A convenient characterisation of convergent upper transition operators

Jasper De Bock[1] **Alexander Erreygers**[1] **Floris Persiau**[2]

[1]Foundations Lab for imprecise probabilities, Ghent University, Belgium
[2]Department of Philosophy, Carnegie Mellon University, United States of America

Abstract

Motivated by its connection to the limit behaviour of imprecise Markov chains, we introduce and study the so-called convergence of upper transition operators: the condition that for any function, the orbit resulting from iterated application of this operator converges. In contrast, the existing notion of 'ergodicity' requires convergence of the orbit to a constant. We derive a very general (and practically verifiable) sufficient condition for convergence in terms of accessibility and lower reachability, and prove that this sufficient condition is also necessary whenever (i) all transient states are absorbed or (ii) the upper transition operator is finitely generated.

Keywords. imprecise Markov chains, upper transition operators, convergence, ergodicity, regularity

1. Introduction

Imprecise Markov chains [3–5, 7, 10, 14] model the uncertain temporal evolution of the state of finite-state discrete-time systems, and they do so in a more general manner than ('precise') Markov chains [9, 12] by allowing for partial probability specifications. Let $\mathcal{X}$ be the finite set of possible states of the system under study. A Markov chain models uncertain dynamics using transition probabilities: for every state $x \in \mathcal{X}$, one has to specify a transition probability mass function (pmf) p_x, and then $p_x(y)$ is the probability of transitioning from state x in the current time step to state y in the next one. These give rise to a *transition operator* (or *kernel*) $T: \mathbb{R}^{\mathcal{X}} \to \mathbb{R}^{\mathcal{X}}$, which maps every function $f \in \mathbb{R}^{\mathcal{X}}$ to

$$Tf : \mathcal{X} \to \mathbb{R} : x \mapsto \mathrm{E}_{p_x}(f).$$

This transition operator comes in handy when determining the expectation of a function f of the state after n time steps, given that the system started in state x:

$$\mathrm{E}\big(f(X_n) \mid X_0 = x\big) = T^n f(x).$$

In contrast, an imprecise Markov chain models uncertain dynamics using a family $(\mathcal{P}^x)_{x\in\mathcal{X}}$ of candidate transition pmfs, with the interpretation that if up to time step n the states were $x_0, \dots, x_n$, the uncertainty about the state X_{n+1} in the next time step $n+1$ is accurately modelled by *some* pmf in $\mathcal{P}^{x_n}$, which, depending on the adopted interpretation, may either depend on $x_0, \dots, x_n$ or only on x_n. Both interpretations give rise to the same range of values for the conditional expectation of a function f of the state after n time steps [7, Theorem 11.4], and one is then typically interested in the supremum $\overline{\mathrm{E}}\big(f(X_n) \mid X_0 = x\big)$ or infimum $\underline{\mathrm{E}}\big(f(X_n) \mid X_0 = x\big)$ of this range, called the *upper and lower expectation*, with

$$\begin{aligned}\overline{\mathrm{E}}\big(f(X_n) \mid X_0 = x\big) &= \overline{T}^n f(x),\\ &= -\underline{\mathrm{E}}\big(-f(X_n) \mid X_0 = x\big),\end{aligned}$$

where $\overline{T}$ is the chain's *upper transition operator*, which maps every function $f \in \mathbb{R}^{\mathcal{X}}$ to

$$\overline{T}f : \mathcal{X} \to \mathbb{R} : x \mapsto \sup\{\mathrm{E}_p(f) : p \in \mathcal{P}^x\}. \tag{1}$$

Studying the limit behaviour of imprecise Markov chains is therefore a matter of understanding the limit behaviour of the orbit $(\overline{T}^n f)_{n\in\mathbb{N}}$. In this contribution, we focus on the convergence of these orbits and develop conditions that guarantee convergence for all $f \in \mathbb{R}^{\mathcal{X}}$.

In order to adhere to the page limit, we've omitted the proof of an intermediary lemma whenever it's straightforward or non-instructive. Furthermore, we also don't give a proof for Lemma 3.1 and Proposition 5.2, since their proofs are quite long and essentially already appear in [6, Proposition 6] and [2, Theorem 5.28], respectively. The interested reader will be happy to find these omitted proofs in the arXiv:2502.04509 preprint of this contribution.

2. Upper transition operators

Since in Equation (1) the upper transition operator $\overline{T}$ is defined as a pointwise supremum over linear operators, it has the following properties:

T1. $\overline{T}(f+g) \leq \overline{T}f + \overline{T}g$ for all $f, g \in \mathbb{R}^{\mathcal{X}}$;

T2. $\overline{T}(\lambda f) = \lambda \overline{T} f$ for all $f \in \mathbb{R}^{\mathcal{X}}$ and $\lambda \in \mathbb{R}_{\geq 0}$;

T3. $\overline{T} f \leq \max f$ for all $f \in \mathbb{R}^{\mathcal{X}}$.

In general, we call any operator $\overline{T}\colon \mathbb{R}^{\mathcal{X}} \to \mathbb{R}^{\mathcal{X}}$ that satisfies T1–T3 an *upper transition operator*,[1] as it can always be thought of as being derived from sets of pmfs: T1–T3 guarantee that for all $x \in \mathcal{X}$, there is at least one *compatible* set of pmfs $\mathcal{P}^x$, in the sense that it satisfies the assignment in Equation (1). There is also a largest one, and this so-called *credal set* $\mathcal{Q}^x$ is the unique closed and convex set of pmfs that is compatible with $\overline{T}$.

2.1. Convergence and ergodicity. As explained in the Introduction, we aim to obtain (necessary and/or sufficient) conditions on the upper transition operator $\overline{T}$ for it to have the following property.

Definition 2.1. An upper transition operator $\overline{T}$ is *convergent* if for all $f \in \mathbb{R}^{\mathcal{X}}$, $(\overline{T}^n f)_{n\in\mathbb{N}}$ converges.

Although we are not the first to study the limit behaviour of orbits $(\overline{T}^n f)_{n\in\mathbb{N}}$, we are not aware of any work that focuses on (what we call) convergence. Most work—at least on upper transition operators for imprecise Markov chains—focuses on the stronger requirement of ergodicity [4, 6, 16]; Škulj [15] does consider non-ergodic imprecise Markov chains, but focusses on invariant sets of distributions.

Definition 2.2. An upper transition operator $\overline{T}$ is *ergodic* if for all $f \in \mathbb{R}^{\mathcal{X}}$, $(\overline{T}^n f)_{n\in\mathbb{N}}$ converges to a constant.

Hermans and De Cooman [6, Proposition 3] study this notion of ergodicity extensively, and they obtain a necessary and sufficient condition that is easy to check, which we'll introduce in Section 3 further on. We aim for similar conditions, but for convergence rather than ergodicity.

Hermans and De Cooman [6] also explain that thanks to T1–T3, an upper transition operator $\overline{T}$ is a '(convex) topical map' or, more generally, a 'sup-norm non-expansive map' [11, 13]. For these more general types of maps, the limit behaviour of orbits has been studied extensively as well [1, 11, 13]. From these references, we know that for any $f \in \mathbb{R}^{\mathcal{X}}$, the orbit $(\overline{T}^n f)_{n\in\mathbb{N}}$ has a finite limit set $\Omega_f = \{\omega_1, \dots, \omega_{p_f}\}$,[2] whose *period* (or cardinality) p_f has a universal upper bound that depends only on the size of $\mathcal{X}$ and whose elements form a cycle: $\omega_2 = \overline{T}\omega_1, \dots, \omega_{p_f} = \overline{T}\omega_{p_f-1}$ and $\omega_{p_f+1} \coloneqq \omega_1 = \overline{T}\omega_{p_f}$. Convergence requires that $p_f = 1$ for all $f \in \mathbb{R}^{\mathcal{X}}$.

2.2. Convenient properties of upper transition operators. Throughout this contribution, we'll make use of several properties of upper transition operators that are well known [6, 17], but which we repeat here for the sake of convenience. An upper transition operator $\overline{T}$ has a corresponding conjugate *lower transition operator* $\underline{T}\colon \mathbb{R}^{\mathcal{X}} \to \mathbb{R}^{\mathcal{X}}$ defined by $\underline{T}f \coloneqq -\overline{T}(-f)$ for all $f \in \mathbb{R}^{\mathcal{X}}$; these names make sense because $\underline{T}f \leq \overline{T}f$ for all $f \in \mathbb{R}^{\mathcal{X}}$. We'll often implicitly use that if $\overline{T}$ is an upper transition operator, then so is its n-fold composition $\overline{T}^n$—which has $\underline{T}^n$ as conjugate lower transition operator. Other important properties of $S \in \{\overline{T}, \underline{T}\}$ are:

T4. $\min f \leq Sf \leq \max f$ for all $f \in \mathbb{R}^{\mathcal{X}}$;

T5. if $f \leq g$ then $Sf \leq Sg$ for all $f, g \in \mathbb{R}^{\mathcal{X}}$;

T6. $S(\mu + f) = \mu + Sf$ for all $f \in \mathbb{R}^{\mathcal{X}}$ and $\mu \in \mathbb{R}$;

T7. $(\max f - \min f)S\mathbb{1}_x + \min f \leq Sf$ for all $f \in \mathbb{R}^{\mathcal{X}}$ and $x \in \arg\max f$.

One particular type of functions we'll use are indicators: the *indicator* $\mathbb{1}_A$ of the set $A \subseteq \mathcal{X}$ maps $x \in \mathcal{X}$ to 1 if $x \in A$ and to 0 otherwise; to ease our notation, we write $\mathbb{1}_x \coloneqq \mathbb{1}_{\{x\}}$ for all $x \in \mathcal{X}$. The final property we list revolves around indicators:

T8. $\overline{T}\mathbb{1}_C = 1 - \underline{T}\mathbb{1}_{\mathcal{X}\setminus C}$ for all $C \subseteq \mathcal{X}$.

3. Characterising ergodicity

The necessary and sufficient condition for ergodicity given by Hermans and De Cooman [6, Proposition 3] is stated in terms of accessibility and reachability relations. Since these relations will also be important in our quest for conditions for convergence, we'll briefly recall them before repeating their result. Throughout this contribution, we'll illustrate most relations and notions with the following running example.

Running example 1. Let $\mathcal{X} \coloneqq \{\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}, \mathsf{e}\}$ and consider the upper transition operator $\overline{T}$ induced by the sets

$$\mathcal{P}^{\mathsf{a}} \coloneqq \{\mathbb{1}_{\mathsf{a}}\},\ \mathcal{P}^{\mathsf{b}} \coloneqq \{\mathbb{1}_{\mathsf{b}}\},\ \mathcal{P}^{\mathsf{c}} \coloneqq \left\{\frac{1}{4}(\mathbb{1}_{\mathsf{a}} + \mathbb{1}_{\mathsf{b}} + \mathbb{1}_{\mathsf{d}} + \mathbb{1}_{\mathsf{e}})\right\}$$
$$\text{and } \mathcal{P}^{\mathsf{d}} \coloneqq \{\mathbb{1}_{\mathsf{c}}, \mathbb{1}_{\mathsf{d}}, \mathbb{1}_{\mathsf{e}}\} \eqqcolon \mathcal{P}^{\mathsf{e}}. \quad \diamond$$

3.1. The upper accessibility graph. An upper transition operator $\overline{T}$ gives rise to a corresponding *upper accessibility graph* $\overline{\mathcal{G}}(\overline{T})$, with $\mathcal{X}$ as nodes and a directed edge between states $x, y \in \mathcal{X}$ if and only if $\overline{T}\mathbb{1}_y(x) > 0$ [6, Definition 5]. A state $y \in \mathcal{X}$ is now said to be *accessible* from a state $x \in \mathcal{X}$ if either $y = x$ or there is a directed path from x to y in $\overline{\mathcal{G}}(\overline{T})$, or equivalently, if $\overline{T}^n\mathbb{1}_y(x) > 0$ for some $n \in \mathbb{Z}_{\geq 0}$ [6, Definition 4 and Proposition 4].

Running example 2. The reader will have no difficulty in verifying that Figure 1 depicts the upper accessibility graph $\overline{\mathcal{G}}(\overline{T})$. ⋄

Two states $x, y \in \mathcal{X}$ *communicate* if y is accessible from x (in $\overline{\mathcal{G}}(\overline{T})$) and vice versa; this equivalence relation partitions the state space $\mathcal{X}$ into equivalence classes C_1, …, C_n, aptly called *communication classes*—or strongly connected components in the theory of directed graphs. For two communication classes C_k and C_ℓ, $y \in C_\ell$ is accessible from $x \in C_k$ if and only if the same is true

[1] Also known as a *sublinear transition operator* or a *sublinear kernel* [5, Definition 5.1].

[2] Defined as the set of accumulation points of the orbit $(\overline{T}^n f)_{n\in\mathbb{N}}$, or equivalently, the limits of the convergent subsequences.

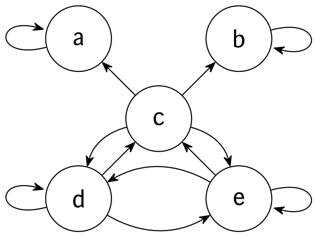


Figure 1. *$\overline{\mathcal{G}}(\overline{T})$ for Running example 1.*

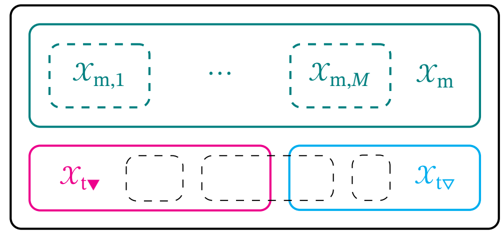


Figure 2. *Venn diagram of the state space $\mathcal{X}$*

for any $y' \in C_\ell$ and $x' \in C_k$; whenever this is the case, we therefore simply say that C_ℓ is accessible from C_k. This accessibility relation induces a partial order on the communication classes: we say that C_ℓ dominates C_k if $C_\ell \neq C_k$ and C_ℓ is accessible from C_k. A communication class C is called *maximal* (or final) if it's undominated with respect to this partial order. Since there are only finitely many communication classes, every non-maximal communication class is dominated by at least one maximal communication class. A state is called *maximal* if it belongs to a maximal communication class and *non-maximal* or *transient* otherwise. We enumerate the maximal communication classes by $\mathcal{X}_{\mathrm{m},1}, \dots, \mathcal{X}_{\mathrm{m},M}$, and we collect their union in $\mathcal{X}_{\mathrm{m}} := \mathcal{X}_{\mathrm{m},1} \cup \cdots \cup \mathcal{X}_{\mathrm{m},M}$.

The *cyclicity* or *period* of a maximal communication class $\mathcal{X}_{\mathrm{m},k}$ is the greatest common divisor of the lengths of the closed directed paths that remain in this class, and the class $\mathcal{X}_{\mathrm{m},k}$ is said to be *regular* if it has cyclicity 1, or equivalently, if there is some $N \in \mathbb{N}$ such that for all $n \geq N$ and $x, y \in \mathcal{X}_{\mathrm{m},k}$, $\overline{T}^n \mathbb{1}_y(x) > 0$ [4, Proposition 4.2].

Running example 3. The graph $\overline{\mathcal{G}}(\overline{T})$ has three communication classes: $\{\mathsf{a}\}$, $\{\mathsf{b}\}$ and $\{\mathsf{c}, \mathsf{d}, \mathsf{e}\}$. The first two are the maximal ones, and these are regular because they obviously have cyclicity 1. ⋄

For small state spaces, it's easy to determine the maximal communication classes and their cyclicity on sight. In general, however, there's several of algorithms that do this (in linear time)—see for example [8, Sections 13.2.3 and 13.3.2].

To make our lives a bit easier, we'll call any non-empty subset C of $\mathcal{X}$ a *class*. Such a class C is *closed*—after T'Joens and De Bock [17, Section 5]—if $\overline{T}\mathbb{1}_y(x) = 0$ for all $x \in C$ and $y \in \mathcal{X} \setminus C$; it follows from T1 and T4 that this is the case if and only if $\overline{T}\mathbb{1}_{\mathcal{X}\setminus C}(x) = 0$ for all $x \in \mathcal{X}$. Clearly, a communication class C is maximal if and only if it's closed; we leave it to the reader to verify that any closed class must be a union of communication classes, but that not all such unions need be closed.

Running example 4. The closed classes in our running example are $\{\mathsf{a}\}$, $\{\mathsf{b}\}$, $\{\mathsf{a}, \mathsf{b}\}$ and $\mathcal{X}$. ⋄

3.2. Lower reachability and absorption. Unfortunately, the upper accessibility graph $\overline{\mathcal{G}}(\overline{T})$ does not suffice to characterise ergodicity; for this, we also need the notions of 'lower reachability' and 'absorption'. A class C is *lower reachable* from a state $x \in \mathcal{X}$ if there is some $n \in \mathbb{N}$ such that $\underline{T}^n \mathbb{1}_C(x) > 0$, and we call C *absorbing* if it is lower reachable from any state $x \in \mathcal{X} \setminus C$ [17, Section 5]. The following result—a slight generalisation of [6, Proposition 6]—provides a convenient recursive method to determine the states from which a *closed* class is lower reachable.

Lemma 3.1. *Consider an upper transition operator $\overline{T}$ with closed class C. Let $(C_n)_{n\in\mathbb{Z}_{\geq 0}}$ be the non-decreasing sequence given by $C_0 := C$ and, for all $n \in \mathbb{Z}_{\geq 0}$, by*

$$\begin{aligned} C_{n+1} &:= C_n \cup \{x \in \mathcal{X} \setminus C_n : \underline{T}\mathbb{1}_{C_n}(x) > 0\} \\ &= \{x \in \mathcal{X} : \underline{T}^{n+1}\mathbb{1}_C(x) > 0\}. \end{aligned}$$

Then after $k \leq |\mathcal{X} \setminus C|$ steps, we reach $C_k = C_{k+1}$, and C_k is the set of states from which C is lower reachable.

As $\mathcal{X}_{\mathrm{m}}$ is the union of the maximal (and therefore closed) communication classes, it is closed as well and we can determine the set C_k of states from which it is lower reachable with the iterative method in Lemma 3.1 for $C = \mathcal{X}_{\mathrm{m}}$. We collect all transient states of this kind in $\mathcal{X}_{\mathrm{t}\blacktriangledown} = C_k \setminus \mathcal{X}_{\mathrm{m}}$; finally, $\mathcal{X}_{\mathrm{t}\triangledown} := \mathcal{X} \setminus (\mathcal{X}_{\mathrm{m}} \cup \mathcal{X}_{\mathrm{t}\blacktriangledown})$ collects the transient states from which $\mathcal{X}_{\mathrm{m}}$ is *not* lower reachable. Note that, as depicted in Figure 2, $\mathcal{X}_{\mathrm{t}\blacktriangledown}$ and $\mathcal{X}_{\mathrm{t}\triangledown}$ need *not* be communication classes, and that $(\mathcal{X}_{\mathrm{m}}, \mathcal{X}_{\mathrm{t}\blacktriangledown}, \mathcal{X}_{\mathrm{t}\triangledown})$ partitions $\mathcal{X}$.

Running example 5. To determine which transient states $\mathcal{X}_{\mathrm{m}}$ is lower reachable from, we apply the recursive method in Lemma 3.1 for $C = \mathcal{X}_{\mathrm{m}} = \{\mathsf{a}, \mathsf{b}\}$:

$$C_1 = \{\mathsf{a}, \mathsf{b}\} \cup \{x \in \{\mathsf{c}, \mathsf{d}, \mathsf{e}\} : \underline{T}\mathbb{1}_{\{\mathsf{a},\mathsf{b}\}}(x) > 0\} = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$$

and

$$C_2 = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\} \cup \{x \in \{\mathsf{d}, \mathsf{e}\} : \underline{T}\mathbb{1}_{\{\mathsf{a},\mathsf{b},\mathsf{c}\}}(x) > 0\} = C_1.$$

Consequently, $\mathcal{X}_{\mathrm{m}}$ is lower reachable from c but not from d and e. This illustrates nicely that $\mathcal{X}_{\mathrm{t}\blacktriangledown} = \{\mathsf{c}\}$ and $\mathcal{X}_{\mathrm{t}\triangledown} = \{\mathsf{d}, \mathsf{e}\}$ need not be communication classes. ⋄

Crucially, the class $\mathcal{X}_{\mathrm{t}\triangledown}$ of transient states from which $\mathcal{X}_{\mathrm{m}}$ is not lower reachable is always empty for a (linear) transition operator T.

Lemma 3.2. *For any linear transition operator T, $\mathcal{X}_{\mathrm{m}}$ is absorbing, so $\mathcal{X}_{\mathrm{t}\blacktriangledown} = \mathcal{X} \setminus \mathcal{X}_{\mathrm{m}}$ and $\mathcal{X}_{\mathrm{t}\triangledown} = \emptyset$.*

Proof. The statement is trivial if $\mathcal{X}_{\mathrm{m}} = \mathcal{X}$, so assume that $\mathcal{X} \setminus \mathcal{X}_{\mathrm{m}} \neq \emptyset$. Any non-maximal $x \in \mathcal{X} \setminus \mathcal{X}_{\mathrm{m}}$ belongs to a non-maximal communication class C, for which we know that it is dominated by some maximal communication class $\mathcal{X}_{\mathrm{m},k}$. For any $y \in \mathcal{X}_{\mathrm{m},k} \subseteq \mathcal{X}_{\mathrm{m}}$, this implies that y is accessible from x, meaning that there is some $n \in \mathbb{N}$ such that $T^n \mathbb{1}_y(x) > 0$ and therefore also $T^n \mathbb{1}_{\mathcal{X}_{\mathrm{m}}}(x) \geq T^n \mathbb{1}_y(x) > 0$. Since T is linear, this shows that $\mathcal{X}_{\mathrm{m}}$ is lower reachable from any $x \in \mathcal{X} \setminus \mathcal{X}_{\mathrm{m}}$, whence $\mathcal{X}_{\mathrm{m}}$ is absorbing. □

3.3. A necessary and sufficient condition for ergodicity. And with that, we are ready to state the necessary and sufficient condition for ergodicity given by Hermans and De Cooman [6, Proposition 3]:

Proposition 3.1. *An upper transition operator $\overline{T}$ is ergodic if and only if it has a single maximal communication class (so $\mathcal{X}_{\mathrm{m}} = \mathcal{X}_{\mathrm{m},1}$) that is absorbing (so $\mathcal{X}_{\mathrm{t}\triangledown} = \emptyset$) and regular.*

Since ergodicity clearly implies convergence, this result provides a sufficient condition for convergence that can be checked easily. However, whenever there's more than one maximal communication class, as in our running example, we're already out of luck.

Running example 6. Since the upper transition operator $\overline{T}$ has two maximal communication classes, Proposition 3.1 tells us it cannot be ergodic. However, it *is* convergent! The reader may set out to verify this explicitly, but it is much more convenient to use the sufficient condition for convergence that we'll establish in Theorem 4.1 further on—see Running example 8. ⋄

4. Conditions for convergence

Our path forward is clear: we set out to establish a sufficient condition for convergence that is more general than the (necessary and sufficient) one for ergodicity in Proposition 3.1. We'll do so in several stages: we first obtain a necessary and sufficient condition in the case of a single communication class, then a necessary and sufficient condition when there's multiple maximal communication classes whose union is absorbing, and finally a sufficient condition in the general case.

4.1. A single communication class. Let us start gently with upper transition operators that have a single communication class—a linear transition operator of this type is called *irreducible* [8, Section 13.2.1]. We know from Proposition 3.1 that in this case regularity is sufficient for ergodicity and therefore convergence, and the following result establishes that it is also necessary for convergence.

Proposition 4.1. *Consider an upper transition operator $\overline{T}$ with a single communication class. Then the following three statements are equivalent:* (i) *$\overline{T}$ is convergent;* (ii) *$\overline{T}$ is ergodic;* (iii) *$\mathcal{X}$ is regular.*

Besides Proposition 3.1, our proof for this result relies on the following lemma, which will come in handy further on as well.

Lemma 4.1. *Consider an upper transition operator $\overline{T}$ with a single communication class. Then for any $f \in \mathbb{R}^{\mathcal{X}}$, $(\overline{T}^n f)_{n\in\mathbb{N}}$ converges if and only if it converges to a constant.*

Proof. It suffices to show that if $\phi := \lim_{n\to+\infty} \overline{T}^n f$ exists, it must be constant. So let us assume *ex absurdo* that ϕ exists but is not constant. Then there are $x, y \in \mathcal{X}$ such that $\phi(x) = \max \phi > \min \phi = \phi(y)$. Since $\mathcal{X}$ is a communication class, there is some $k \in \mathbb{N}$ such that $\overline{T}^k \mathbb{1}_x(y) > 0$. It therefore follows that

$$\overline{T}^k \phi(y) \overset{\text{T7}}{\geq} (\max \phi - \min \phi) \overline{T}^k \mathbb{1}_x(y) + \min \phi > \min \phi.$$

Meanwhile, $\overline{T}\phi = \phi$, and hence, $\phi(y) = \overline{T}^k \phi(y) > \min \phi$, which contradicts the fact that $\phi(y) = \min \phi$. □

Proving Proposition 4.1 is now a piece of cake.

Proof of Proposition 4.1. The equivalence of convergence and ergodicity is an immediate consequence of Lemma 4.1. That ergodicity is equivalent to regularity, on the other hand, follows directly from Proposition 3.1 □

Whenever one of the three equivalent conditions in Proposition 4.1 holds, we can also say something about the (constant) limit of $(\overline{T}^n f)_{n\in\mathbb{N}}$ in relation to f.

Proposition 4.2. *Consider an upper transition operator $\overline{T}$ such that $\mathcal{X}$ is a regular communication class. Then for any $f \in \mathbb{R}^{\mathcal{X}}$, $(\overline{T}^n f)_{n\in\mathbb{N}}$ converges to a constant function $\phi \geq \min f$ and, unless f is constant, $\phi > \min f$.*

Proof. Since $\mathcal{X}$ is a regular communication class, we know from Proposition 4.1 that $\overline{T}$ is ergodic, which implies that $(\overline{T}^n f)_{n\in\mathbb{N}}$ converges to a constant function ϕ.

Consider any $x \in \arg\max f$. Since $\mathcal{X}$ is regular, there is some $k \in \mathbb{N}$ such that $\min \overline{T}^k \mathbb{1}_x > 0$. Consequently, with $\alpha := \max f - \min f \geq 0$,

$$\phi = \lim_{n\to+\infty} \overline{T}^n \overline{T}^k f \overset{\text{T4}}{\geq} \min \overline{T}^k f \overset{\text{T7}}{\geq} \alpha \min \overline{T}^k \mathbb{1}_x + \min f.$$

Hence, indeed, $\phi \geq \min f$; if f is not constant, then $\alpha > 0$ and therefore $\phi > \min f$. □

4.2. Restriction. In the remainder of this section we move beyond upper transition operators with a single communication class. As a first step, instead of studying the convergence of $\overline{T}$ for all $x \in \mathcal{X}$, we zoom in on particular subsets of states.

To do so, we introduce the notion of restriction of functions, sets of pmfs and upper transition operators. For $A, B \subseteq \mathcal{X}$ with $A \supseteq B$ and $f \in \mathbb{R}^A$, $f{\downarrow_B}$ denotes the *restriction of f to B*:

$$f{\downarrow_B} : B \to \mathbb{R} : x \mapsto f(x).$$

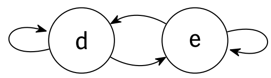


Figure 3. *$\overline{\mathcal{G}}(\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}})$ for Running example 7.*

For any family $(\mathcal{P}^x)_{x\in\mathcal{X}}$ of sets of pmfs, class $C \subseteq \mathcal{X}$ and $x \in \mathcal{X}$, we let

$$\mathcal{P}^x_C \coloneqq \{p\!\downarrow_C \colon\ p \in \mathcal{P}^x, (\forall y \in \mathcal{X} \setminus C)\ p(y) = 0\} \tag{2}$$

be the set that collects the restriction to the class C of those pmfs $p \in \mathcal{P}^x$ whose support $\{y \in \mathcal{X} \colon\ p(y) > 0\}$ is contained in C. Consider now an upper transition operator $\overline{T}$ with corresponding family $(\mathcal{Q}^x)_{x\in\mathcal{X}}$ of credal sets. For any class $C \subseteq \mathcal{X}$ such that $\mathcal{Q}^x_C \neq \emptyset$ for all $x \in C$, we let $\overline{T}_C$ be the upper transition operator corresponding to the family $(\mathcal{Q}^x_C)_{x\in C}$:

$$\overline{T}_C g(x) = \max\{\mathrm{E}_p(g) \colon\ p \in \mathcal{Q}^x_C\} \text{ for all } g \in \mathbb{R}^C, x \in C.$$

It follows almost immediately from the definition of $\overline{T}_C$ and $(\mathcal{Q}^x_C)_{x\in\mathcal{X}}$ that

$$(\overline{T}_C)^n(f\!\downarrow_C) \leq (\overline{T}^n f)\!\downarrow_C \quad \text{for all } n \in \mathbb{N}, f \in \mathbb{R}^{\mathcal{X}}. \tag{3}$$

Proof. The base case $n = 1$ follows immediately from the definition of $\overline{T}_C$ and $\mathcal{P}^x_C$. The inductive step follows from this base case and T5. □

One thing that is particularly useful in practice, is that one can obtain $\overline{T}_C$ with *any* compatible family of sets of pmfs, as long as these sets are closed.

Lemma 4.2. *Consider an upper transition operator $\overline{T}$, let $(\mathcal{P}^x)_{x\in\mathcal{X}}$ be any compatible family of* closed *sets of pmfs and fix some class $C \subseteq \mathcal{X}$. Then $\overline{T}_C$ is well defined if and only if for all $x \in C$, $\overline{T}\mathbb{1}_C(x) = 1$, or equivalently, $\mathcal{P}^x_C \neq \emptyset$; whenever this is the case, $(\mathcal{P}^x_C)_{x\in C}$ is compatible with $\overline{T}_C$.*

Running example 7. Recall that $\mathcal{X}_{\mathrm{m}} = \mathcal{X}_{\mathrm{m},1} \cup \mathcal{X}_{\mathrm{m},2} = \{\mathsf{a}, \mathsf{b}\}$, $\mathcal{X}_{\mathrm{t}\blacktriangledown} = \{\mathsf{c}\}$ and $\mathcal{X}_{\mathrm{t}\triangledown} = \{\mathsf{d}, \mathsf{e}\}$. Applying Equation (2) for $x \in C = \mathcal{X}_{\mathrm{t}\triangledown}$ we find that

$$\mathcal{P}^{\mathsf{d}}_{\mathcal{X}_{\mathrm{t}\triangledown}} = \mathcal{P}^{\mathsf{e}}_{\mathcal{X}_{\mathrm{t}\triangledown}} = \{\mathbb{1}_{\mathsf{d}}, \mathbb{1}_{\mathsf{e}}\}.$$

Lemma 4.2 therefore implies that $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}}$ is well defined—which is no coincidence, as we'll see in Lemma 4.6—and, in particular, that $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}} f = \max f$ for all $f \in \mathbb{R}^{\mathcal{X}_{\mathrm{t}\triangledown}}$; its upper accessibility graph is depicted in Figure 3. ⋄

With these notions of restriction in place, we can now formalize what we mean by zooming in on C.

Definition 4.1. An upper transition operator $\overline{T}$ is *convergent (ergodic) on $C \subseteq \mathcal{X}$* if, for all $f \in \mathbb{R}^{\mathcal{X}}$, $((\overline{T}^n f)\!\downarrow_C)_{n\in\mathbb{N}}$ converges (to a constant).

Whenever C is a maximal communication class, these restricted notions of convergence and/or ergodicity can be conveniently characterised in terms of $\overline{T}_C$.

Lemma 4.3. *Let C be one of the maximal communication classes of an upper transition operator $\overline{T}$. Then $\overline{T}_C$ is well defined and*

$$(\overline{T}_C)^n(f\!\downarrow_C) = (\overline{T}^n f)\!\downarrow_C \quad \textit{for all } n \in \mathbb{N}, f \in \mathbb{R}^{\mathcal{X}}.$$

Consequently,

(i) *C is the sole communication class for $\overline{T}_C$;*

(ii) *C is regular for $\overline{T}_C$ if and only if it's regular for $\overline{T}$;*

(iii) *$\overline{T}_C$ is ergodic if and only if $\overline{T}$ is ergodic on C;*

(iv) *$\overline{T}_C$ is convergent if and only if $\overline{T}$ is convergent on C.*

Combined with Proposition 4.1, Lemma 4.3 yields a convenient necessary and sufficient condition for convergence on $\mathcal{X}_{\mathrm{m}}$.

Proposition 4.3. *An upper transition operator $\overline{T}$ is convergent on $\mathcal{X}_{\mathrm{m}}$ if and only if $\mathcal{X}_{\mathrm{m},1}, \dots, \mathcal{X}_{\mathrm{m},M}$ are regular.*

Proof. $\overline{T}$ is clearly convergent on $\mathcal{X}_{\mathrm{m}}$ if and only if it is convergent on all $\mathcal{X}_{\mathrm{m},1}, \dots, \mathcal{X}_{\mathrm{m},M}$. The result now follows because, for all $k \in \{1, \dots, M\}$, $\mathcal{X}_{\mathrm{m},k}$ is a closed communication class, so it follows from Lemma 4.3 (for $C = \mathcal{X}_{\mathrm{m},k}$) and Proposition 4.1 (applied to $\overline{T}_{\mathcal{X}_{\mathrm{m},k}}$) that $\overline{T}$ is convergent on $\mathcal{X}_{\mathrm{m},k}$ if and only if $\mathcal{X}_{\mathrm{m},k}$ is regular (for $\overline{T}$). □

4.3. Maximal classes are absorbing. With convergence on $\mathcal{X}_{\mathrm{m}}$ completely covered, we now move to upper transition operators for which the closed class $\mathcal{X}_{\mathrm{m}}$ is absorbing. Recall from Lemma 3.2 that this is *always* the case for linear transition operators.

More generally, we first look at convergence on an *arbitrary* closed class C that is absorbing.

Lemma 4.4. *Let C be an absorbing closed class for the upper transition operator $\overline{T}$. Then for any $f \in \mathbb{R}^{\mathcal{X}}$, if $(\overline{T}^n f)_{n\in\mathbb{N}}$ converges on C, it converges (on $\mathcal{X}$) as well.*

In our proof for this result, we'll also rely on [17, Lemma 39], which we repeat here for the sake of clarity.

Lemma 4.5. *Consider an upper transition operator $\overline{T}$ with an absorbing closed class C. Then, for all $\epsilon \in \mathbb{R}_{>0}$, there is some $n_\epsilon \in \mathbb{N}$ such that for all $n \geq n_\epsilon$ and all $f \in \mathbb{R}^{\mathcal{X}}$, $|\overline{T}^n f - \overline{T}^n(f\mathbb{1}_C)| \leq \max|f|\epsilon$.*

Proof of Lemma 4.4. It suffices to show that the limit set $\Omega_f = \{\omega_1, \dots, \omega_{p_f}\}$ has period $p_f = 1$.

For all $k \in \{1, \dots, p_f\}$, since $\omega_k = \lim_{n\to+\infty} \overline{T}^{p_f n}\omega_k$ by definition, it follows from Lemma 4.5 that $\omega_k = \lim_{n\to+\infty} \overline{T}^{p_f n}(\omega_k \mathbb{1}_C)$. Since $\omega_1\!\downarrow_C = \dots = \omega_{p_f}\!\downarrow_C$ by assumption, and therefore also $\omega_1\mathbb{1}_C = \dots = \omega_{p_f}\mathbb{1}_C$, this implies that $\omega_1 = \dots = \omega_{p_f}$, proving that $p_f = 1$. □

One immediate and interesting consequence of Lemma 4.4 is that for an absorbing closed class C, convergence on C is equivalent to convergence (on $\mathcal{X}$).

Corollary 4.1. *Let C be a closed class that is absorbing. Then $\overline{T}$ is convergent on C if and only if it is convergent.*

Of course, the prime example of such an absorbing closed class is the union of all maximal communication classes $\mathcal{X}_{\mathrm{m}}$ whenever $\mathcal{X}_{\mathrm{t}\triangledown} = \emptyset$. Combined with Proposition 4.3, this yields a necessary and sufficient condition for convergence for the case $\mathcal{X}_{\mathrm{t}\triangledown} = \emptyset$.

Corollary 4.2. *Consider any upper transition operator $\overline{T}$ such that $\mathcal{X}_{\mathrm{t}\triangledown} = \emptyset$. Then $\overline{T}$ is convergent if and only if $\mathcal{X}_{\mathrm{m},1}, \dots, \mathcal{X}_{\mathrm{m},M}$ are regular.*

For linear transition operators, for which we know from Lemma 3.2 that $\mathcal{X}_{\mathrm{t}\triangledown} = \emptyset$, this result specialises to the following necessary and sufficient condition for convergence.

Corollary 4.3. *A linear transition operator T is convergent if and only if $\mathcal{X}_{\mathrm{m},1}, \dots, \mathcal{X}_{\mathrm{m},M}$ are regular.*

4.4. Maximal classes are not absorbing. We have one more step to take: to allow for transient states from which $\mathcal{X}_{\mathrm{m}}$ is not lower reachable, meaning that $\mathcal{X}_{\mathrm{t}\triangledown} \neq \emptyset$; as we've indicated before, this can only happen for non-linear transition operators. As a first step, we establish that whenever $\mathcal{X}_{\mathrm{t}\triangledown} \neq \emptyset$, we can always restrict the upper transition operator $\overline{T}$ to $\mathcal{X}_{\mathrm{t}\triangledown}$.

Lemma 4.6. *For any upper transition operator $\overline{T}$ with $\mathcal{X}_{\mathrm{t}\triangledown} \neq \emptyset$, $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}}$ is well defined.*

Proof. Let $(\mathcal{Q}^x)_{x\in\mathcal{X}}$ be the family of credal sets corresponding to $\overline{T}$. We need to show that $\mathcal{Q}^x_{\mathcal{X}_{\mathrm{t}\triangledown}} \neq \emptyset$ for all $x \in \mathcal{X}_{\mathrm{t}\triangledown}$. So fix any $x \in \mathcal{X}_{\mathrm{t}\triangledown}$. Let $(C_n)_{n\in\mathbb{Z}_{\geq 0}}$ be as defined in Lemma 3.1 with $C_0 = \mathcal{X}_{\mathrm{m}}$, and recall from there that $\mathcal{X}_{\mathrm{t}\triangledown} = \mathcal{X} \setminus C_n$ for all $n \geq k$. Recall furthermore that

$$C_k = C_{k+1} = C_k \cup \{y \in \mathcal{X} \setminus C_k : \underline{T}\mathbb{1}_{C_k}(y) > 0\},$$

which, due to T4, implies that $\underline{T}\mathbb{1}_{C_k}(x) = 0$ because $x \in \mathcal{X}_{\mathrm{t}\triangledown} = \mathcal{X} \setminus C_k$. Consider any $p \in \mathcal{Q}^x$ such that $\mathrm{E}_p(\mathbb{1}_{C_k}) = \min_{q\in\mathcal{Q}^x} \mathrm{E}_q(\mathbb{1}_{C_k}) = \underline{T}\mathbb{1}_{C_k}(x) = 0$. Then clearly $p(y) = 0$ for all $y \in C_k = \mathcal{X} \setminus \mathcal{X}_{\mathrm{t}\triangledown}$, so $p{\downarrow}_{\mathcal{X}_{\mathrm{t}\triangledown}} \in \mathcal{Q}^x_{\mathcal{X}_{\mathrm{t}\triangledown}}$ by the definition in Equation (2), implying that $\mathcal{Q}^x_{\mathcal{X}_{\mathrm{t}\triangledown}} \neq \emptyset$. □

Our next step is to use $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}}$ to decompose $\mathcal{X}_{\mathrm{t}\triangledown}$ into its own maximal and transient states, similarly to how we used $\overline{T}$ to decompose $\mathcal{X}$ into $\mathcal{X}_{\mathrm{m},1}, \dots, \mathcal{X}_{\mathrm{m},M}, \mathcal{X}_{\mathrm{t}\blacktriangledown}$ and $\mathcal{X}_{\mathrm{t}\triangledown}$. During this process, to avoid confusion, we will adopt $\mathcal{X}^1_{\mathrm{m},1}, \dots, \mathcal{X}^1_{\mathrm{m},M_1}, \mathcal{X}^1_{\mathrm{t}\blacktriangledown}$ and $\mathcal{X}^1_{\mathrm{t}\triangledown}$ as an alternative notation for $\mathcal{X}_{\mathrm{m},1}, \dots, \mathcal{X}_{\mathrm{m},M}, \mathcal{X}_{\mathrm{t}\blacktriangledown}$ and $\mathcal{X}_{\mathrm{t}\triangledown}$, respectively. We now repeat the subdivision from before, but for $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}}$ and $\mathcal{X}^1_{\mathrm{t}\triangledown}$ rather than $\overline{T}$ and $\mathcal{X}$: we let $\mathcal{X}^2_{\mathrm{m},1}, \dots, \mathcal{X}^2_{\mathrm{m},M_2}$ denote the maximal classes of $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}}$, let $\mathcal{X}^2_{\mathrm{m}}$ denote the union of these maximal classes, let $\mathcal{X}^2_{\mathrm{t}\blacktriangledown}$ collect all states in $\mathcal{X}^1_{\mathrm{t}\triangledown} \setminus \mathcal{X}^2_{\mathrm{m}}$ from which $\mathcal{X}^2_{\mathrm{m}}$ is lower reachable by $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}}$, and let $\mathcal{X}^2_{\mathrm{t}\triangledown}$ collect all states in $\mathcal{X}^1_{\mathrm{t}\triangledown} \setminus \mathcal{X}^2_{\mathrm{m}}$ from which $\mathcal{X}^2_{\mathrm{m}}$ isn't lower reachable by $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}}$. If $\mathcal{X}^2_{\mathrm{t}\triangledown} \neq \emptyset$, we repeat this process with $(\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}})_{\mathcal{X}^2_{\mathrm{t}\triangledown}}$—so the restriction of $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}}$ to $\mathcal{X}^2_{\mathrm{t}\triangledown}$—to similarly decompose $\mathcal{X}^2_{\mathrm{t}\triangledown}$. This notation is a bit unwieldy though. Luckily, the following lemma implies that we can use the simpler notation $\overline{T}_{\mathcal{X}^2_{\mathrm{t}\triangledown}}$ instead because $\overline{T}_{\mathcal{X}^2_{\mathrm{t}\triangledown}} = (\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}})_{\mathcal{X}^2_{\mathrm{t}\triangledown}}$.

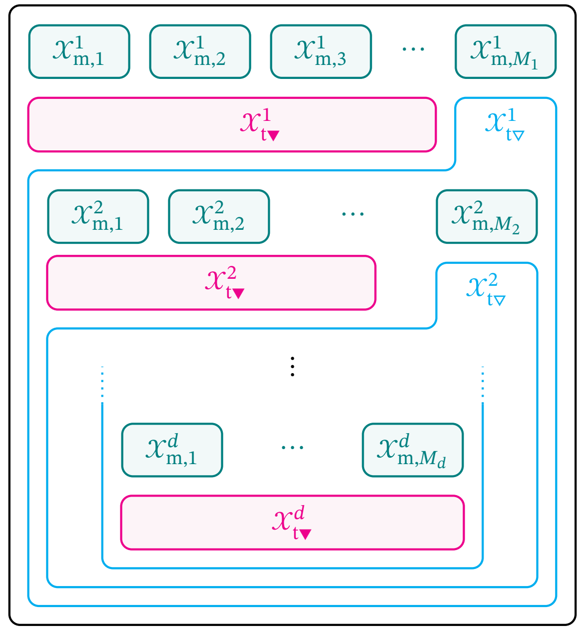


Figure 4. *Modified Venn diagram of the state space $\mathcal{X}$*

Lemma 4.7. *Consider any upper transition operator $\overline{T}$ and any two classes $C, D \subseteq \mathcal{X}$ such that $C \supseteq D$. If $\overline{T}_C$ and $(\overline{T}_C)_D$ are well defined, then $(\overline{T}_C)_D = \overline{T}_D$.*

We continue the process of repeated subdivisions until we reach a depth $d \in \mathbb{N}$ such that $\mathcal{X}^d_{\mathrm{t}\triangledown} = \emptyset$—this is always the case, as $\mathcal{X}$ is finite. Then by construction, $\mathcal{X}^1_{\mathrm{m},1}, \mathcal{X}^1_{\mathrm{m},2}, \dots, \mathcal{X}^d_{\mathrm{m},M_d}, \mathcal{X}^1_{\mathrm{t}\blacktriangledown}, \dots, \mathcal{X}^d_{\mathrm{t}\blacktriangledown}$ partitions $\mathcal{X}$, as depicted in Figure 4. Finally, we let $\mathcal{X}^*_{\mathrm{m}} := \cup_{\ell=1}^d \cup_{k=1}^{M_\ell} \mathcal{X}^\ell_{\mathrm{m},k}$ and $\mathcal{X}^*_{\mathrm{t}\blacktriangledown} := \cup_{\ell=1}^d \mathcal{X}^\ell_{\mathrm{t}\blacktriangledown}$ and, for the sake of convenience, let $\mathcal{X}^0_{\mathrm{t}\triangledown} := \mathcal{X}$.

This partitioning of the state space $\mathcal{X}$ allows us to present the following (general) sufficient condition for convergence of an upper transition operator.

Theorem 4.1. *Consider any upper transition operator $\overline{T}$. If $\mathcal{X}^\ell_{\mathrm{m},1}, \dots, \mathcal{X}^\ell_{\mathrm{m},M_\ell}$ are $\overline{T}_{\mathcal{X}^{\ell-1}_{\mathrm{t}\triangledown}}$-regular for all $\ell \in \{1, \dots, d\}$, then $\overline{T}$ is convergent.*

Before heading to its proof, let's apply this result to our running example.

Running example 8. It's clear from Figure 3 that $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}}$ has a single maximal communication class that is given by $\{\mathsf{d},\mathsf{e}\}$, so $\mathcal{X}_{\mathrm{m}}^2 = \mathcal{X}_{\mathrm{m},1}^2 = \{\mathsf{d},\mathsf{e}\}$ and $\mathcal{X}_{\mathrm{t}\blacktriangledown}^2 = \mathcal{X}_{\mathrm{t}\triangledown}^2 = \emptyset$. This finishes our decomposition of the state space. To conclude that $\overline{T}$ is convergent, let's check whether $\mathcal{X}_{\mathrm{m},1}^1$ and $\mathcal{X}_{\mathrm{m},2}^1$ are both $\overline{T}$-regular, and whether $\mathcal{X}_{\mathrm{m}}^2$ is $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}^1}$-regular. Since $\mathcal{X}_{\mathrm{m},1}^1$ and $\mathcal{X}_{\mathrm{m},2}^1$ both consist of a single element, it is immediate that they are $\overline{T}$-regular. Furthermore, since the upper accessibility graph $\overline{\mathcal{G}}(\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}^1})$ in Figure 3 has cyclicity 1, $\mathcal{X}_{\mathrm{m}}^2$ is $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}^1}$-regular. ⋄

As a first step towards proving Theorem 4.1, we present a result that enables us to use the $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}^{\ell-1}}$-regularity of $\mathcal{X}_{\mathrm{m},k}^{\ell}$ to establish the convergence (and even ergodicity) of $\overline{T}$ on $\mathcal{X}_{\mathrm{m},k}^{\ell}$.

Lemma 4.8. *Consider an upper transition operator $\overline{T}$ and a non-empty class C such that $\overline{T}_C$ is defined. If a maximal class of $\overline{T}_C$ is $\overline{T}_C$-regular, then $\overline{T}$ is ergodic on this class.*

Proof. Consider a maximal class $D \subseteq C$ of $\overline{T}_C$ that is $\overline{T}_C$-regular, fix any $f \in \mathbb{R}^{\mathcal{X}}$ and consider its limit set $\Omega_f = \{\omega_1, \dots, \omega_{p_f}\}$. It suffices to show that there is a constant $\phi \in \mathbb{R}^D$ such that $\phi = \omega_1{\downarrow_D} = \cdots = \omega_{p_f}{\downarrow_D}$.

Fix any $k \in \{1, \dots, p_f\}$. It's immediate from Proposition 4.2 and Lemmas 4.3 and 4.7 that $((\overline{T}_D)^n(\omega_k{\downarrow_D}))_{n\in\mathbb{N}}$ converges to some constant function, say $\phi_k \in \mathbb{R}^D$, such that $\phi_k \geq \min \omega_k{\downarrow_D}$ and, unless $\omega_k{\downarrow_D}$ is a constant, $\phi_k > \min \omega_k{\downarrow_D}$. Furthermore, due to Equation (3),

$$\phi_k = \lim_{n\to+\infty} (\overline{T}_D)^{p_f n}\omega_k{\downarrow_D} \leq \lim_{n\to+\infty} (\overline{T}^{p_f n}\omega_k){\downarrow_D} = \omega_k{\downarrow_D}.$$

It must therefore be that $\omega_k{\downarrow_D}$ is a constant, because otherwise $\min \omega_k{\downarrow_D} < \phi_k \leq \omega_k{\downarrow_D}$. Since $\min \omega_k{\downarrow_D} \leq \phi_k \leq \omega_k{\downarrow_D}$ and $\omega_k{\downarrow_D}$ and ϕ_k are constants, it follows that $\phi_k = \omega_k{\downarrow_D}$ is a constant function.

It therefore suffices to show that $\phi_1 = \cdots = \phi_{p_f}$. To this end, observe that for every $k \in \{1, \dots, p_f\}$, it follows from Equation (3) that, with $\phi_{p_f+1} \coloneqq \phi_1$,

$$\begin{aligned}\phi_k &= \lim_{n\to+\infty} (\overline{T}_D)^{n+1}\omega_k{\downarrow_D} \\ &= \lim_{n\to+\infty} (\overline{T}_D)^{n}\overline{T}_D\omega_k{\downarrow_D} \\ &\leq \limsup_{n\to+\infty} (\overline{T}_D)^{n}(\overline{T}\omega_k){\downarrow_D} \\ &= \lim_{n\to+\infty} (\overline{T}_D)^{n}\omega_{k+1}{\downarrow_D} = \phi_{k+1},\end{aligned}$$

whence $\phi_1 \leq \phi_2 \leq \cdots \leq \phi_{p_f} \leq \phi_1$, as required. □

If $\mathcal{X}_{\mathrm{m}}^*$ were an absorbing closed class for $\overline{T}$, we could combine this result with Lemma 4.4 to prove Theorem 4.1; it is however not. We therefore derive another upper transition operator $\overline{T}^*$ from $\overline{T}$ as follows: for all $f \in \mathbb{R}^{\mathcal{X}}$, let $\overline{T}^*f(x) \coloneqq \overline{T}f(x)$ for all $x \in \mathcal{X}_{\mathrm{t}\blacktriangledown}^*$ and let $\overline{T}^*f(x) \coloneqq \overline{T}_{\mathcal{X}_{\mathrm{m},k}^{\ell}}(f{\downarrow_{\mathcal{X}_{\mathrm{m},k}^{\ell}}})(x)$ for all $x \in \mathcal{X}_{\mathrm{m},k}^{\ell}$, with $\ell \in \{1, \dots, d\}$ and $k \in \{1, \dots, M_\ell\}$; it's immediate from Lemmas 4.3, 4.6 and 4.7 that $\overline{T}^*$ is well defined and it's easy to verify that it's an upper transition operator. By Equation (3), this definition ensures that

$$\overline{T}^*f \leq \overline{T}f \quad \text{for all } f \in \mathbb{R}^{\mathcal{X}}, \tag{4}$$

with equality on $\mathcal{X}_{\mathrm{t}\blacktriangledown}^*$. That $\overline{T}^*$ does satisfy the conditions in Lemma 4.4 is our next result.

Lemma 4.9. *Consider an upper transition operator $\overline{T}$. Then $\mathcal{X}_{\mathrm{m}}^*$ is an absorbing closed class for $\overline{T}^*$.*

Proof. For all $\ell \in \{1, \dots, d\}$ and $k \in \{1, \dots, M_\ell\}$, $\mathcal{X}_{\mathrm{m},k}^{\ell}$ is a closed class for $\overline{T}^*$ because, for all $x \in \mathcal{X}_{\mathrm{m},k}^{\ell}$ and $y \in \mathcal{X} \setminus \mathcal{X}_{\mathrm{m},k}^{\ell}$, $\overline{T}^*\mathbb{1}_y(x) = \overline{T}_{\mathcal{X}_{\mathrm{m},k}^{\ell}}(0)(x) \overset{\text{T4}}{=} 0$. Since $\mathcal{X}_{\mathrm{m}}^*$ is a union of such closed classes, it is itself closed as well.

To show that $\mathcal{X}_{\mathrm{m}}^*$ is absorbing for $\overline{T}^*$, we'll prove by induction that, for all $\ell \in \{1, \dots, d\}$, $\mathcal{X}_{\mathrm{m}}^*$ is lower reachable by $\overline{T}^*$ from all states in $\mathcal{X}_{\mathrm{t}\blacktriangledown}^{1:\ell} \coloneqq \cup_{i=1}^{\ell}\mathcal{X}_{\mathrm{t}\blacktriangledown}^{i}$. For $\ell = d$, we then find that $\mathcal{X}_{\mathrm{m}}^*$ is lower reachable by $\overline{T}^*$ from all states in $\cup_{i=1}^{d}\mathcal{X}_{\mathrm{t}\blacktriangledown}^{i} = \mathcal{X}_{\mathrm{t}\blacktriangledown}^* = \mathcal{X} \setminus \mathcal{X}_{\mathrm{m}}^*$, or equivalently, that $\mathcal{X}_{\mathrm{m}}^*$ is indeed absorbing for $\overline{T}^*$.

For the base case $\ell = 1$, we need to show that $\mathcal{X}_{\mathrm{m}}^*$ is lower reachable by $\overline{T}^*$ from all states in $\mathcal{X}_{\mathrm{t}\blacktriangledown}^1$. Consider any $x \in \mathcal{X}_{\mathrm{t}\blacktriangledown}^1 = \mathcal{X}_{\mathrm{t}\blacktriangledown}$. Since $\mathcal{X}_{\mathrm{m}}$ is lower reachable by $\overline{T}$ from all states in $\mathcal{X}_{\mathrm{t}\blacktriangledown}$, there is some $n \in \mathbb{N}$ such that $\underline{T}^n\mathbb{1}_{\mathcal{X}_{\mathrm{m}}}(x) > 0$. Since $\mathcal{X}_{\mathrm{m}} \subseteq \mathcal{X}_{\mathrm{m}}^*$ and $\underline{T}^* \geq \underline{T}$ [Equation (4) and conjugacy], it follows that

$$(\underline{T}^*)^n\mathbb{1}_{\mathcal{X}_{\mathrm{m}}^*}(x) \geq \underline{T}^n\mathbb{1}_{\mathcal{X}_{\mathrm{m}}^*}(x) \overset{\text{T5}}{\geq} \underline{T}^n\mathbb{1}_{\mathcal{X}_{\mathrm{m}}}(x) > 0,$$

whence $\mathcal{X}_{\mathrm{m}}^*$ is indeed lower reachable by $\overline{T}^*$ from x.

For the induction step, we assume that $\mathcal{X}_{\mathrm{m}}^*$ is lower reachable by $\overline{T}^*$ from all states in $\mathcal{X}_{\mathrm{t}\blacktriangledown}^{1:\ell}$ for some $1 \leq \ell < d$, and set out to prove that the same is then true for $\mathcal{X}_{\mathrm{t}\blacktriangledown}^{1:\ell+1} = \mathcal{X}_{\mathrm{t}\blacktriangledown}^{1:\ell} \cup \mathcal{X}_{\mathrm{t}\blacktriangledown}^{\ell+1}$.

By definition, $\mathcal{X}_{\mathrm{t}\blacktriangledown}^{\ell+1}$ contains the states in $\mathcal{X}_{\mathrm{t}\triangledown}^{\ell} \setminus \mathcal{X}_{\mathrm{m}}^{\ell+1}$ from which $\mathcal{X}_{\mathrm{m}}^{\ell+1}$ is lower reachable by $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}^{\ell}}$. Hence, if we let $C_0 \coloneqq \mathcal{X}_{\mathrm{m}}^{\ell+1}$ and, for all $n \in \mathbb{Z}_{\geq 0}$,

$$C_{n+1} \coloneqq C_n \cup \{x \in \mathcal{X}_{\mathrm{t}\triangledown}^{\ell} \setminus C_n : \underline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}^{\ell}}\mathbb{1}_{C_n}(x) > 0\},$$

then since $\mathcal{X}_{\mathrm{m}}^{\ell+1}$ is closed for $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}^{\ell}}$, Lemma 3.1 says that $\mathcal{X}_{\mathrm{t}\blacktriangledown}^{\ell+1} = C_k \setminus \mathcal{X}_{\mathrm{m}}^{\ell+1}$ for some $k \in \mathbb{Z}_{\geq 0}$.

We'll prove by induction over n that $C_n \setminus \mathcal{X}_{\mathrm{m}}^{\ell+1}$ is a set of states from which $\mathcal{X}_{\mathrm{m}}^*$ is lower reachable by $\overline{T}^*$. For $n = k$, we then find that $\mathcal{X}_{\mathrm{m}}^*$ is lower reachable by $\overline{T}^*$ from all states in $\mathcal{X}_{\mathrm{t}\blacktriangledown}^{\ell+1} = C_k \setminus \mathcal{X}_{\mathrm{m}}^{\ell+1}$.

The base case $n = 0$ is trivially true because $C_0 \setminus \mathcal{X}_{\mathrm{m}}^{\ell+1} = \emptyset$. For the induction step, we assume that, for some $n \in \{0, \dots, k-1\}$, $\mathcal{X}_{\mathrm{m}}^*$ is lower reachable from $C_n \setminus \mathcal{X}_{\mathrm{m}}^{\ell+1}$ by $\overline{T}^*$, and set out to prove that $\mathcal{X}_{\mathrm{m}}^*$ is then lower reachable from $C_{n+1} \setminus C_n$ by $\overline{T}^*$ as well.

To that end, consider any $x \in C_{n+1} \setminus C_n$. Since $\mathcal{X}_{\mathrm{m}}^*$ is a closed class for $\overline{T}^*$, we know from Lemma 3.1 that there is some $K \in \mathbb{N}$ such that $(\underline{T}^*)^K \mathbb{1}_{\mathcal{X}^*_{\mathrm{m}}}(y) > 0$ for all $y \in \mathcal{X}^*_{\mathrm{m}}$ and all $y \in \mathcal{X} \setminus \mathcal{X}^*_{\mathrm{m}}$ from which $\mathcal{X}^*_{\mathrm{m}}$ is lower reachable by $\overline{T}^*$. We now set out to prove that $(\underline{T}^*)^{K+1} \mathbb{1}_{\mathcal{X}^*_{\mathrm{m}}}(x) > 0$, thereby indeed establishing that $\mathcal{X}^*_{\mathrm{m}}$ is lower reachable from x by $\overline{T}^*$, as required.

Since it follows from our induction hypothesises that $\mathcal{X}^{1:\ell}_{\mathrm{t}\blacktriangledown}$ and $C_n \setminus \mathcal{X}^{\ell+1}_{\mathrm{m}}$ are sets from which $\mathcal{X}^*_{\mathrm{m}}$ is lower reachable by $\overline{T}^*$, we know that $(\underline{T}^*)^K \mathbb{1}_{\mathcal{X}^*_{\mathrm{m}}}(y) > 0$ for all $y \in (\mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}) \cup C_n$. Since $(\underline{T}^*)^K \mathbb{1}_{\mathcal{X}^*_{\mathrm{m}}} \geq 0$ by T4, this implies that there is some $\alpha > 0$ such that $(\underline{T}^*)^K \mathbb{1}_{\mathcal{X}^*_{\mathrm{m}}} \geq \alpha \mathbb{1}_{(\mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}) \cup C_n}$. Hence,

$$\begin{aligned}(\underline{T}^*)^{K+1} \mathbb{1}_{\mathcal{X}^*_{\mathrm{m}}}(x) &= \underline{T}^*\big((\underline{T}^*)^K \mathbb{1}_{\mathcal{X}^*_{\mathrm{m}}}\big)(x) \\ &\overset{\text{T5}}{\geq} \underline{T}^*\big(\alpha \mathbb{1}_{(\mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}) \cup C_n}\big)(x) \\ &\overset{\text{T2}}{=} \alpha \underline{T}^*\big(\mathbb{1}_{(\mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}) \cup C_n}\big)(x) \\ &= \alpha \underline{T}\big(\mathbb{1}_{(\mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}) \cup C_n}\big)(x), \qquad (5)\end{aligned}$$

where the last equality follows from the definition of $\overline{T}^*$ because $x \in C_{n+1} \setminus C_n \subseteq C_k \setminus \mathcal{X}^{\ell+1}_{\mathrm{m}} = \mathcal{X}^{\ell+1}_{\mathrm{t}\blacktriangledown} \subseteq \mathcal{X}^*_{\mathrm{t}\blacktriangledown}$.

Let $(\mathcal{Q}^y)_{y \in \mathcal{X}}$ be the family of credal sets that correspond with $\overline{T}$. Since $\mathcal{Q}^x$ is closed by definition, there is some $p \in \mathcal{Q}^x$ such that

$$\begin{aligned}\underline{T}\big(\mathbb{1}_{(\mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}) \cup C_n}\big)(x) &= \mathrm{E}_p\big(\mathbb{1}_{(\mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}) \cup C_n}\big) \\ &= \mathrm{E}_p\big(\mathbb{1}_{\mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}}\big) + \mathrm{E}_p(\mathbb{1}_{C_n}),\end{aligned}$$

where for the second equality we used that $(\mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}) \cap C_n = \emptyset$, which holds because $C_n \subseteq \mathcal{X}^\ell_{\mathrm{t}\triangledown}$ by definition. Since both terms in the sum are clearly non-negative, it suffices to establish that at least one of them is positive to show that $\underline{T}^*\big(\mathbb{1}_{(\mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}) \cup C_n}\big)(x) > 0$, which then implies $(\underline{T}^*)^{K+1} \mathbb{1}_{\mathcal{X}^*_{\mathrm{m}}}(x) > 0$ due to Equation (5). To that end, we'll assume that $p(y) = 0$ for all $y \in \mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}$, or equivalently, that $\mathrm{E}_p(\mathbb{1}_{\mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}}) = 0$ and show that this implies that $\mathrm{E}_p(\mathbb{1}_{C_n}) > 0$.

Since $p(y) = 0$ for all $y \in \mathcal{X} \setminus \mathcal{X}^\ell_{\mathrm{t}\triangledown}$, Equation (2) tells us that $p{\downarrow}_{\mathcal{X}^\ell_{\mathrm{t}\triangledown}} \in \mathcal{P}^x_{\mathcal{X}^\ell_{\mathrm{t}\triangledown}}$. This implies that

$$\mathrm{E}_p(\mathbb{1}_{C_n}) = \mathrm{E}_{p{\downarrow}_{\mathcal{X}^\ell_{\mathrm{t}\triangledown}}}(\mathbb{1}_{C_n}) \geq \underline{T}_{\mathcal{X}^\ell_{\mathrm{t}\triangledown}} \mathbb{1}_{C_n}(x) > 0,$$

where for the final inequality we use the fact that $x \in C_{n+1} \setminus C_n$. □

Finally, then, we can lay out our proof for Theorem 4.1.

Proof of Theorem 4.1. Fix $f \in \mathbb{R}^{\mathcal{X}}$ and consider its limit set $\Omega_f = \{\omega_1, \dots, \omega_{p_f}\}$. It suffices to show that $p_f = 1$.

For any $\ell \in \{1, \dots, d\}$ and $k \in \{1, \dots, M_\ell\}$, since $\mathcal{X}^\ell_{\mathrm{m},k}$ is a maximal class of $\overline{T}_{\mathcal{X}^{\ell-1}_{\mathrm{t}\triangledown}}$ that is $\overline{T}_{\mathcal{X}^{\ell-1}_{\mathrm{t}\triangledown}}$-regular, it's immediate from Lemma 4.8 that $\overline{T}$ is ergodic on $\mathcal{X}^\ell_{\mathrm{m},k}$, which implies that $(\overline{T}^n \omega_1)_{n \in \mathbb{N}}$ converges to a constant on $\mathcal{X}^\ell_{\mathrm{m},k}$, and hence, $\omega_1{\downarrow}_{\mathcal{X}^\ell_{\mathrm{m},k}} = \dots = \omega_{p_f}{\downarrow}_{\mathcal{X}^\ell_{\mathrm{m},k}}$ is constant. Consequently, for all $i \in \{1, \dots, p_f\}$,

$$\omega_{i+1}(x) = \omega_i(x) \overset{\text{T4}}{=} \overline{T}_{\mathcal{X}^\ell_{\mathrm{m},k}}(\omega_i{\downarrow}_{\mathcal{X}^\ell_{\mathrm{m},k}})(x) = \overline{T}^* \omega_i(x)$$

for all $x \in \mathcal{X}^\ell_{\mathrm{m},k}$. Since also $\omega_{i+1}(x) = \overline{T}\omega_i(x) = \overline{T}^*\omega_i(x)$ for all $i \in \{1, \dots, p_f\}$ and $x \in \mathcal{X}^*_{\mathrm{t}\triangledown}$, we conclude that $\omega_{i+1} = \overline{T}^*\omega_i$ for all $i \in \{1, \dots, p_f\}$. Consequently, the limit set of $((\overline{T}^*)^n \omega_1)_{n \in \mathbb{N}}$ is $\{\omega_1, \dots, \omega_{p_f}\}$.

Since $\omega_1{\downarrow}_{\mathcal{X}^\ell_{\mathrm{m},k}} = \dots = \omega_{p_f}{\downarrow}_{\mathcal{X}^\ell_{\mathrm{m},k}}$ for all $\ell \in \{1, \dots, d\}$ and $k \in \{1, \dots, M_\ell\}$, this implies that $((\overline{T}^*)^n \omega_1)_{n \in \mathbb{N}}$ converges on $\mathcal{X}^*_{\mathrm{m}}$. Now recall from Lemma 4.9 that for $\overline{T}^*$, $\mathcal{X}^*_{\mathrm{m}}$ is an absorbing closed class. Consequently, it follows from Lemma 4.4 that $((\overline{T}^*)^n \omega_1)_{n \in \mathbb{N}}$ converges, so $\omega_1 = \dots = \omega_{p_f}$ and therefore $p_f = 1$. □

5. The Finitely Generated Case

A natural follow-up question is whether the sufficient condition in Theorem 4.1 is also a necessary one. This is the case, at least if the upper transition operator $\overline{T}$ is *finitely generated*, meaning that it is compatible with a family $(\mathcal{P}^x)_{x \in \mathcal{X}}$ of *finite* sets of pmfs.[3]

Proposition 5.1. *Consider a finitely generated upper transition operator $\overline{T}$. If $\overline{T}$ is convergent, then for all $\ell \in \{1, \dots, d\}$, $\mathcal{X}^\ell_{\mathrm{m},1}, \dots, \mathcal{X}^\ell_{\mathrm{m},M_\ell}$ are $\overline{T}_{\mathcal{X}^{\ell-1}_{\mathrm{t}\triangledown}}$-regular.*

Our proof follows relatively straightforward from some intermediary results. Since it is rather instructive, we run through it in the main text. First, recall from Proposition 4.3 that as $\overline{T}$ is convergent, the maximal communication classes $\mathcal{X}^1_{\mathrm{m},k}$ are regular for $\overline{T}_{\mathcal{X}^0_{\mathrm{t}\triangledown}} = \overline{T}$. So if $d = 1$, we're already done. If on the other hand $d > 1$—and therefore $\mathcal{X}_{\mathrm{t}\triangledown} \neq \emptyset$—we turn our attention to the behaviour on $\mathcal{X}_{\mathrm{t}\triangledown}$.

Proposition 5.2. *Consider a finitely generated upper transition operator $\overline{T}$. If $\overline{T}$ is convergent and $\mathcal{X}_{\mathrm{t}\triangledown} \neq \emptyset$, then $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}}$ is convergent.*

Since $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}} = \overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}}$ is convergent, we may again use Proposition 4.3 to infer that $\mathcal{X}^2_{\mathrm{m},1}, \dots, \mathcal{X}^2_{\mathrm{m},M_2}$ are regular for $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}}$. Now if $d = 2$, we're done. If on the other hand $d > 2$, we want to repeat the argument, and for this we need that $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}}$ is finitely generated. This is however clearly the case. Indeed, as $\overline{T}$ is finitely generated, it is compatible with a family $(\mathcal{P}^x)_{x \in \mathcal{X}}$ of sets of pmfs that are all finite (and therefore closed). Since restrictions of finite credal sets are finite themselves, Lemma 4.2 therefore implies that $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}}$ is finitely generated as well.

[3] In fact, the reader may want to verify that it suffices for $\overline{T}$ to be *sufficiently finitely generated*, meaning that it is compatible with a family $(\mathcal{P}^x)_{x \in \mathcal{X}}$ of sets of pmfs such that for all $x \in \mathcal{X}_{\mathrm{t}\triangledown}$, $\mathcal{P}^x$ is a finite set.

Since $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}}$ is also convergent, Proposition 5.2 tells us that $\overline{T}_{\mathcal{X}^2_{\mathrm{t}\triangledown}}$ is convergent, and then Proposition 4.3 establishes that $\mathcal{X}^3_{\mathrm{m},1}, \dots, \mathcal{X}^3_{\mathrm{m},M_3}$ are $\overline{T}_{\mathcal{X}^2_{\mathrm{t}\triangledown}}$-regular.

Repeated application of the same argument until depth d—with $\mathcal{X}^d_{\mathrm{t}\triangledown} = \emptyset$—eventually results in (a proof for) Proposition 5.1.

To conclude our treatment for finitely generated upper transition operators, we combine Theorem 4.1 and Propositions 5.1 and 5.2 into the following strong result.

Theorem 5.1. *For a finitely generated upper transition operator $\overline{T}$, the following are equivalent:*

(i) *$\overline{T}$ is convergent;*

(ii) *the maximal communication classes $\mathcal{X}_{\mathrm{m},1}, \dots, \mathcal{X}_{\mathrm{m},M}$ are regular and $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}}$ is convergent;*

(iii) *for all $\ell \in \{1, \dots, d\}$, the maximal communication classes $\mathcal{X}^\ell_{\mathrm{m},1}, \dots, \mathcal{X}^\ell_{\mathrm{m},M_\ell}$ are $\overline{T}_{\mathcal{X}^{\ell-1}_{\mathrm{t}\triangledown}}$-regular.*

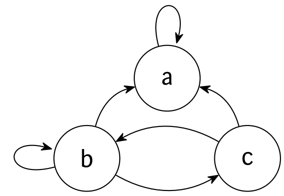
Theorem 5.1 immediately leads to an algorithm to determine whether a finitely generated upper transition operator is convergent. It consists in recursively taking the following steps, starting from $\ell = 1$: (i) construct the upper accessibility graph $\overline{\mathcal{G}}(\overline{T}_{\mathcal{X}^{\ell-1}_{\mathrm{t}\triangledown}})$, (ii) determine the maximal communication classes $\mathcal{X}^\ell_{\mathrm{m},1}, \dots, \mathcal{X}^\ell_{\mathrm{m},M_\ell}$ and their cyclicity using one of the standard algorithms; (iii) determine $\mathcal{X}^\ell_{\mathrm{t}\triangledown}$ with the recursive procedure in Lemma 3.1; and (iv) if $\mathcal{X}^\ell_{\mathrm{t}\triangledown} \neq \emptyset$ determine $\overline{T}_{\mathcal{X}^\ell_{\mathrm{t}\triangledown}}$ [via finite sets of pmfs thanks to Lemmas 4.2 and 4.7], increment ℓ and repeat, otherwise stop.

6. Wrapping things up

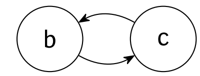
The results above give rise to another follow-up question: can we generalise Proposition 5.2 (and then also Proposition 5.1 and Theorem 5.1) to the general case of upper transition operators that need not be finitely generated? As is clear from Proposition 4.3, it's always necessary that the maximal communication classes are regular. Unfortunately, though, for an upper transition operator that is *not* finitely generated, it's no longer necessary for convergence that $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}}$ is convergent, making it impossible to generalise Proposition 5.2 to this case. What follows is a straightforward example of a convergent upper transition operator $\overline{T}$ that is not finitely generated such that $\overline{T}_{\mathcal{X}_{\mathrm{t}\triangledown}}$ is not convergent—because its sole communication class is not regular.

6.1. Counterexample. Let $\mathcal{X} := \{\mathrm{a}, \mathrm{b}, \mathrm{c}\}$ and $p_\epsilon := \epsilon^2 \mathbb{1}_{\mathrm{a}} + \epsilon \mathbb{1}_{\mathrm{b}} + (1 - \epsilon - \epsilon^2)\mathbb{1}_{\mathrm{c}}$ for all $\epsilon \in [0, 1]$. Then the sets of transition pmfs

$$\mathcal{P}^{\mathrm{a}} := \{\mathbb{1}_{\mathrm{a}}\}, \mathcal{P}^{\mathrm{b}} := \{\mathbb{1}_{\mathrm{a}}\} \cup \{p_\epsilon \colon \epsilon \in [0, 1/2]\}, \mathcal{P}^{\mathrm{c}} := \{\mathbb{1}_{\mathrm{a}}, \mathbb{1}_{\mathrm{b}}\}$$



(a) $\overline{\mathcal{G}}(\overline{T})$



(b) $\overline{\mathcal{G}}(\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}})$

Figure 5

induce the upper transition operator $\overline{T}$ given for all $f \in \mathbb{R}^{\mathcal{X}}$ and $x \in \mathcal{X}$ by

$$\overline{T}f(x) = \begin{cases} f(\mathrm{a}) & \text{if } x = \mathrm{a}, \\ \max\{f(\mathrm{a})\} \cup \{\mathrm{E}_{p_\epsilon}(f) \colon \epsilon \in [0, 1/2]\} & \text{if } x = \mathrm{b}, \\ \max\{f(\mathrm{a}), f(\mathrm{b})\} & \text{if } x = \mathrm{c}. \end{cases}$$

From the upper accessibility graph $\overline{\mathcal{G}}(\overline{T})$ depicted in Figure 5a, it's clear that the upper transition operator $\overline{T}$ has one maximal class: $\mathcal{X}^1_{\mathrm{m}} = \mathcal{X}^1_{\mathrm{m},1} = \{\mathrm{a}\}$. Since T8 implies that $\underline{T}\mathbb{1}_{\mathrm{a}} = 1 - \overline{T}\mathbb{1}_{\{\mathrm{b},\mathrm{c}\}} = 1 - \mathbb{1}_{\{\mathrm{b},\mathrm{c}\}} = \mathbb{1}_{\mathrm{a}}$, it follows that

$$\mathcal{X}^1_{\mathrm{t}\blacktriangledown} = \{x \in \{\mathrm{b}, \mathrm{c}\} \colon (\exists n \in \mathbb{N})\, \underline{T}^n \mathbb{1}_{\mathrm{a}}(x) > 0\} = \emptyset,$$

so $\mathcal{X}^1_{\mathrm{t}\triangledown} = \{\mathrm{b}, \mathrm{c}\}$.

Since $\mathcal{P}^{\mathrm{b}}$ and $\mathcal{P}^{\mathrm{c}}$ are closed, Lemma 4.2 implies that $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}}$ is compatible with the restricted sets of pmfs

$$\mathcal{P}^{\mathrm{b}}_{\mathcal{X}^1_{\mathrm{t}\triangledown}} = \{\mathbb{1}_{\mathrm{c}}\} \text{ and } \mathcal{P}^{\mathrm{c}}_{\mathcal{X}^1_{\mathrm{t}\triangledown}} = \{\mathbb{1}_{\mathrm{b}}\};$$

it is therefore easy to see that for all $g \in \mathbb{R}^{\mathcal{X}^1_{\mathrm{t}\triangledown}}$,

$$\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}} g(\mathrm{b}) = g(\mathrm{c}) \quad \text{and} \quad \overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}} g(\mathrm{c}) = g(\mathrm{b}).$$

Its upper accessibility graph $\overline{\mathcal{G}}(\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}})$ is depicted in Figure 5b. It has only one (and therefore maximal) communication class $\mathcal{X}^2_{\mathrm{m}} = \{\mathrm{b}, \mathrm{c}\} = \mathcal{X}^1_{\mathrm{t}\triangledown}$, so $\mathcal{X}^2_{\mathrm{t}\blacktriangledown} = \emptyset = \mathcal{X}^2_{\mathrm{t}\triangledown}$. Since it's clear that $\mathcal{X}^2_{\mathrm{m}}$ has cyclicity 2, it is not regular and therefore, due to Proposition 4.1, $\overline{T}_{\mathcal{X}^1_{\mathrm{t}\triangledown}}$ is not convergent. Consequently, the sufficient condition for convergence in Theorem 4.1 is *not* met, nor is the necessary condition in Proposition 5.2.

Nonetheless, as we will now show, $\overline{T}$ is convergent. In particular, for all $f \in \mathbb{R}^{\mathcal{X}}$ and $x \in \mathcal{X}$,

$$\lim_{n \to +\infty} \overline{T}^n f(x) = \begin{cases} f(\mathrm{a}) & \text{if } x = \mathrm{a} \\ \max f & \text{if } x \neq \mathrm{a}. \end{cases}$$

Since $\overline{T}f(\mathrm{a}) = f(\mathrm{a}) = \min \overline{T}f$ and $\max \overline{T}f = \max f$, we can assume without loss of generality that $\min f = f(\mathrm{a})$. It follows immediately from the expression for $\overline{T}$ that $\overline{T}^n f(\mathrm{a}) = f(\mathrm{a})$ for all $n \in \mathbb{N}$, so we can focus on the value in b and c.

If $f(\mathsf{a}) \leq f(\mathsf{b}) = f(\mathsf{c}) = \max f$, then it follows from the expression for $\overline{T}$ that $\overline{T}^n f = f$ for all $n \in \mathbb{N}$. Hence, it remains for us to look at the limiting behaviour on b and c in case $f(\mathsf{a}) < \max f$ with $f(\mathsf{b}) \neq f(\mathsf{c})$.

For all $n \in \mathbb{N}$, let $f_n^- := \min\{\overline{T}^n f(\mathsf{b}), \overline{T}^n f(\mathsf{c})\}$ and $f_n^+ := \max\{\overline{T}^n f(\mathsf{b}), \overline{T}^n f(\mathsf{c})\}$. We continue with a general observation: for any $h \in \mathbb{R}^{\mathcal{X}}$, we write $h = (h(\mathsf{a}), h(\mathsf{b}), h(\mathsf{c}))$ and observe that

$$h(\mathsf{a}) \leq h(\mathsf{b}) < h(\mathsf{c}) \Rightarrow \overline{T}h = (h(\mathsf{a}), h(\mathsf{c}), h(\mathsf{b})); \tag{6}$$

$$h(\mathsf{a}) \leq h(\mathsf{c}) < h(\mathsf{b}) \Rightarrow \begin{cases} \overline{T}h(\mathsf{a}) = h(\mathsf{a}), \\ h(\mathsf{c}) < \overline{T}h(\mathsf{b}) \leq h(\mathsf{b}), \\ \overline{T}h(\mathsf{c}) = h(\mathsf{b}). \end{cases} \tag{7}$$

From Equations (6) and (7) and the definition of $\overline{T}$, it follows that (i) $f_n^+ = \max f$ for all $n \in \mathbb{N}$; and (ii) the sequence $(f_n^-)_{n\in\mathbb{N}}$ is non-decreasing in $[f(\mathsf{a}), \max f]$, and therefore converges to a limit λ^- with $f(\mathsf{a}) \leq \lambda^- \leq \max f$. We now need to show that $\lambda^- = \max f$, so we assume towards contradiction that $f(\mathsf{a}) \leq \lambda^- < \max f$. On the one hand, the orbit $(\overline{T}^n f)_{n\in\mathbb{N}}$ then has a limit set $\Omega_f = \{\omega_1, \omega_2\}$ of period $p_f = 2$, which alternates between $\omega_1 = h_\mathsf{c} := (f(\mathsf{a}), \max f, \lambda^-)$ and $\omega_2 = h_\mathsf{b} := (f(\mathsf{a}), \lambda^-, \max f)$. This implies in particular that $h_\mathsf{b} = \omega_2 = \overline{T}\omega_1 = \overline{T}h_\mathsf{c}$. On the other hand, it follows from Equation (7) that $\overline{T}h_\mathsf{c}(\mathsf{b}) > h_\mathsf{c}(\mathsf{c}) = h_\mathsf{b}(\mathsf{b})$; the contradiction we were after!

6.2. Outlook. In future work, we'd like to find out whether for the general, not necessarily finitely generated case, there is some—necessarily other—equivalent characterisation of convergence that can also be easily checked. Our preliminary research has already revealed that it is indeed possible to come up with an interesting equivalent condition, but we've not yet succeeded at translating this condition into one that can be easily verified in practice.

We also plan to scrutinise the relation between our Theorem 4.1 and Akian and Gaubert's [1] Theorem 5.5, which says that $\overline{T}$ is convergent if all of the strongly connected components of their 'critical graph' $\mathcal{G}^c(\overline{T})$ have cyclicity 1, where $\mathcal{G}^c(\overline{T})$ is defined in terms of the accessibility graphs of all the linear transition operators T that are dominated by $\overline{T}$. While we strongly believe our condition is more convenient to verify than theirs, there seems to be a strong connection between our condition and theirs, and we wouldn't be surprised if they turn out to be equivalent.

Additional Author Information

Acknowledgements. A lot of the ideas and some of the results—for example the counterexample in Section 6.1, Propositions 4.1 and 4.3, Corollary 4.2, and Proposition 5.2—in this contribution have their roots in Hanne Asselman's Master's thesis [2], which was supervised by the authors. These ideas were further developed by the authors and Natan T'Joens during a first research retreat and by the authors during a second research retreat. The authors thank Hanne and Natan for their valuable input, as well as for their blessing to continue this line of research without them.

Our thanks also goes out to the three anonymous reviewers for their constructive comments.

Jasper De Bock's research is supported by Research Foundation – Flanders (FWO), project number G0ACO25N. Alexander Erreygers is a post-doctoral research fellow of Ghent University's Special Research Fund; his research was also supported by Research Foundation – Flanders (FWO), project number 3G028919. Floris Persiau is currently a post-doctoral fellow of the Belgian American Educational Foundation; his research was also supported by Research Foundation – Flanders (FWO), project number 11H5523N.

Author contributions. The order in which the authors are listed is alphabetical; it is not intended to reflect the extent of their contribution. All authors have contributed equally.

References

[1] Marianne Akian and Stéphane Gaubert. "Spectral theorem for convex monotone homogeneous maps, and ergodic control". In: *Nonlinear Analysis: Theory, Methods & Applications* 52.2 (2003), pp. 637–679. DOI: 10.1016/S0362-546X(02)00170-0.

[2] Hanne Asselman. "The limit behaviour of imprecise Markov chains: a new perspective that goes beyond ergodicity". Master's thesis. Ghent University, 2024. URL: https://lib.ugent.be/catalog/rug01:003214571.

[3] Gert de Cooman, Jasper De Bock, and Stavros Lopatatzidis. "Imprecise stochastic processes in discrete time: global models, imprecise Markov chains, and ergodic theorems". In: *International Journal of Approximate Reasoning* 76 (2016), pp. 18–46. DOI: 10.1016/j.ijar.2016.04.009.

[4] Gert de Cooman, Filip Hermans, and Erik Quaeghebeur. "Imprecise Markov chains and their limit behavior". In: *Probability in the Engineering and Informational Sciences* 23.4 (2009), pp. 597–635. DOI: 10.1017/S0269964809990039.

[5] Robert Denk, Michael Kupper, and Max Nendel. "Kolmogorov-type and general extension results for nonlinear expectations". In: *Banach Journal of Mathematical Analysis* 12.3 (2018), pp. 515–540. DOI: 10.1215/17358787-2017-0024.

[6] Filip Hermans and Gert de Cooman. “Characterisation of ergodic upper transition operators”. In: *International Journal of Approximate Reasoning* 53.4 (2012), pp. 573–583. DOI: 10.1016/j.ijar.2011.12.008.

[7] Filip Hermans and Damjan Škulj. “Stochastic processes”. In: *Introduction to Imprecise Probabilities*. Ed. by Thomas Augustin, Frank P. A. Coolen, Gert de Cooman, and Matthias C. M. Troffaes. Wiley, 2014. Chap. 11.

[8] James P. Jarvis and Douglas R. Shier. “Graph-Theoretic Analysis of Finite Markov Chains”. In: *Applied Mathematical Modeling*. Ed. by K. Ted Wallenius and Douglas R. Shier. CRC Press, 1999. DOI: 10.1201/9780367802363.

[9] John G. Kemeny and J. Laurie Snell. *Finite Markov Chains*. Springer-Verlag, 1960.

[10] Igor O. Kozine and Lev V. Utkin. “Interval-valued finite Markov chains”. In: *Reliable Computing* 8.2 (2002), pp. 97–113. DOI: 10.1023/A:1014745904458.

[11] Bas Lemmens and Michael Scheutzow. “On the dynamics of sup-norm non-expansive maps”. In: *Ergodic Theory and Dynamical Systems* 25.3 (2005), pp. 861–871. DOI: 10.1017/s0143385704000665.

[12] James R. Norris. *Markov Chains*. Cambridge University Press, 1997. DOI: 10.1017/CBO9780511810633.

[13] Robert Sine. “A nonlinear Perron–Frobenius Theorem”. In: *Proceedings of the American Mathematical Society* 109.2 (1990), pp. 331–336. DOI: 10.2307/2047992. JSTOR: 2047992.

[14] Damjan Škulj. “Discrete time Markov chains with interval probabilities”. In: *International Journal of Approximate Reasoning* 50.8 (2009), pp. 1314–1329. DOI: 10.1016/j.ijar.2009.06.007.

[15] Damjan Škulj. “A classification of invariant distributions and convergence of imprecise Markov chains”. In: *Linear Algebra and its Applications* 439.9 (2013), pp. 2542–2561. DOI: 10.1016/j.laa.2013.07.001.

[16] Damjan Škulj and Robert Hable. “Coefficients of ergodicity for Markov chains with uncertain parameters”. In: *Metrika* 76.1 (2011), pp. 107–133. DOI: 10.1007/s00184-011-0378-0.

[17] Natan T’Joens and Jasper De Bock. “Average behaviour in discrete-time imprecise Markov chains: A study of weak ergodicity”. In: *International Journal of Approximate Reasoning* 132 (2021), pp. 181–205. DOI: 10.1016/j.ijar.2021.03.001.