of ergodicity in terms of accessibility. We refer to [1] for a detailed discussion of accessibility relations and their connections with upper transition operators.

Definition 4. Consider an upper transition operator T on $\mathscr{L}(\mathscr{X})$, and two states x and y in $\mathscr{X}$. We say that y is *accessible* from x *in n steps*, and denote this as $x \xrightarrow{n} y$, if $\mathrm{T}^n \mathbb{I}_{\{y\}}(x) > 0$. We say that state y is *accessible* from state x, and denote this as $x \to y$, if $\mathrm{T}^n \mathbb{I}_{\{y\}}(x) > 0$ for some $n \in \mathbb{N}_0$. We say that x and y *communicate*, and denote this as $x \leftrightarrow y$, if both $x \to y$ and $y \to x$.

The relation $\to$ is a weak order (reflexive and transitive), and consequently $\leftrightarrow$ is an equivalence relation. The equivalence classes for this relation are called *communication classes*: sets of $\mathscr{X}$ for which every element has access to any other element. The accessibility relation induces a partial order on these communication classes.

In the case of finite-state, discrete-time Markov chains, this partial order gives us clues about the ergodicity of the Markov chain. Indeed for such a Markov chain to be ergodic, it is necessary and sufficient that [1] it should be *top class regular*, meaning that: (i) there should be only one *maximal* or *undominated* communication class—elements of a maximal communication class have no access to states not in that class—, in which case we call this unique maximal class $\mathscr{R}$ the *top class*; and (ii) the top class $\mathscr{R}$ should be regular, meaning that after some time k, all elements of this class become accessible to each other in any number of steps: for all x and y in $\mathscr{R}$ and for all $n \geq k$, $x \xrightarrow{n} y$.

For upper transition operators, it turns out that top class regularity is a necessary condition for ergodicity. However, top class regularity is by itself not a sufficient condition: we need some garantee that the top class will eventually be reached—a requirement that is automatically fulfilled in finite-state discrete-time Markov chains.

Proposition 3. *An upper transition operator* T *is ergodic if and only if it is* regularly absorbing*, meaning that it satisfies the following properties:*

(TCR) *it is* top class regular*:*

$$\mathscr{R} := \left\{ x \in \mathscr{X} \colon (\exists n \in \mathbb{N})(\forall k \geq n) \min \mathrm{T}^k \mathbb{I}_{\{x\}} > 0 \right\} \neq \emptyset,$$

(TCA) *it is* top class absorbing*: with* $\mathscr{R}^c := \mathscr{X} \setminus \mathscr{R}$*,*

$$(\forall y \in \mathscr{R}^c)(\exists n \in \mathbb{N})\, \mathrm{T}^n \mathbb{I}_{\mathscr{R}^c}(y) < 1.$$

For a proof that (TCR) is equivalent to $\mathscr{R} \neq \emptyset$, we refer to [1, Prop. 4.3]. (TCA) means that for every element y not in the top class, there is a strictly positive *lower* probability $1 - \mathrm{T}\mathbb{I}_{\mathscr{R}^c}(y)$ to access an element in the top class.

Proof. (TCR) $\wedge$ (TCA) $\Rightarrow$ (ER). Consider any fixed point ξ of T^k, where $k \in \mathbb{N}$. We infer from Prop. 2 that we have to show that ξ is constant. Using (T5), (T4) and (T3) we construct from $\xi \geq \min \xi + [\xi(x) - \min \xi]\mathbb{I}_{\{x\}}$ the following inequality, which holds for all $n \in \mathbb{N}$ and all $x \in \mathscr{X}$:

$$\mathrm{T}^{nk}\xi \geq \min \xi + [\xi(x) - \min \xi]\, \mathrm{T}^{nk}\mathbb{I}_{\{x\}}.$$

Hence, by taking the minimum on both sides of this inequality and using that $\mathrm{T}^{nk}\xi = \xi$, we find that

$$0 \geq [\xi(x) - \min \xi] \min \mathrm{T}^{nk}\mathbb{I}_{\{x\}}.$$

We infer from (TCR) that by taking n large enough, we can ensure that $\min \mathrm{T}^{kn}\mathbb{I}_{\{x\}} > 0$ if $x \in \mathscr{R}$. So we already find that $\xi(x) = \min \xi$ for all $x \in \mathscr{R}$.

If ξ reaches its maximum on $\mathscr{R}$, then $\max \xi = \min \xi$ so ξ is indeed constant. Let us therefore assume that the maximum of ξ is not reached in $\mathscr{R}$. Using (T5), (T4) and (T3) we construct from $\xi \leq \max \xi - [\max \xi - \max_{x \in \mathscr{R}} \xi(x)]\mathbb{I}_{\mathscr{R}}$ and $-\mathbb{I}_{\mathscr{R}} = \mathbb{I}_{\mathscr{R}^c} - 1$ the following inequality, which holds for all $n \in \mathbb{N}$:

$$\mathrm{T}^n\xi \leq \max \xi + \left[\max \xi - \max_{x \in \mathscr{R}} \xi(x)\right] (\mathrm{T}^n \mathbb{I}_{\mathscr{R}^c} - 1).$$

By taking the maximum over $\mathscr{R}^c$ on both sides we get

$$0 = \max_{y \in \mathscr{R}^c} \mathrm{T}^n \xi(y) - \max \xi \leq \left[\max \xi - \max_{x \in \mathscr{R}} \xi(x)\right] \left(\max_{y \in \mathscr{R}^c} \mathrm{T}^n \mathbb{I}_{\mathscr{R}^c}(y) - 1\right).$$

If we choose $n = \max\left\{n_y \colon y \in \mathscr{R}^c, \mathrm{T}^{n_y}\mathbb{I}_{\mathscr{R}^c}(y) < 1\right\}$ then we see that for every $y \in \mathscr{R}^c$

$$\mathrm{T}^n \mathbb{I}_{\mathscr{R}^c}(y) = \mathrm{T}^{n_y}[(\mathbb{I}_{\mathscr{R}} + \mathbb{I}_{\mathscr{R}^c})\mathrm{T}^{n-n_y}\mathbb{I}_{\mathscr{R}^c}](y) = \mathrm{T}^{n_y}[\mathbb{I}_{\mathscr{R}^c}\mathrm{T}^{n-n_y}\mathbb{I}_{\mathscr{R}^c}](y) \leq \mathrm{T}^{n_y}\mathbb{I}_{\mathscr{R}^c}(y) < 1.$$

But this means that $\max_{x \in \mathscr{R}^c} \mathrm{T}^n \mathbb{I}_{\mathscr{R}^c}(y) - 1 < 0$ and consequently

$$\max \xi = \max_{x \in \mathscr{R}} \xi = \max \min \xi = \min \xi.$$

(ER) $\Rightarrow$ (TCR) $\wedge$ (TCA). We will use contraposition and show first that $\neg$(TCR) $\Rightarrow$ $\neg$(ER). Then we will show that $\neg$(TCA) $\wedge$ (TCR) $\Rightarrow$ $\neg$(ER).

$\neg$(TCR) $\Rightarrow$ $\neg$(ER). Not being top class regular means that $\mathscr{R} = \emptyset$, which is equivalent to

$$(\forall x \in \mathscr{X})(\forall n \in \mathbb{N})(\exists k \geq n)(\exists z \in \mathscr{X})\mathrm{T}^k\mathbb{I}_{\{x\}}(z) = 0.$$

Since we infer from $\mathbb{I}_{\{x\}} \geq 0$ and (T1) that $\mathrm{T}^k\mathbb{I}_{\{x\}} \geq 0$, this implies that $\liminf_{n\to\infty} \min \mathrm{T}^n\mathbb{I}_{\{x\}} = 0$. But for any $n \in \mathbb{N}$, $\mathrm{T}^{n+1}\mathbb{I}_{\{x\}} = \mathrm{T}(\mathrm{T}^n\mathbb{I}_{\{x\}}) \geq \min \mathrm{T}^n\mathbb{I}_{\{x\}}$ by (T1), and therefore also $\min \mathrm{T}^{n+1}\mathbb{I}_{\{x\}} \geq \min \mathrm{T}^n\mathbb{I}_{\{x\}}$. This implies that the sequence $\min \mathrm{T}^n\mathbb{I}_{\{x\}}$ is non-decreasing, and bounded above [by 1], and therefore convergent. This leads to the conclusion that

$$(\forall x \in \mathscr{X}) \lim_{n\to\infty} \min \mathrm{T}^n\mathbb{I}_{\{x\}} = 0. \tag{1}$$

We also infer from (T1) and (T2) that $1 = \mathrm{T}^k\mathbb{I}_{\mathscr{X}} \leq \sum_{x\in\mathscr{X}} \mathrm{T}^k\mathbb{I}_{\{x\}}$. Since the cardinality $|\mathscr{X}|$ of the state space is finite, this means that for all $z \in \mathscr{X}$ and all $n \in \mathbb{N}$ there is some $x \in \mathscr{X}$ such that $\mathrm{T}^n\mathbb{I}_{\{x\}}(z) \geq 1/|\mathscr{X}|$. This tells us that $\max \mathrm{T}^n\mathbb{I}_{\{x\}} \geq 1/|\mathscr{X}|$. Since we can infer from a similar argument as before that the sequence $\max \mathrm{T}^n\mathbb{I}_{\{x\}}$ converges, this tells us that

$$(\forall x \in \mathscr{X}) \lim_{n\to\infty} \max \mathrm{T}^n\mathbb{I}_{\{x\}} \geq \frac{1}{|\mathscr{X}|}. \tag{2}$$

Combining Eqs. (1) and (2) tells us that $\lim_{n\to\infty}(\max \mathrm{T}^n\mathbb{I}_{\{x\}} - \min \mathrm{T}^n\mathbb{I}_{\{x\}}) > 0$, so T cannot be ergodic.

$\neg$(TCA) $\wedge$ (TCR) $\Rightarrow$ $\neg$(ER). Since T is not top class absorbing, we know that there is some $y \in \mathscr{R}^c$ such that $\mathrm{T}^n\mathbb{I}_{\mathscr{R}^c}(y) = 1$ for all $n \in \mathbb{N}$. As the top class $\mathscr{R}$ is non-empty, we know that there is some $x \in \mathscr{R}$, and this x has no access to any state outside the maximal communication class $\mathscr{R}$: $\mathrm{T}^n\mathbb{I}_{\mathscr{R}^c}(x) = 0$ for all $n \in \mathbb{N}$. Consequently

$$\lim_{n\to\infty} (\max \mathrm{T}^n\mathbb{I}_{\mathscr{R}^c} - \min \mathrm{T}^n\mathbb{I}_{\mathscr{R}^c}) = 1 - 0 > 0,$$

so T cannot be ergodic. □

3.1. **Checking top class regularity.**

Checking for top class regularity directly using the definition would involve calculating for every state x the maps $\mathrm{T}\mathbb{I}_{\{x\}}$, $\mathrm{T}^2\mathbb{I}_{\{x\}}$, ..., $\mathrm{T}^n\mathbb{I}_{\{x\}}$ until a first number $n = n_x$ is found for which $\min \mathrm{T}^{n_x}\mathbb{I}_{\{x\}} > 0$. Unfortunately, it is not clear whether this procedure is guaranteed to terminate after a certain number of iterations, or whether we can stop checking after a fixed number of iterations. In this section, we want to take a closer look at this problem.

The next proposition shows that all the information we need in order to check top class regularity is incorporated in a single application of T to the atoms of $\mathscr{X}$.

Proposition 4. *Let* T *be an upper transition operator on $\mathscr{L}(\mathscr{X})$, $n \in \mathbb{N}$ and $x, y \in \mathscr{X}$. Then $\mathrm{T}^n\mathbb{I}_{\{y\}}(x) > 0$ if and only if there is some sequence x_0, x_1, x_2, ..., x_{n-1}, x_n in $\mathscr{X}$ with $x_0 = x$ and $x_n = y$ such that $\mathrm{T}\mathbb{I}_{\{x_{k+1}\}}(x_k) > 0$ for all $k \in \{0, 1, \dots, n-1\}$.*

Proof. Sufficiency. Fix k and ℓ in $\mathbb{N}$, and u and v in $\mathscr{X}$. Since $\mathrm{T}^\ell\mathbb{I}_{\{y\}} = \sum_{z\in\mathscr{X}} \mathbb{I}_{\{z\}}\mathrm{T}^\ell\mathbb{I}_{\{y\}}(z) \geq \mathbb{I}_{\{v\}}\mathrm{T}^\ell\mathbb{I}_{\{y\}}(v)$, it follows from (T5), (T2) and (T3) that $\mathrm{T}^{k+\ell}\mathbb{I}_{\{y\}} \geq \mathrm{T}^k\mathbb{I}_{\{v\}}\mathrm{T}^\ell\mathbb{I}_{\{y\}}(v)$ and therefore $\mathrm{T}^{k+\ell}\mathbb{I}_{\{y\}}(x) \geq \mathrm{T}^k\mathbb{I}_{\{v\}}(x)\mathrm{T}^\ell\mathbb{I}_{\{y\}}(v)$. Applying this inequality repeatedly, we get:

$$\mathrm{T}^n\mathbb{I}_{\{y\}}(x) \geq \prod_{k=0}^{n-1} \mathrm{T}\mathbb{I}_{\{x_{k+1}\}}(x_k)$$

for any sequence x_0, x_1, x_2, ..., x_{n-1}, x_n in $\mathscr{X}$ with $x_0 = x$ and $x_n = y$. It follows that the left-hand side is positive as soon as all factors on the right-hand side are.

Necessity. We infer using (T2) and (T3) that

$$\mathrm{T}^n\mathbb{I}_{\{y\}}(x) = \mathrm{T}\left(\sum_{x_1\in\mathscr{X}} \mathbb{I}_{\{x_1\}}\mathrm{T}^{n-1}\mathbb{I}_{\{y\}}(x_1)\right)(x) \leq \sum_{x_1\in\mathscr{X}} \mathrm{T}^{n-1}\mathbb{I}_{\{y\}}(x_1)\mathrm{T}\mathbb{I}_{\{x_1\}}(x),$$

and repeating the same argument recursively leads to

$$\mathrm{T}^n\mathbb{I}_{\{y\}}(x) \leq \sum_{\substack{x_0,x_1,\dots,x_{n-1},x_n\in\mathscr{X}\\ x_0=x,x_n=y}} \prod_{k=0}^{n-1} \mathrm{T}\mathbb{I}_{\{x_{k+1}\}}(x_k).$$

Since all the factors (and therefore all terms) on the right-hand side are non-negative by (T1) and (T5), the positivity of the left-hand side implies that there must be at least one positive term on the right-hand side, all of whose factors must therefore be positive. □

This proposition not only implies that the set $\left\{\mathrm{T}\mathbb{I}_{\{x\}}\colon x\in\mathscr{X}\right\}$ completely determines the accessibility relation $\to$, but also that it determines the 'accessibility in n steps' relation $\overset{n}{\to}$. In other words, not only the communication and maximal classes can be determined from $\{\mathrm{T}\mathbb{I}_x\colon x\in\mathscr{X}\}$, but also their regularity.

Definition 5. A stochastic matrix $M\in\mathbb{R}^{\mathscr{X}\times\mathscr{X}}$ *represents* an upper transition operator T on $\mathscr{L}(\mathscr{X})$ if it has the following form:

$$M_{x,y} := \begin{cases} c_{x,y} & \text{if } \mathrm{T}\mathbb{I}_{\{y\}}(x)>0,\\ 0 & \text{if } \mathrm{T}\mathbb{I}_{\{y\}}(x)=0,\end{cases}$$

with $c_{x,y}>0$ and $\sum_{y\in\mathscr{X}} c_{x,y}=1$.

If a stochastic matrix M represents T then we know that $M_{x,y}>0$ if and only if $\mathrm{T}\mathbb{I}_{\{y\}}(x)$. We gather from Prop. 4 that accessibility in n steps is completely determined by $\{\mathrm{T}\mathbb{I}_x\colon x\in\mathscr{X}\}$, so we may conclude that the finite-state, discrete-time Markov chain with matrix M and the upper transition operator T will invoke exactly the same 'accessibility in n steps' relations $\overset{n}{\to}$. This means that they will have the same communication classes, the same maximal classes and the same regular class. This makes it possible to use the entire machinery of finite-state, discrete-time Markov chains to decide upon top class regularity for upper transition operators. We are led to the following immediate conclusion [2, theorem 1.7].

Proposition 5 (Top class regularity)**.** *Consider an upper transition operator* T *and any stochastic matrix M that represents it. Then the following statements are equivalent:* (i) T *is top class regular;* (ii) *M is regular; and* (iii) *M has exactly one eigenvalue with modulus 1.*

Clearly this single eigenvalue with modulus 1 has to be 1 itself because M is a stochastic matrix. The problem corresponds to checking whether there exists $\phi\in\mathbb{R}$ such that

$$\frac{\det(e^{j\phi}\mathbb{I}-M)}{e^{j\phi}-1}=0$$

with $\mathbb{I}$ the identity matrix and $j^2=-1$.

Example 1. Let $\mathscr{X}:=\{x,y\}$ and $\mathrm{T}f:=f(x)\mathbb{I}_{\{x\}}+\max\{f(x),f(y)\}\mathbb{I}_{\{y\}}$ for all $f\in\mathscr{L}(\mathscr{X})$. Then $\mathrm{T}\mathbb{I}_{\{x\}}=\mathbb{I}_{\mathscr{X}}$ and $\mathrm{T}\mathbb{I}_{\{y\}}=\mathbb{I}_{\{y\}}$. This means that the stochastic matrix $M=\begin{pmatrix}1 & 0\\ 1/2 & 1/2\end{pmatrix}$ represents T. Since M has eigenvalues 1 and $1/2$, we conclude that T is top class regular.

3.2. **Checking for top class absorption.** We now present a computationally cheap procedure to check for top class absorption.

Proposition 6 (Top class absorption)**.** *Let* T *be an upper transition operator with regular top class $\mathscr{R}$. Consider the nested sequence of subsets of $\mathscr{R}^c$ defined by the iterative scheme:*

$$A_0 := \mathscr{R}^c$$
$$A_{n+1} := \{a\in A_n\colon \mathrm{T}\mathbb{I}_{A_n}(a)=1\},\quad n\geq 0.$$

Then T *is not top class absorbing if there is some $n\geq 0$ for which $A_n=A_{n+1}\neq\emptyset$. If on the other hand $A_n=\emptyset$ for some $n\geq 0$ then* T *is top class absorbing. In any case the conclusion will be reached in at most $|\mathscr{R}^c|$ iterations.*

Proof. We start by showing inductively that under the given assumptions, the statement

$$H_n: \quad \mathbb{I}_{A_n}\mathrm{T}^n\mathbb{I}_{\mathscr{R}^c} = \mathbb{I}_{A_n} \text{ and } (\forall a \in A_{n+1}^c)\mathrm{T}\mathbb{I}_{A_n}(a) < 1 \text{ and } (\forall a \in A_n^c)\mathrm{T}^n\mathbb{I}_{\mathscr{R}^c}(a) < 1$$

holds for all $n \geq 0$. We first prove that the state H_n holds for $n = 0$. The first and third statemenst of H_0 hold trivially. For the second statement, we have to prove that $\mathrm{T}\mathbb{I}_{A_0}(a) < 1$ for all $a \in A_1^c = A_0^c \cup A_0 \setminus A_1$. On $A_0 \setminus A_1$, the desired inequality holds by definition. On $A_0^c = \mathscr{R}$ it holds because there $\mathrm{T}\mathbb{I}_{A_0}$ is zero: no state in the top class $\mathscr{R}$ has access to any state outside it.

Next, we prove that $H_n \Rightarrow H_{n+1}$. First of all,

$$\mathrm{T}^{n+1}\mathbb{I}_{A_0} = \mathrm{T}(\mathrm{T}^n\mathbb{I}_{A_0}) = \mathrm{T}[\mathbb{I}_{A_n}\mathrm{T}^n\mathbb{I}_{A_0} + \mathbb{I}_{A_n^c}\mathrm{T}^n\mathbb{I}_{A_0}] = \mathrm{T}[\mathbb{I}_{A_n} + \mathbb{I}_{A_n^c}\mathrm{T}^n\mathbb{I}_{A_0}], \tag{3}$$

where the last equality follows from the induction hypothesis H_n. It follows from the definition of A_{n+1} that $\mathbb{I}_{A_{n+1}}\mathrm{T}\mathbb{I}_{A_n} = \mathbb{I}_{A_{n+1}}$, and therefore

$$\begin{aligned}\mathbb{I}_{A_{n+1}} = \mathbb{I}_{A_{n+1}}\mathrm{T}[\mathbb{I}_{A_n} + \mathbb{I}_{A_n^c}\mathrm{T}^n\mathbb{I}_{A_0} - \mathbb{I}_{A_n^c}\mathrm{T}^n\mathbb{I}_{A_0}] &\leq \mathbb{I}_{A_{n+1}}\mathrm{T}[\mathbb{I}_{A_n} + \mathbb{I}_{A_n^c}\mathrm{T}^n\mathbb{I}_{A_0}] + \mathbb{I}_{A_{n+1}}\mathrm{T}[-\mathbb{I}_{A_n^c}\mathrm{T}^n\mathbb{I}_{A_0}] \\ &= \mathbb{I}_{A_{n+1}}\mathrm{T}^{n+1}\mathbb{I}_{A_0} + \mathbb{I}_{A_{n+1}}\mathrm{T}[-\mathbb{I}_{A_n^c}\mathrm{T}^n\mathbb{I}_{A_0}] \\ &\leq \mathbb{I}_{A_{n+1}}\mathrm{T}^{n+1}\mathbb{I}_{A_0} \leq \mathbb{I}_{A_{n+1}},\end{aligned}$$

where the first inequality follows from (T2), the second from the fact that $-\mathbb{I}_{A_n^c}\mathrm{T}^n\mathbb{I}_{A_0} \leq 0$ and therefore $\mathbb{I}_{A_{n+1}}\mathrm{T}[-\mathbb{I}_{A_n^c}\mathrm{T}^n\mathbb{I}_{A_0}] \leq 0$ [use (T1) and (T5)], and the third from $\mathrm{T}^{n+1}\mathbb{I}_{A_0} \leq 1$ [use (T5)]. The second equality follows from Eq. (3). Hence indeed $\mathbb{I}_{A_{n+1}} = \mathbb{I}_{A_{n+1}}\mathrm{T}^{n+1}\mathbb{I}_{A_0}$.

Next, observe that $A_{n+2}^c = A_{n+1}^c \cup A_{n+1} \setminus A_{n+2}$. By definition, $\mathrm{T}\mathbb{I}_{A_{n+1}}(a) < 1$ for all $a \in A_{n+1} \setminus A_{n+2}$. It also follows from the induction hypothesis H_n that $\mathrm{T}\mathbb{I}_{A_n}(a) < 1$ for all $a \in A_{n+1}^c$. But since $A_{n+1} \subseteq A_n$, it follows from (T5) that $\mathrm{T}\mathbb{I}_{A_{n+1}} \leq \mathrm{T}\mathbb{I}_{A_n}$, and therefore also $\mathrm{T}\mathbb{I}_{A_{n+1}}(a) < 1$ for all $a \in A_{n+1}^c$. Hence indeed $\mathrm{T}\mathbb{I}_{A_{n+1}}(a) < 1$ for all $a \in A_{n+2}^c$.

To finish the induction proof, let $\beta := \max_{b \in A_n^c}\mathrm{T}^n\mathbb{I}_{\mathscr{R}^c}(a)$, then $\beta < 1$ by the induction hypothesis H_n. Since $\mathbb{I}_{A_0^c}\mathrm{T}^n\mathbb{I}_{A_0} = \mathbb{I}_{\mathscr{R}}\mathrm{T}^n\mathbb{I}_{\mathscr{R}^c} = 0$ [because no element in the maximal class $\mathscr{R}$ has access to any element outside it], we infer from Eq. (3) that

$$\mathrm{T}^{n+1}\mathbb{I}_{A_0} = \mathrm{T}[\mathbb{I}_{A_n} + \mathbb{I}_{A_n^c}\mathrm{T}^n\mathbb{I}_{A_0}] \leq \mathrm{T}[\mathbb{I}_{A_n} + \beta\mathbb{I}_{A_n^c}] = \mathrm{T}[\beta + (1-\beta)\mathbb{I}_{A_n}] \leq \beta + (1-\beta)\mathrm{T}\mathbb{I}_{A_n}.$$

Consider any $a \in A_{n+1}^c$, then by definition $\mathrm{T}\mathbb{I}_{A_n}(a) < 1$ by the induction hypothesis H_n, and therefore $\mathrm{T}^{n+1}\mathbb{I}_{A_0}(a) \leq \beta + (1-\beta)\mathrm{T}\mathbb{I}_{A_n}(a) < 1$ since also $\beta < 1$. We conclude that H_{n+1} holds too.

To continue the proof, we observe that $A_0, A_1, \dots, A_n, \dots$ is a non-increasing sequence, and that A_0 is finite. This implies that there must be some first $k \in \mathbb{N}$ such that $A_{k+1} = A_k$. Clearly, $k \leq |A_0|$. We now prove by induction that $G_n: \mathbb{I}_{A_k}\mathrm{T}^{n+k}\mathbb{I}_{A_0} = \mathbb{I}_{A_k}$ for all $n \geq 0$. The statement G_n clearly holds for $n = 0$. We show that $G_n \Rightarrow G_{n+1}$. First of all,

$$\mathrm{T}^{n+k+1}\mathbb{I}_{A_0} = \mathrm{T}(\mathrm{T}^{n+k}\mathbb{I}_{A_0}) = \mathrm{T}[\mathbb{I}_{A_k}\mathrm{T}^{n+k}\mathbb{I}_{A_0} + \mathbb{I}_{A_k^c}\mathrm{T}^{n+k}\mathbb{I}_{A_0}] = \mathrm{T}[\mathbb{I}_{A_k} + \mathbb{I}_{A_k^c}\mathrm{T}^{n+k}\mathbb{I}_{A_0}],$$

where the last equality follows from the induction hypothesis G_n. As before, it follows from the definition of A_{k+1} that $\mathbb{I}_{A_{k+1}}\mathrm{T}\mathbb{I}_{A_k} = \mathbb{I}_{A_{k+1}}$, and therefore $\mathbb{I}_{A_k}\mathrm{T}\mathbb{I}_{A_k} = \mathbb{I}_{A_k}$, so

$$\begin{aligned}\mathbb{I}_{A_k} = \mathbb{I}_{A_k}\mathrm{T}[\mathbb{I}_{A_k} + \mathbb{I}_{A_k^c}\mathrm{T}^{n+k}\mathbb{I}_{A_0} - \mathbb{I}_{A_k^c}\mathrm{T}^{n+k}\mathbb{I}_{A_0}] &\leq \mathbb{I}_{A_k}\mathrm{T}[\mathbb{I}_{A_nk} + \mathbb{I}_{A_k^c}\mathrm{T}^{n+k}\mathbb{I}_{A_0}] + \mathbb{I}_{A_k}\mathrm{T}[-\mathbb{I}_{A_k^c}\mathrm{T}^{n+k}\mathbb{I}_{A_0}] \\ &= \mathbb{I}_{A_k}\mathrm{T}^{n+k+1}\mathbb{I}_{A_0} + \mathbb{I}_{A_k}\mathrm{T}[-\mathbb{I}_{A_k^c}\mathrm{T}^{n+k}\mathbb{I}_{A_0}] \\ &\leq \mathbb{I}_{A_k}\mathrm{T}^{n+k+1}\mathbb{I}_{A_0} \leq \mathbb{I}_{A_k},\end{aligned}$$

where the first inequality follows from (T2), the second from the fact that $-\mathbb{I}_{A_k^c}\mathrm{T}^{n+k}\mathbb{I}_{A_0} \leq 0$ and therefore $\mathbb{I}_{A_k}\mathrm{T}[-\mathbb{I}_{A_k^c}\mathrm{T}^{n+k}\mathbb{I}_{A_0}] \leq 0$ [use (T1) and (T5)], and the third from $\mathrm{T}^{n+k+1}\mathbb{I}_{A_0} \leq 1$ [use (T5)]. Hence indeed $\mathbb{I}_{A_k} = \mathbb{I}_{A_k}\mathrm{T}^{n+k+1}\mathbb{I}_{A_0}$.

There are now two possibilities. The first is that $A_k \neq \emptyset$. It follows from the arguments above that for any element a of A_k, $\mathrm{T}^\ell\mathbb{I}_{\mathscr{R}^c}(a) = 1$ for all $\ell \in \mathbb{N}$, which implies that T cannot be top class absorbing. The second possibility is that $A_k = \emptyset$. It follows from the argument above that $\mathrm{T}^k\mathbb{I}_{\mathscr{R}^c}(a) < 1$ for all $a \in A_k^c = \mathscr{X}$ which implies that T is top class absorbing. □

Example 2. Define $\mathrm{T}f = \max\{Mf\colon L \le M \le U, M\mathbb{I}_{\mathscr{X}} = \mathbb{I}_{\mathscr{X}}\}$ where L and U are given by

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1/4 & 1/4 & 0 & 0 \\ 1/2 & 1/4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 1/4 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 3/4 & 1/2 & 0 & 0 \\ 3/4 & 1/2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1/4 & 3/4 & 0 & 0 & 1/4 \end{pmatrix}.$$

In this particular case, any representing matrix' non zero elements correspond to those of U. For example

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

The characteric function of M is

$$\chi_M(s) = (s-1)(s-1/3)^2(s-(1+\sqrt{7})/6)(s-(1-\sqrt{7})/6)$$

whence T is top class regular and from the form of M it can be seen immediately that the first element of the state vector is absorbing which implies that $\mathbb{I}_{\mathscr{R}} = \begin{pmatrix}1 & 0 & 0 & 0 & 0\end{pmatrix}^T$.

To check for top class absorbingness, we start iterating:

(itr 1) $\mathrm{T}\mathbb{I}_{\mathscr{R}^c} = \begin{pmatrix}0 & 1 & 1/2 & 1 & 1\end{pmatrix}^T$ whence $\mathbb{I}_{A_1} = \begin{pmatrix}0 & 1 & 0 & 1 & 1\end{pmatrix}^T$,
(itr 2) $\mathrm{T}\mathbb{I}_{A_1} = \begin{pmatrix}0 & 3/4 & 1/2 & 1 & 1\end{pmatrix}^T$ whence $\mathbb{I}_{A_2} = \begin{pmatrix}0 & 0 & 0 & 1 & 1\end{pmatrix}^T$,
(itr 3) $\mathrm{T}\mathbb{I}_{A_2} = \begin{pmatrix}0 & 0 & 0 & 1 & 1/4\end{pmatrix}^T$ whence $\mathbb{I}_{A_3} = \begin{pmatrix}0 & 0 & 0 & 1 & 0\end{pmatrix}^T$,
(itr 4) $\mathrm{T}\mathbb{I}_{A_3} = \begin{pmatrix}0 & 0 & 0 & 1 & 0\end{pmatrix}^T$ whence $\mathbb{I}_{A_4} = \begin{pmatrix}0 & 0 & 0 & 1 & 0\end{pmatrix}^T$.

Because $A_4 = A_3 \neq \emptyset$ we conclude that T is not top class absorbing and therefore not ergodic.

4. Coefficient of ergodicity

It is clear that ergodicity would follow immediately from Banach's fixed point theorem if T were contractive instead of non-expansive. With this in mind, one might think that conditions for ergodicity might coincide with contractiveness of T. This is not true as the next example shows.

Example 3. Consider the upper transition operator $\mathrm{T} = \mathbb{I}_{\mathscr{X}} \max$. Since $\mathrm{T}^k\mathbb{I}_{\{x\}} = \mathbb{I}_{\mathscr{X}}$ for all $x \in \mathscr{X}$ and all $k \in \mathbb{N}$, we find that $\mathscr{R} = \{x \in \mathscr{X}\colon (\exists n \in \mathbb{N})(\forall k \ge n) \min \mathrm{T}^k\mathbb{I}_{\{x\}} > 0\} = \mathscr{X}$, so T is both top class regular [because $\mathscr{R} \neq \emptyset$] and top class absorbing [trivially because $\mathscr{R}^c = \emptyset$], and therefore ergodic by Prop. 3. But T is non-expansive and not contractive, since $\|\mathrm{T}f\|_\infty = |\max f| \le \max|f| = \|f\|_\infty$, and therefore $\|\mathrm{T}f\|_\infty = \|f\|_\infty$ if $f \ge 0$.

Also, we should not aim for the uniqueness of the fixed point as we know that all constant functions on $\mathscr{L}(\mathscr{X})$ are fixed points. Under slight modifications, the reasoning developed in the contraction mapping theorem can still be used however. In the first stage, it is proven that

$$\left\|\mathrm{T}^{k+1}f - \mathrm{T}^k f\right\|_\infty \le \alpha_f q^k, \tag{4}$$

with α_f a constant depending on f and $0 \le q < 1$. Subsequently, equation 5 is used to show that $\{\mathrm{T}^n f\}$ is a Cauchy sequence and because of the compactness and hence completeness of the space it is concluded that $\lim_{k\to\infty}\mathrm{T}^k f$ exists and we will denote it $\mathrm{T}_\infty f$. Clearly also $\mathrm{T}_\infty f$ is a fixed point of T. By using exactly the same reasoning as used to prove equation 5 plus the fact that $\mathrm{T}_\infty\mathrm{T}^k f = \mathrm{T}_\infty f$ it follows that

$$\left\|\mathrm{T}^k f - \mathrm{T}_\infty f\right\|_\infty \le \beta_f q^k. \tag{5}$$

If we are looking for ergodicity then we do not only want the sequence $\{\mathrm{T}^k f\}$ to converge, but also that $\max \mathrm{T}_\infty f = \min \mathrm{T}_\infty f$. To join this requirement with the contraction assumption, Škulj and Hable introduced the coefficient of ergodicity [7] which they defined as

$$\rho(\mathrm{T}) := \max_{0\le f\le 1} (\max \mathrm{T}f - \min \mathrm{T}f).$$

By doing so, the next proposition –which is very similar to the first step in the contraction theorem– can be shown.

Proposition 7. *If there exist $n \in \mathbb{N}$ such that $\rho(\mathrm{T}^n) < 1$ then the next inequality holds for $k \ge n$:*

$$\left\| \mathrm{T}^{k+1} f - \mathrm{T}^k f \right\|_\infty \le \alpha q^k,$$

where $q := \rho(\mathrm{T}^n)^{1/n}$ and $\alpha := \rho(\mathrm{T}^n)^{-1}(\max f - \min f)\max_{0\le g\le 1} \|g - \mathrm{T}g\|_\infty$.

Proof. We decompose $k = qn + r$ where r is the remainder when dividing k by n. Because T_∞ is T-invariant

$$\begin{aligned}\left\| \mathrm{T}^{k+1} f - \mathrm{T}^k f \right\|_\infty &= \left\| \mathrm{T}^{nq+r+1} f - \mathrm{T}^{nq+r} f \right\|_\infty,\\ &\le \left\| \mathrm{T}^{nq+1} f - \mathrm{T}^{nq} f \right\|_\infty,\end{aligned}$$

where the last step is a result from (T8). If we define $\tilde{f} := \frac{\mathrm{T}^{nq} f - \min \mathrm{T}^{nq} f}{\max \mathrm{T}^{nq} f - \min \mathrm{T}^{nq} f}$ then we see that

$$\begin{aligned}\left\| \mathrm{T}^{nq+1} f - \mathrm{T}^{nq} f \right\|_\infty &= (\max \mathrm{T}^{nq} f - \min \mathrm{T}^{nq} f) \left\| \mathrm{T}^{nq+1} \tilde{f} - \mathrm{T}^{nq} \tilde{f} \right\|_\infty,\\ &\le \frac{\max \mathrm{T}^{nq} f - \min \mathrm{T}^{nq} f}{\max f - \min f} (\max f - \min f) \left\| \mathrm{T}\tilde{f} - \tilde{f} \right\|_\infty,\\ &\le \rho(\mathrm{T}^{nq})(\max f - \min f) \max_{0\le g\le 1} \|\mathrm{T}g - g\|_\infty.\end{aligned}$$

Here we used the positive homogeneity of $\|\cdot\|_\infty$, the non-expansiveness of T and the fact that $\max \mathrm{T}^{nq} f - \min \mathrm{T}^{nq} f \le \max f - \min f$.

A close examination of the coefficient of ergodicity shows us that

$$\begin{aligned}\rho(\mathrm{T}^{nq}) &= \max_{0\le f\le 1} \Big(\max \mathrm{T}[\mathrm{T}^{n(q-1)} f - \min \mathrm{T}^{n(q-1)} f] - \min \mathrm{T}[\mathrm{T}^{n(q-1)} f - \min \mathrm{T}^{n(q-1)} f] \Big),\\ &\le \max_{0\le f\le 1} [\mathrm{T}^{n(q-1)} f - \min \mathrm{T}^{n(q-1)} f] \max_{0\le g\le 1} (\max \mathrm{T}^n g - \min \mathrm{T}^n g),\\ &\le \max_{0\le f\le 1} [\max \mathrm{T}^{n(q-1)} f - \min \mathrm{T}^{n(q-1)} f] \max_{0\le g\le 1} (\max \mathrm{T}^n g - \min \mathrm{T}^n g),\\ &= \rho(\mathrm{T}^{n(q-1)})\,\rho(\mathrm{T}^n),\end{aligned}$$

which we can use as the inductive step to show that $\rho(\mathrm{T}^{nq}) \le \rho(\mathrm{T}^n)^q$. To conclude the proof we notice that

$$\rho(\mathrm{T}^{nq}) \le \rho(\mathrm{T}^n)^q = \rho(\mathrm{T}^n)^{\frac{1}{n}nq} = \rho(\mathrm{T}^n)^{-\frac{r}{n}} \rho(\mathrm{T}^n)^{\frac{1}{n}(nq+r)} \le \rho(\mathrm{T}^n)^{-1} \rho(\mathrm{T}^n)^{\frac{1}{n}k}$$

□

From the last line of the proof it follows moreover that if there exists n such that $\rho(\mathrm{T}^n) < 1$ then $\lim_{k\to\infty} \rho(\mathrm{T}^k) = 0$ which means that $\max \mathrm{T}_\infty f - \min \mathrm{T}_\infty f = 0$ and thus that T must be ergodic.

As explained before, the previous proposition can be used to say something about the speed of convergence or in other words that all fixed points are constant functions. Consequently, $\rho(\mathrm{T}^n) < 1$ is a sufficient condition for ergodicity. Similar to the previous proof and using $\mathrm{T}_\infty \mathrm{T} = \mathrm{T}\mathrm{T}_\infty = \mathrm{T}\infty$ we get the speed of convergence result.

Proposition 8. *l If there exist $n \in \mathbb{N}$ such that $\rho(\mathrm{T}^n) < 1$ then* T *is ergodic and for all $k \ge n$*

$$\left\| \mathrm{T}^k f - \mathrm{T}_\infty f \right\|_\infty \le \beta q^k,$$

where $q := \rho(\mathrm{T}^n)^{1/n}$ *and* $\beta := \rho(\mathrm{T}^n)^{-1}(\max f - \min f) \max_{0 \le g \le 1} \|g - \mathrm{T}_\infty g\|_\infty$.

5. CONCLUSION

In this paper we gave different conditions under which an upper transition operator –which corresponds to a set of non-stationary Markov chains– is ergodic. We showed that ergodicity is completely determined by the eigenvalues and functions of the transition operator as is the case in classical Markov chains which opens the door to a spectral theorem for upper transition operators. Unfortunately, it is not known how to calculate these eigenvalues which is why we developed an alternative test for ergodicity which needs at most $2|\mathscr{X}| - 1$ evaluations of the upper transition operator. The algorithm consists of two steps in which the first step checks for top class regularity which is done by building a representing stochastic matrix and solving a linear eigenvalue problem and a second step which checks for top class absorption.

Another approach that has been documented in the literature calculates the coefficient of ergodicity and checks whether there exists a power of the transition operators such that this coefficient becomes strictly smaller than zero. Because calculating this coefficient involves an optimization problem over an hypercube of the space and because it is not known whether there exists a maximal power till which the coefficient needs to be checked, the coefficient of ergodicity is merely a theoretical measure. Interesting about the coefficient of ergodicity however is that it gives an upper bound on the speed of convergence.

What has not been investigated yet are the conditions under which $\{\mathrm{T}^n f\}$ will converge in general. Extrapolating from the conditions for convergence given in this paper, one might conjecture that there is convergence if and only if all classes that are not regular are absorbed by a union of classes that are regular. Also worth investigating is whether there are possibilities to get stochastic matrices that are representing the upper transition operator in a quantitative way rather than in a qualitative way as the ones we defined. These matrices could then for example be used to make statements about the rate of convergence from individual elements to the absorbing class.

ACKNOWLEDGEMENTS

This paper presents research results of the Belgian programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its authors.

REFERENCES

[1] Gert de Cooman, Filip Hermans, and Erik Quaeghebeur. Imprecise markov chains and their limit behavior. *Probability in the Engineering and Informational Sciences*, 23(4):597–635, 2009.
[2] Darald J. Hartfiel. *Markov Set-Chains*. Number 1695 in Lecture Notes in Mathematics. Springer, Berlin, 1998.
[3] John G. Kemeny and J. Laurie Snell. *Finite Markov Chains*. Undergraduate Text in Mathematics. Springer-Verlag, New York, 1976. Reprint of the 1960 Edition.
[4] B Lemmens and M Scheutzow. On the dynamics of sup-norm non-expansive maps. *Ergod Theor Dyn Syst*, 25:861–871, Jan 2005.
[5] E. Seneta. *Non-negative Matrices and Markov Chains (Springer Series in Statistics)*. Springer, January 2006.
[6] R Sine. A nonlinear Perron–Frobenius theorem. *Proceedings of the American Mathematical Society*, 109(2):331–336, January 1990.
[7] Damjan Škulj and Robert Hable. Coefficients of ergodicity for imprecise markov chains. In Thomas Augustin, Frank P.A. Coolen, Moral Serafin, and Matthias C.M. Troffaes, editors, *ISIPTA '09 – Proceedings of the Sixth International Symposium on Imprecise Probability: Theories and Applications*, pages 377–386. SIPTA, 2009.

SYSTeMS RESEARCH GROUP, GHENT UNIVERSITY, TECHNOLOGIEPARK–ZWIJNAARDE 914, 9052 ZWIJNAARDE, BELGIUM
E-mail address: `{filip.hermans,gert.decooman}@UGent.be`