Characterising the Convergence of Imprecise Markov Chains

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

[1]Foundations Lab for imprecise probabilities, Ghent University, Ghent, Belgium

Abstract

Motivated by their connection to the limit behaviour of imprecise Markov chains, we study the asymptotic behaviour of upper transition operators. Our focus is on their convergence: the condition that for every real function, the sequence generated by repeated application of the operator admits a well-defined limit. This notion is strictly weaker than classical ergodicity, which enforces convergence to a constant, and therefore requires a different analytic treatment. We develop a full characterisation of convergence in terms of graph-theoretic relations induced by the operator: accessibility and lower reachability. The resulting criterion is both necessary and sufficient, and applies to arbitrary upper transition operators without structural restrictions, thus strengthening earlier work that provided only sufficient conditions for the unrestricted case.

1 INTRODUCTION

Imprecise Markov chains [Kozine and Utkin, 2002, Škulj, 2009, De Cooman and Hermans, 2008, De Cooman et al., 2009, 2016, Denk et al., 2018, Hermans and Škulj, 2014] are a generalisation of classical Markov chains that allow for partially specified transition probabilities. One way to parametrize them is in terms of a so-called 'upper transition operator', a sublinear operator that generalises the transition matrix of a classical Markov chain. The study of the limit behaviour of an imprecise Markov chains then amounts to the study of the convergence properties of this upper transition operator, which is why we focus on the latter in this contribution. Our main result Theorem 12 is a practically verifiable necessary and sufficient condition for the convergence of the iterates of such an upper transition operator, thereby providing a complete characterisation of the convergence of imprecise Markov chains. Our work builds on recent work by De Bock et al. [2025] that studies the same type of convergence; their main results are necessary and sufficient conditions for specific types of upper transition operators, a generally applicable condition that is sufficient for convergence (their Theorem 4.1) and a counterexample showing that this condition is not necessary in general.

The remainder of this contribution is structured as follows. Section 2 gives a brief primer on imprecise Markov chains, which serves as motivating context, and introduces upper transition operators and their properties. Section 3 introduces the key notion of convergence for upper transition operators, and repeats the sufficient conditions given by De Bock et al. [2025, Theorem 4.1]. In Section 4 we introduce a new accessibility relation, and go on to show that this relation, combined with lower reachability, allows us to characterise the convergence of upper transition operators.

Proofs of all results, as well as some additional intermediary lemmas, are available in the Supplementary Material.

2 IMPRECISE MARKOV CHAINS

Throughout this contribution, we fix some finite set $\mathcal{X}$, which we call the *state space*. One very convenient way to model the uncertainty about the time-evolution of an infinite sequence $(X_n)_{n\in\mathbb{Z}_{\geq 0}}$ of $\mathcal{X}$-valued variables is a Markov chain [Kemeny and Snell, 1960, Norris, 1997]. Such a model is completely determined by an initial probability mass function $q_\square$ which describes the uncertainty regarding X_0 and, for every state $x \in \mathcal{X}$, a transition probability mass function q_x which describes the uncertainty about X_{n+1} given $X_n = x$. It is well-known that these local uncertainty models, combined with the Markov assumption, give rise to a global uncertainty model; in this case a probability measure P on the σ-algebra generated by the cylinder events, or equivalently, an expectation operator E_P on a suitable class of (extended) real-valued variables on the set of paths $\mathcal{X}^{\mathbb{Z}_{\geq 0}}$. An imprecise Markov chain relaxes this characterisation by allowing for partial specification of these probability mass functions.

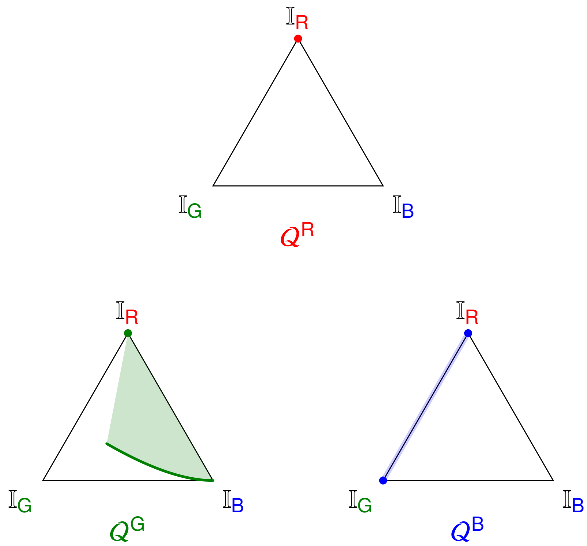


Figure 1: Ternary plots depicting the set Q^x (opaque dots and lines) and the corresponding credal set—that is, convex hull—(transparent shading) for each state $x \in \{\mathsf{R}, \mathsf{G}, \mathsf{B}\}$ in the running example.

2.1 PARTIALLY SPECIFIED MODELS

Formally, an imprecise Markov chain is characterised by a set $Q^{\square}$ of initial probability mass functions and, for every state $x \in \mathcal{X}$, a set Q^x of transition probability mass functions. The following example provides concrete instances of such sets of transition probability mass functions, which we've borrowed from De Bock et al. [2025, Section 6.1]; we will use them throughout this contribution to illustrate new concepts and results as we come across them.

Example 1. Consider the ternary state space $\mathcal{X} := \{\mathsf{R}, \mathsf{G}, \mathsf{B}\}$. For all $x \in \mathcal{X}$, let $\mathbb{I}_x : \mathcal{X} \to \{0, 1\}$ be the indicator of $\{x\}$, which only takes on the value 1 on $\{x\}$—equivalently, $\mathbb{I}_x$ is the mass function on $\mathcal{X}$ with all mass on x. For all $\epsilon \in [0, 1]$, we consider the mass function on $\mathcal{X}$ defined by $q_\epsilon := \epsilon^2 \mathbb{I}_{\mathsf{R}} + \epsilon \mathbb{I}_{\mathsf{G}} + (1 - \epsilon - \epsilon^2)\mathbb{I}_{\mathsf{B}}$. As sets of transition mass functions, we take $Q^{\mathsf{R}} := \{\mathbb{I}_{\mathsf{R}}\}$, $Q^{\mathsf{G}} := \{\mathbb{I}_{\mathsf{R}}\} \cup \{q_\epsilon : \epsilon \in [0, \frac{1}{2}]\}$ and $Q^{\mathsf{B}} := \{\mathbb{I}_{\mathsf{R}}, \mathbb{I}_{\mathsf{G}}\}$, which are depicted in Figure 1.

Here too, as in the 'precise' case, the idea is to extend these local uncertainty models $Q^{\square}$ and $(Q^x)_{x \in \mathcal{X}}$ to a global one, which now—at least in a measure-theoretic setting [De Bock and T'Joens, 2021, Section 2]—takes the form of a *set* of probability measures on the σ-algebra generated by the cylinder events. Crucially, there are three ways to make this transition from local to global information, each leading to a different global uncertainty model—that is, set of probability measures. In the context of the present contribution, their formal definition is not crucial; the interested reader can find more details in, for example, [De Bock and T'Joens, 2021, De Cooman and Hermans, 2008, Hermans and Škulj, 2014]. It suffices to know that the resulting sets of probability measures are the imprecise Markov chain under repetition independence $\mathcal{P}^{\mathrm{ri}}$, complete independence $\mathcal{P}^{\mathrm{ci}}$ and epistemic irrelevance $\mathcal{P}^{\mathrm{ei}}$, and that these models are nested: $\mathcal{P}^{\mathrm{ri}} \subseteq \mathcal{P}^{\mathrm{ci}} \subseteq \mathcal{P}^{\mathrm{ei}}$.

Given such an imprecise Markov chain $\mathcal{P} \in \{\mathcal{P}^{\mathrm{ri}}, \mathcal{P}^{\mathrm{ci}}, \mathcal{P}^{\mathrm{ei}}\}$, a variable $g : \mathcal{X}^{\mathbb{Z}_{\geq 0}} \to \mathbb{R} \cup \{-\infty, +\infty\}$ and some initial state x, one is then often interested in the supremum $\overline{\mathrm{E}}_{\mathcal{P}}(g \mid X_0 = x)$ and infimum $\underline{\mathrm{E}}_{\mathcal{P}}(g \mid X_0 = x)$ of $\mathrm{E}_P(g \mid X_0 = x)$ as P ranges over $\mathcal{P}$. For more details on how to compute these for particular instances of g, we refer the interested reader to [De Bock et al., 2021, Krak, 2022, Sangalli et al., 2026] and references therein. One of the fundamental examples of such variables g are those of the form $f(X_n)$, with f some real-valued function on the state space $\mathcal{X}$. Let us have a closer look at these.

2.2 UPPER TRANSITION OPERATORS

We collect the real-valued functions on $\mathcal{X}$ in $\mathbb{R}^{\mathcal{X}}$; since $\mathcal{X}$ is finite, any such function $f \in \mathbb{R}^{\mathcal{X}}$ is bounded. One important example are indicators: the *indicator* $\mathbb{I}_A$ of a set $A \subseteq \mathcal{X}$ maps $x \in \mathcal{X}$ to 1 if $x \in A$ and to 0 otherwise; to ease our notation, as we did in Example 1, we write $\mathbb{I}_x := \mathbb{I}_{\{x\}}$ for all $x \in \mathcal{X}$.

For imprecise Markov chains under complete independence and epistemic irrelevance, it turns out that one can easily compute the supremum and infimum conditional expectation of $f(X_n)$: for every imprecise Markov chain $\mathcal{P} \in \{\mathcal{P}^{\mathrm{ci}}, \mathcal{P}^{\mathrm{ei}}\}$, initial state $x \in \mathcal{X}$, time point $n \in \mathbb{Z}_{\geq 0}$ and function $f \in \mathbb{R}^{\mathcal{X}}$,

$$\begin{aligned} \overline{\mathrm{E}}_{\mathcal{P}}(f(X_n) \mid X_0 = x) &= \overline{\mathrm{T}}^n f(x) \\ &= -\underline{\mathrm{E}}_{\mathcal{P}}(-f(X_n) \mid X_0 = x), \end{aligned} \tag{1}$$

where the operator $\overline{\mathrm{T}} : \mathbb{R}^{\mathcal{X}} \to \mathbb{R}^{\mathcal{X}}$ maps $f \in \mathbb{R}^{\mathcal{X}}$ to

$$\overline{\mathrm{T}} f : \mathcal{X} \to \mathbb{R} : z \mapsto \sup\{\mathrm{E}_q(f) : q \in Q^z\}. \tag{2}$$

For example, for any target set $A \subseteq \mathcal{X}$, initial state $x \in \mathcal{X}$ and time point $n \in \mathbb{Z}_{\geq 0}$, the supremum and infimum of

$$\{P(X_n \in A \mid X_0 = x) : P \in \mathcal{P}\}$$

are equal to $\overline{\mathrm{T}}^n \mathbb{I}_A(x)$ and $-\overline{\mathrm{T}}^n(-\mathbb{I}_A)(x)$, respectively; these are typically referred to as the upper and lower probability of $X_n \in A$ given $X_0 = x$.

Example 2. In our running example, the upper transition operator $\overline{\mathrm{T}}$ induced by the family $(Q^x)_{x \in \mathcal{X}}$ is given for all $f \in \mathbb{R}^{\mathcal{X}}$ and $x \in \mathcal{X}$ by

$$\overline{\mathrm{T}} f(x) = \begin{cases} f(\mathsf{R}) & \text{if } x = \mathsf{R}, \\ \max\{f(\mathsf{R})\} \cup \{\mathrm{E}_{q_\epsilon}(f) : \epsilon \in [0, \frac{1}{2}]\} & \text{if } x = \mathsf{G}, \\ \max\{f(\mathsf{R}), f(\mathsf{G})\} & \text{if } x = \mathsf{B}. \end{cases}$$

It is not difficult to verify that the operator $\overline{\mathrm{T}}$ defined by (2) is an upper transition operator in the sense of [De Cooman and Hermans, 2008, Hermans and De Cooman, 2012].

Definition 1. An *upper transition operator* is a map $\overline{\mathrm{T}}$ from $\mathbb{R}^{\mathcal{X}}$ to $\mathbb{R}^{\mathcal{X}}$ such that

T1. $\overline{\mathrm{T}}f \le \max f$ for all $f \in \mathbb{R}^{\mathcal{X}}$;

T2. $\overline{\mathrm{T}}(f+g) \le \overline{\mathrm{T}}f + \overline{\mathrm{T}}g$ for all $g \in \mathbb{R}^{\mathcal{X}}$;

T3. $\overline{\mathrm{T}}(\lambda f) = \lambda\overline{\mathrm{T}}f$ for all $\lambda \in \mathbb{R}_{\ge 0}$, $f \in \mathbb{R}^{\mathcal{X}}$.

Even when an upper transition operator $\overline{\mathrm{T}}$ is not derived from a family $(\mathcal{Q}^x)_{x\in\mathcal{X}}$ of sets of probability mass functions through (2), it can still be thought of as being derived from such a family; the point-wise largest such family is that of the closed and convex *credal sets* given by

$$\left\{q \in \Sigma_{\mathcal{X}} : (\forall f \in \mathbb{R}^{\mathcal{X}})\ \mathrm{E}_q(f) \le \overline{\mathrm{T}}f(x)\right\} \quad \text{for all } x \in \mathcal{X},$$

where $\Sigma_{\mathcal{X}}$ is the set of all probability mass functions on $\mathcal{X}$. Following De Bock et al. [2025, Section 5], we call $\overline{\mathrm{T}}$ finitely generated if for every state $x \in \mathcal{X}$, the corresponding credal set is a convex polytope—that is, the convex hull of a finite set of probability mass functions.

Example 3. The sets $\mathcal{Q}^{\mathsf{R}}$, $\mathcal{Q}^{\mathsf{G}}$ and $\mathcal{Q}^{\mathsf{B}}$ and their corresponding credal sets—that is, the closure of their convex hull—are depicted in Figure 1. These figures clearly show that $\overline{\mathrm{T}}$ is *not* finitely generated, since $\mathcal{Q}^{\mathsf{G}}$ is not a convex polytope.

The three defining properties (T1)–(T3) of an upper transition operator $\overline{\mathrm{T}}$ imply a number of other useful properties which we will rely on in the remainder. When stating these, it will be helpful to also use the *conjugate lower transition operator* $\underline{\mathrm{T}}$, defined by

T4. $\underline{\mathrm{T}}f \coloneqq -\overline{\mathrm{T}}(-f)$ for all $f \in \mathbb{R}^{\mathcal{X}}$.

If $\overline{\mathrm{T}}$ is an upper transition operator, then so is its n-fold composition $\overline{\mathrm{T}}^n$, and its conjugate lower transition operator is $\underline{\mathrm{T}}^n$; we will often use this basic fact implicitly in the remainder. From the (well-known) properties of upper and conjugate lower transition operators $\overline{\mathrm{T}}$ and $\underline{\mathrm{T}}$ that De Bock et al. [2025, Section 2.2] list, we will need the following ones: with $\mathrm{S} \in \{\overline{\mathrm{T}}, \underline{\mathrm{T}}\}$,

T5. $\min f \le \mathrm{S}f \le \max f$ for all $f \in \mathbb{R}^{\mathcal{X}}$;

T6. if $f \le g$ then $\mathrm{S}f \le \mathrm{S}g$ for all $f, g \in \mathbb{R}^{\mathcal{X}}$;

T7. $\mathrm{S}(\mu + f) = \mu + \mathrm{S}f$ for all $f \in \mathbb{R}^{\mathcal{X}}$ and $\mu \in \mathbb{R}$;

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

3 CONVERGENCE

Just like for classical Markov chains, for imprecise Markov chains $\mathcal{P}$ one is often interested in $\overline{\mathrm{E}}_{\mathcal{P}}(f(X_n) \mid X_0 = x)$ for large n. For imprecise Markov chains under complete independence and repetition independence, (1) tells us that we then only need to look at $\overline{\mathrm{T}}^n f(x)$ for large n. This motivates why in the remainder of this contribution we will only be interested in the limit behaviour of the sequences $(\overline{\mathrm{T}}^n f)_{n\in\mathbb{N}}$ in $\mathbb{R}^{\mathcal{X}}$ for $f \in \mathbb{R}^{\mathcal{X}}$, given some upper transition operator $\overline{\mathrm{T}}$.

We are by no means the first to investigate this limit behaviour. Hermans and De Cooman [2012] already noted that any upper transition operator $\overline{\mathrm{T}}$ is a '(convex) topological map' [Akian and Gaubert, 2003]—or, even more general, a 'sup-norm non-expansive map' [Sine, 1990, Lemmens and Scheutzow, 2005]—thanks to (T1)–(T3). Consequently, for any $f \in \mathbb{R}^{\mathcal{X}}$, the sequence $(\overline{\mathrm{T}}^n f)_{n\in\mathbb{N}}$ has a finite limit set $\Omega_f = \{\omega_1, \dots, \omega_{p_f}\}$ whose period p_f has a universal upper bound (functionally dependent of $|\mathcal{X}|$). The points in this limit set form a cycle: $\omega_2 = \overline{\mathrm{T}}\omega_1, \dots, \omega_{p_f} = \overline{\mathrm{T}}\omega_{p_f-1}$ and $\omega_1 = \overline{\mathrm{T}}\omega_{p_f} \eqqcolon \omega_{p_f+1}$.

Hermans and De Cooman [2012] introduced the notion of *ergodicity* for upper transition operators, which requires that for all $f \in \mathbb{R}^{\mathcal{X}}$, the sequence $(\overline{\mathrm{T}}^n f)_{n\in\mathbb{N}}$ converges to a constant function, or equivalently, the limit set Ω_f is a singleton (so $p_f = 1$) containing a constant function. Their main result is an easy to check necessary and sufficient condition for ergodicity.

In their wake, De Bock et al. [2025, Definition 2.1] introduced the notion of convergence, which also requires that all limit sets Ω_f are singletons, but drops the requirement that the limit functions should be constant.

Definition 2. The upper transition operator $\overline{\mathrm{T}}$ is *convergent* if for all $f \in \mathbb{R}^{\mathcal{X}}$, $(\overline{\mathrm{T}}^n f)_{n\in\mathbb{N}}$ converges.

De Bock et al. [2025] go on to give an alternative characterisation of convergence in terms of a sufficient condition, which they prove to be also necessary for finitely-generated upper transition operators. To state this condition, we need to introduce some additional terminology. Following T'Joens and De Bock [2021, Section 5], we call any non-empty subset C of $\mathcal{X}$ a *class*.

3.1 ACCESSIBILITY AND REGULARITY

Here and in the remainder, we will need the notion of an accessibility relation, which is a well-known concept in the theory of Markov chains. We essentially follow De Cooman et al. [2009, Section 4], but adhere to a slightly different axiomatisation for accessibility relations; the reader will have no difficulty in verifying the equivalence of their Eqns. (26)–(27) and our (A1)–(A2).

An *accessibility relation* on some class A is a ternary relation $\cdot \xrightarrow{\cdot} \cdot \subseteq A \times \mathbb{N} \times A$ such that for all $x, y, z \in A$ and $n, m \in \mathbb{N}$,

A1. $x \xrightarrow{n} y \wedge y \xrightarrow{m} z \implies x \xrightarrow{n+m} z$;

A2. there is some $y_x \in A$ such that $x \xrightarrow{1} y_x$.

For every two $x, y \in A$, x has access to y if either $x = y$

or $x \xrightarrow{n} y$ for some $n \in \mathbb{N}$, and x and y *communicate* if x has access to y and vice versa; we denote this by $x \to y$ and $x \leftrightarrow y$, respectively. The binary relation $\to$ on A is a preorder, and its associated relation $\leftrightarrow$ is an equivalence relation, which partitions A in equivalence classes, called *communication classes*. The preorder $\to$ on A induces a partial order on this partition, also denoted by $\to$, and the communication classes that are *undominated* in this preorder are also called *maximal*. Such a maximal communication class C is *regular* if there is some $N \in \mathbb{N}$ such that $x \xrightarrow{n} y$ for all $n \geq N$ and $x, y \in C$. In general, we call any class $C \subseteq A$ closed (for $\to$) if no state $x \in C$ has access to a state outside C. Note that every maximal communication class is closed, and that a union of closed classes is again a closed class.

The accessibility relations in this contribution all satisfy an additional property that implies (A1), which makes them particularly easy to visualise: for all $x, y \in A$ and $n \in \mathbb{N}$, $x \xrightarrow{n} y$ if and only if there is some sequence $x = z_0, \dots, z_n = y$ of states in A such that $z_{i-1} \xrightarrow{1} z_i$ for all $i \in \{1, \dots, n\}$. If this is the case, $\cdot \stackrel{\cdot}{\to} \cdot$ is completely determined by the binary relation $\cdot \xrightarrow{1} \cdot$, and we can then visualise it with a directed graph with nodes A and a directed edge from x to y if $x \xrightarrow{1} y$. Checking whether $x \xrightarrow{n} y$ then amounts to checking whether there is a directed path of length n from x to y in this graph. The communication classes of $\cdot \stackrel{\cdot}{\to} \cdot$ can also be easily obtained using this graph: they are exactly the strongly connected components of it. Furthermore, if we define the cyclicity of a maximal communication class C as the greatest common divisor of the lengths of the closed directed paths that remain in this class, then C is regular if and only if it has cyclicity 1.

De Cooman et al. [2009, Section 4.2] give one particular example of an accessibility relation on $\mathcal{X}$: the one defined for all $x, y \in \mathcal{X}$ and $n \in \mathbb{N}$ by

$$x \overset{n}{\rightsquigarrow} y \Leftrightarrow \overline{\mathrm{T}}^n \mathbb{I}_y(x) > 0. \tag{3}$$

Taking into account (1), this means x has access to y in n steps if and only if, according to an imprecise Markov chain $\mathcal{P}$ with upper transition operator $\overline{\mathrm{T}}$, the upper probability of transitioning from x to y in n steps is (strictly) positive. Hermans and De Cooman [2012, Proposition 4] prove that this relation has the additional property that allows us to visualise it with the directed graph with nodes $\mathcal{X}$ and a directed edge from x to y if $x \overset{1}{\rightsquigarrow} y$, which we call the upper accessibility graph and denote by $\mathcal{G}(\overline{\mathrm{T}})$.

Example 4. The upper accessibility graph $\mathcal{G}(\overline{\mathrm{T}})$ of the upper transition operator $\overline{\mathrm{T}}$ is depicted in Figure 2. This graph clearly has two communication classes $\{\mathsf{R}\}$ and $\{\mathsf{G}, \mathsf{B}\}$, and the former dominates the latter. Since the maximal communication class $\{\mathsf{R}\}$ is a singleton, it is trivially regular.

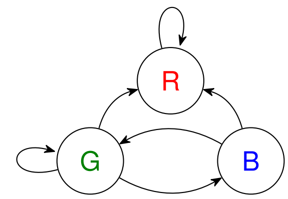


Figure 2: The upper accessibility graph $\mathcal{G}(\overline{\mathrm{T}})$ in the running example.

3.2 LOWER REACHABILITY

To state their sufficient condition for convergence, De Bock et al. [2025, Theorem 3.1] not only need (upper) accessibility, but also the following notion of 'lower reachability' inspired by Hermans and De Cooman [2012, Proposition 3] and T'Joens and De Bock [2021, Section 5]: a class $C \subseteq \mathcal{X}$ is called *lower reachable* from a state $x \in \mathcal{X}$ if there is some $n \in \mathbb{N}$ such that $\underline{\mathrm{T}}^n \mathbb{I}_C(x) > 0$.

For a class C that is closed with respect to $\cdot \stackrel{\cdot}{\rightsquigarrow} \cdot$, the set of states for which this closed class C is lower reachable can be determined with the help of the recursive scheme in [De Bock et al., 2025, Lemma 3.1]—see also Proposition 6 in [Hermans and De Cooman, 2012]. The following result establishes that this recursive scheme is valid for more general classes, which will turn out to be quite helpful further on.

Lemma 3. *Consider a class B such that $\underline{\mathrm{T}}^n \mathbb{I}_B(b) > 0$ for all $n \in \mathbb{N}$ and $b \in B$. For all $n \in \mathbb{Z}_{\geq 0}$, let*

$$B_n := \{x \in \mathcal{X} : \underline{\mathrm{T}}^n \mathbb{I}_B(x) > 0\}.$$

Then for all $n \in \mathbb{Z}_{\geq 0}$, $B \subseteq B_n$ and

$$\begin{aligned} B_{n+1} &= \{x \in \mathcal{X} : \underline{\mathrm{T}} \mathbb{I}_{B_n}(x) > 0\} \\ &= B_n \cup \{x \in B_n^{\mathrm{c}} : \underline{\mathrm{T}} \mathbb{I}_{B_n}(x) > 0\}. \end{aligned}$$

Consequently, there is some $k \leq |B^{\mathrm{c}}|$ for which $B_k = B_{k+1}$, and B_k is the set of states from which B is lower reachable.

For the sake of completeness, we verify that any closed class C satisfies the condition in Lemma 3.

Lemma 4. *For every closed class C,*

$$\underline{\mathrm{T}}^n \mathbb{I}_C(x) = 1 \quad \textit{for all } n \in \mathbb{N}, x \in \mathcal{X}.$$

Example 5. Recall that $\{\mathsf{R}\}$ is the unique maximal communication class of $\cdot \stackrel{\cdot}{\rightsquigarrow} \cdot$, which is of course closed. Following Lemma 3.1 in [De Bock et al., 2025]—or Lemmas 3 and 4 here—we look at the sequence $(B_n)_{n \in \mathbb{Z}_{\geq 0}}$ defined by the initial condition $B_0 := C$ and, for all $n \in \mathbb{Z}_{\geq 0}$, the recursive relation

$$B_{n+1} := B_n \cup \{x \in B_n^{\mathrm{c}} : \underline{\mathrm{T}} \mathbb{I}_{B_n}(x) > 0\}.$$

Since it follows from (T8) that

$$\underline{\mathrm{T}}\mathbb{I}_{\mathsf{R}} = 1 - \overline{\mathrm{T}}\mathbb{I}_{\{\mathsf{G},\mathsf{B}\}} = 1 - \mathbb{I}_{\{\mathsf{G},\mathsf{B}\}} = \mathbb{I}_{\mathsf{R}},$$

it follows from this definition that $B_n = \{\mathsf{R}\}$ for all $n \in \mathbb{Z}_{\geq 0}$. Consequently, the unique maximal communication class $\{\mathsf{R}\}$ is only lower reachable from $\{\mathsf{R}\}$.

As the previous example shows, the class B_k obtained by applying the algorithm in Lemma 3 need not yield the entire state space $\mathcal{X}$, but possibly only a subset of it; that is, B need not be lower reachable from all states. The set B_k^{c} of states from which B is not lower reachable is then of course non-empty and has the interesting property of being a so-called 'filtering class'.

Definition 5. A *filtering class* is a class A such that

$$A = \{x \in \mathcal{X} \colon \overline{\mathrm{T}}\mathbb{I}_A(x) = 1\} = \{x \in \mathcal{X} \colon \underline{\mathrm{T}}\mathbb{I}_{A^{\mathrm{c}}}(x) = 0\}. \quad (4)$$

That this is indeed the case for B_k^{c} follows from the fact that $B_k = B_{k+1}$, and the relation between B_n and B_{n+1} given in Lemma 3. Since $\overline{\mathrm{T}}\mathbb{I}_{\mathcal{X}} = \overline{\mathrm{T}}1 = 1$ by (T5), another example of a filtering class is the entire state space $\mathcal{X}$.

At first sight the term 'filtering' might be somewhat far-fetched, but it has some intuitive motivation behind it: when you are inside such a class, the upper probability of staying inside is one, and when you are outside of it, the upper probability of returning to it is less than one, which is why we can think of such a class as a 'filter' that allows probability mass to stay inside it, but filters out at least part of the probability mass that tries to enter it from outside.

Lemma 6. *For every class A, the following statements are equivalent:*

(i) *A is a filtering class;*

(ii) *$A = \{x \in \mathcal{X} \colon (\forall n \in \mathbb{Z}_{\geq 0})\ \overline{\mathrm{T}}^n\mathbb{I}_A(x) = 1\}$;*

(iii) *$A^{\mathrm{c}} = \{x \in \mathcal{X} \colon (\forall n \in \mathbb{Z}_{\geq 0})\ \overline{\mathrm{T}}^n\mathbb{I}_A(x) < 1\}$;*

(iv) *$(\exists \beta_A \in \,]0,1])(\forall n \in \mathbb{Z}_{\geq 0})\ \mathbb{I}_{A^{\mathrm{c}}} \geq \underline{\mathrm{T}}^n\mathbb{I}_{A^{\mathrm{c}}} \geq \beta_A^n\mathbb{I}_{A^{\mathrm{c}}}$.*

Example 6. Recall that the unique maximal communication class $\{\mathsf{R}\}$ is only lower reachable from $\{\mathsf{R}\}$. This implies that $\{\mathsf{G},\mathsf{B}\}$ is a filtering class. Indeed, for $x = \mathsf{G}$ or $x = \mathsf{B}$, $\overline{\mathrm{T}}\mathbb{I}_{\{\mathsf{G},\mathsf{B}\}}(x) = 1$, meaning that there is at least one compatible probability model that keeps all probability mass in $\{\mathsf{G},\mathsf{B}\}$; for $x = \mathsf{R}$, on the other hand, $\overline{\mathrm{T}}\mathbb{I}_{\{\mathsf{G},\mathsf{B}\}}(\mathsf{R}) < 1$, meaning that probability mass in $\{\mathsf{R}\}$ cannot be entirely transferred to $\{\mathsf{G},\mathsf{B}\}$.

3.3 A SUFFICIENT CONDITION FOR CONVERGENCE

De Bock et al. [2025] consider a number of different cases for which they establish sufficient (and sometimes necessary) conditions for convergence. The distinction between these cases is based on the structure of the communication classes of $\cdot \overset{\cdot}{\rightsquigarrow} \cdot$ and their lower reachability.

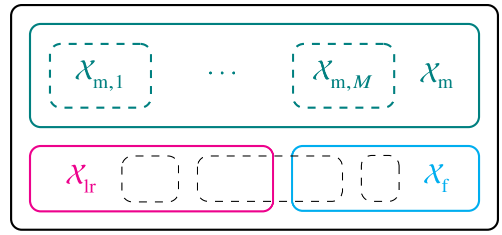


Figure 3: Partitioning of the state space $\mathcal{X}$, where dashed lines indicate the communication classes of $\cdot \overset{\cdot}{\rightsquigarrow} \cdot$.

They start by considering the maximal communication classes $\mathcal{X}_{\mathrm{m},1}, \dots, \mathcal{X}_{\mathrm{m},M}$ corresponding to $\cdot \overset{\cdot}{\rightsquigarrow} \cdot$, and their union $\mathcal{X}_{\mathrm{m}} \coloneqq \mathcal{X}_{\mathrm{m},1} \cup \cdots \cup \mathcal{X}_{\mathrm{m},M}$. Next, they subdivide the remaining states in $\mathcal{X} \setminus \mathcal{X}_{\mathrm{m}}$ into two sets: those from which $\mathcal{X}_{\mathrm{m}}$ is lower reachable, and those from which it isn't, denoted by $\mathcal{X}_{\mathrm{lr}}$ and $\mathcal{X}_{\mathrm{f}}$, respectively. As indicated by the subscript f, the set $\mathcal{X}_{\mathrm{f}}$ will then be a filtering class. Figure 3 provides a schematic depiction of the resulting partition of $\mathcal{X}$.

Example 7. In our running example, there is only one maximal communication class $\{\mathsf{R}\}$, so $\mathcal{X}_{\mathrm{m}} = \mathcal{X}_{\mathrm{m},1} = \{\mathsf{R}\}$. Since $\mathcal{X}_{\mathrm{m}}$ is not lower reachable from either of the remaining states G and B, we have $\mathcal{X}_{\mathrm{lr}} = \emptyset$ and $\mathcal{X}_{\mathrm{f}} = \{\mathsf{G},\mathsf{B}\}$.

The least restrictive assumption under which they provide a necessary and sufficient condition for convergence, is the assumption that $\mathcal{X}_{\mathrm{f}} = \emptyset$, meaning that the union of maximal communication classes is lower reachable from every state $x \in \mathcal{X}$. In that case, they find that $\overline{\mathrm{T}}$ is convergent if and only each of the maximal communication classes $\mathcal{X}_{\mathrm{m},\ell}, \dots, \mathcal{X}_{\mathrm{m},M}$ is regular [De Bock et al., 2025, Corollary 4.2]. The intuition behind this result is that if the union of maximal communication classes is lower reachable from every state, then the limit behaviour of $(\overline{\mathrm{T}}^n f)_{n\in\mathbb{N}}$ is completely determined by the limit behaviour of $(\overline{\mathrm{T}}^n f)_{n\in\mathbb{N}}$ on these maximal communication classes, and the latter is completely determined by their regularity; in particular, $(\overline{\mathrm{T}}^n f)_{n\in\mathbb{N}}$ converges on a maximal communication class if and only if this class is regular.

As illustrated by our example, the union of maximal communication classes is not always lower reachable from every state. Whenever this is the case, the limit behaviour of $(\overline{\mathrm{T}}^n f)_{n\in\mathbb{N}}$ is also influenced by the behaviour on the remaining states $\mathcal{X}_{\mathrm{f}}$; Theorem 4.1 in [De Bock et al., 2025] then provides a sufficient condition for convergence that is based on a further decomposition of $\mathcal{X}_{\mathrm{f}}$. De Bock et al. [2025, Theorem 5.1] show that this sufficient condition is also necessary if $\overline{\mathrm{T}}$ is finitely generated, and give a counterexample—which we've adopted as running example—which shows that this condition is not necessary.

The first part of their sufficient condition is again the requirement that, similarly to the case where $\mathcal{X}_\mathrm{f} = \emptyset$, the maximal communication classes $\mathcal{X}_{\mathrm{m},1}, \dots, \mathcal{X}_{\mathrm{m},M}$ need to be regular, but this is now no longer sufficient for convergence. The problem is that $\mathcal{X}_\mathrm{f}$ is a filtering class, which implies that it is possible to stay in $\mathcal{X}_\mathrm{f}$ for an arbitrarily long time, with upper probability one. To guarantee convergence, we therefore need to ensure that 'the behaviour on $\mathcal{X}_\mathrm{f}$' is convergent as well.

To achieve this, De Bock et al. [2025, Equation (2) and Lemma 4.2] introduce a new upper transition operator $\overline{\mathrm{T}}_\mathrm{f}$ that is obtained by 'cutting' $\mathcal{X}_\mathrm{f}$ from $\mathcal{X}$ and deriving an upper transition operator on $\mathbb{R}^{\mathcal{X}_\mathrm{f}}$ from the original one $\overline{\mathrm{T}}$. More formally, for all $x \in \mathcal{X}_\mathrm{f}$, with $\mathcal{Q}^x_\mathrm{f} \subseteq \Sigma_{\mathcal{X}_\mathrm{f}}$ the set of probability mass functions on $\mathcal{X}_\mathrm{f}$ obtained by taking the restriction to $\mathcal{X}_\mathrm{f}$ of every probability mass function in the closure of $\mathcal{Q}^x$ that assigns zero mass to $\mathcal{X}^\mathrm{c}_\mathrm{f}$, they define

$$\overline{\mathrm{T}}_\mathrm{f} g(x) \coloneqq \sup\{\mathrm{E}_q(g) \colon q \in \mathcal{Q}^x_\mathrm{f}\} \quad \text{for all } g \in \mathbb{R}^{\mathcal{X}_\mathrm{f}}. \quad (5)$$

Since this new upper transition operator $\overline{\mathrm{T}}_\mathrm{f}$ captures the behaviour inside $\mathcal{X}_\mathrm{f}$, they then proceed to require that $\overline{\mathrm{T}}_\mathrm{f}$ is convergent, and show that this, together with the regularity of the maximal communication classes, is sufficient for convergence of $\overline{\mathrm{T}}$. In this way, the problem has been reduced from checking the convergence of $\overline{\mathrm{T}}$ to checking the convergence of $\overline{\mathrm{T}}_\mathrm{f}$. This problem can then be similarly decomposed, until there are no more states from which the union of maximal communication classes is not lower reachable, in which case we can apply the result for that case.

Example 8. In our running example, with $\mathcal{X}_\mathrm{f} = \{\mathsf{G}, \mathsf{B}\}$ and $\mathcal{X}^\mathrm{c}_\mathrm{f} = \{\mathsf{R}\}$, the upper transition operator $\overline{\mathrm{T}}_\mathrm{f}$ obtained by cutting $\overline{\mathrm{T}}$ on $\mathcal{X}_\mathrm{f}$ is rather simple. The only probability mass function in $\mathcal{Q}^\mathsf{G}$ that assigns probability zero to R is $\mathbb{I}_\mathsf{B}$, and similarly, the only probability mass function in $\mathcal{Q}^\mathsf{B}$ that assigns zero mass to R is $\mathbb{I}_\mathsf{G}$. Consequently, the upper transition operator $\overline{\mathrm{T}}_\mathrm{f}$ is given by $\overline{\mathrm{T}}_\mathrm{f} g(\mathsf{G}) = g(\mathsf{B})$ and $\overline{\mathrm{T}}_\mathrm{f} g(\mathsf{B}) = g(\mathsf{G})$ for all $g \in \mathbb{R}^{\{\mathsf{G},\mathsf{B}\}}$; its upper accessibility graph $\mathcal{G}(\overline{\mathrm{T}}_\mathrm{f})$ is depicted in Figure 4. This graph has a single (and hence maximal) communication class $\{\mathsf{G}, \mathsf{B}\}$, which is not regular since G only has acces to itself after an even number of steps, and similarly for B. Consequently, $\overline{\mathrm{T}}_\mathrm{f}$ is not convergent, and therefore the sufficient condition of De Bock et al. [2025, Theorem 4.1] for convergence of $\overline{\mathrm{T}}$ is not satisfied. This implies that this sufficient condition is not necessary, because $\overline{\mathrm{T}}$ is in fact convergent, as De Bock et al. [2025, Example 5.1] demonstrate by explicitly verifying the definition of convergence—that is, after some tedious calculations they painstakingly managed to show that for all $f \in \mathbb{R}^\mathcal{X}$, the iterates $(\overline{\mathrm{T}}^n f)_{n\in\mathbb{N}}$ converge to a limit.

Motivated by this counterexample, we will now develop a slightly weaker condition that is both necessary and sufficient for convergence, even without the assumption that $\overline{\mathrm{T}}$ is finitely generated.

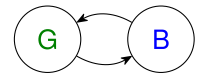


Figure 4: The graph $\mathcal{G}(\overline{\mathrm{T}}_\mathrm{f})$ for our running example.

4 A NECESSARY AND SUFFICIENT CONDITION

The reason behind the fact that the sufficient condition of De Bock et al. [2025, Theorem 4.1] is not necessary, is that it requires the behaviour inside $\mathcal{X}_\mathrm{f}$ to be convergent: as we saw in our counterexample, even if the behaviour inside $\mathcal{X}_\mathrm{f}$ is cyclic—and hence not convergent—it's still possible for $\overline{\mathrm{T}}$ to be convergent. The intuitive reason why this is the case is rather subtle, and best understood by looking at the graphical representation of $\mathcal{Q}^\mathsf{G}$ in Figure 1. There is only one probability mass function that assigns probability zero to R, and this probability mass function assigns all mass to B, forcing a step to B that, combined with the reversed edge, leads to cyclic behaviour inside $\mathcal{X}_\mathrm{f}$. If we would also have an edge from G to itself, this cycle could be broken, making $\{\mathsf{G}, \mathsf{B}\}$ regular and $\overline{\mathrm{T}}$ convergent. Of course, such an edge from G to itself would require a probability mass function in $\mathcal{Q}^\mathsf{G}$ that assigns mass to G but not to R, which is not the case. However, the crucial insight is that this is almost the case, in the sense that there are probability mass functions in $\mathcal{Q}^\mathsf{G}$ that do assign mass to G, at the 'cost' of assigning only a comparatively negligible amount of mass to R. This is a result of the curvature of the boundary of $\mathcal{Q}^\mathsf{G}$ in Figure 1, which is tangent to the line connecting $\mathbb{I}_\mathsf{B}$ and $\mathbb{I}_\mathsf{G}$. Inspired by this observation, we will now develop a slightly weaker notion of accessibility that allows us to capture this idea of 'almost accessibility' relative to a filtering class, and use it to establish a necessary and sufficient condition for convergence that is weaker than the sufficient one of De Bock et al. [2025, Theorem 4.1].

4.1 ACCESSIBILITY REVISITED

The accessibility relation we will use is a more general form of the 'standard' one for upper transition operators defined in (3). For any class A, we consider the corresponding ternary relation $\cdot \xrightarrow{A,\cdot} \cdot$ defined for all $x, y \in A$ and $n \in \mathbb{N}$ by

$$x \xrightarrow{A,n} y \Leftrightarrow (\forall \alpha \in \mathbb{R}_{>0})\, \overline{\mathrm{T}}^n(-\alpha \mathbb{I}_{A^\mathrm{c}} + \mathbb{I}_y)(x) > 0.$$

Note that for $A = \mathcal{X}$, this definition reduces to the one in (3); for $A \subset \mathcal{X}$, the extra term $-\alpha \mathbb{I}_{A^\mathrm{c}}$ in the argument of $\overline{\mathrm{T}}^n$, with $\alpha \in \mathbb{R}_{>0}$ arbitrarily large, formalises the above idea of allowing one to assign mass to A^c as long as it is negligible compared to the mass assigned to y.

We now set out to show that $\cdot \xrightarrow{A,\cdot} \cdot$ is indeed an accessibility relation and that it furthermore has the additional property that allows us to visualise it with a directed graph, at least under the assumption that A is a filtering class. A first consequence of this assumption is that our ternary relation satisfies (A2).

Lemma 7. *If A is a filtering class, then for all $a \in A$ there is some $x \in A$ such that $a \xrightarrow{A,1} x$.*

Property (A1) follows immediately from the following generalisation of Proposition 4 in [Hermans and De Cooman, 2012].

Lemma 8. *Consider some filtering class A. Then for all $x, y \in \mathcal{X}$ and $n \in \mathbb{N}$, $x \xrightarrow{A,n} y$ if and only if there is some sequence $x = z_0, \ldots, z_n = y$ in $\mathcal{X}$ such that $z_{k-1} \xrightarrow{A,1} z_k$ for all $k \in \{1, \ldots, n\}$.*

Since the ternary relation $\cdot \xrightarrow{A,\cdot} \cdot$ satisfies (A1) and (A2), it is indeed an accessibility relation.

Proposition 9. *For every filtering class A, $\cdot \xrightarrow{A,\cdot} \cdot$ is an accessibility relation.*

For any filtering class A, Lemma 8 also tells us that the corresponding accessibility relation $\cdot \xrightarrow{A,\cdot} \cdot$ has the additional property mentioned in Section 3.1, so it is completely determined by the graph with A as nodes and a directed edge from x to y if $x \xrightarrow{A,1} y$, which we will denote by $\mathcal{G}_A(\overline{\mathrm{T}})$.

Moreover, if $\overline{\mathrm{T}}$ is finitely generated, this graph $\mathcal{G}_A(\overline{\mathrm{T}})$ is identical to the upper accessibility graph $\mathcal{G}(\overline{\mathrm{T}}_A)$ of the upper transition operator $\overline{\mathrm{T}}_A$ obtained by 'cutting' A from $\mathcal{X}$. Generalising (5) from $\mathcal{X}_{\mathrm{f}}$ to arbitrary filtering classes A, De Bock et al. [2025, Equation (2) and Lemma 4.2] define this upper transition operator $\overline{\mathrm{T}}_A \colon \mathbb{R}^A \to \mathbb{R}^A$ as

$$\overline{\mathrm{T}}_A g(a) \coloneqq \max\{\mathrm{E}_q(g) \colon q \in \mathcal{Q}_A^a\} \quad \text{for all } g \in \mathbb{R}^A, a \in A,$$

where $\mathcal{Q}_A^a \subseteq \Sigma_A$ is obtained by taking the restriction to A of all probability mass functions in the closure of $\mathcal{Q}^a$ that assign zero mass to A^{c}.

Lemma 10. *If A is a filtering class and $\overline{\mathrm{T}}$ is finitely generated, then $\mathcal{G}_A(\overline{\mathrm{T}}) = \mathcal{G}(\overline{\mathrm{T}}_A)$.*

If $\overline{\mathrm{T}}$ is not finitely generated, however, $\mathcal{G}_A(\overline{\mathrm{T}})$ and $\mathcal{G}(\overline{\mathrm{T}}_A)$ can differ: the former can (only) have *more* edges than the latter. Indeed, take any directed edge in $\mathcal{G}(\overline{\mathrm{T}}_A)$ from x to y, meaning that $\overline{\mathrm{T}}_A \mathbb{I}_y(x) > 0$, and let $\mathcal{Q}^x$ be the (closed) credal set of x that corresponds to $\overline{\mathrm{T}}$. Then by definition of $\overline{\mathrm{T}}_A$, there is some probability mass function q in $\mathcal{Q}^x$ such that $\mathrm{E}_q(\mathbb{I}_{A^{\mathrm{c}}}) = 0$ and $q(y) > 0$. Therefore, for all $\alpha \in \mathbb{R}_{>0}$,

$$\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^{\mathrm{c}}} + \mathbb{I}_y)(x) \geq \mathrm{E}_q(-\alpha\mathbb{I}_{A^{\mathrm{c}}} + \mathbb{I}_y) = -\alpha\mathrm{E}_q(\mathbb{I}_{A^{\mathrm{c}}}) + q(y) > 0,$$

so there is a directed edge from x to y in $\mathcal{G}_A(\overline{\mathrm{T}})$. Our running example illustrates nicely that this inclusion $\mathcal{G}(\overline{\mathrm{T}}_A) \subseteq \mathcal{G}_A(\overline{\mathrm{T}})$ is sometimes a strict one.

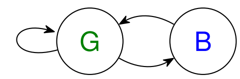


Figure 5: The graph $\mathcal{G}_{\{\mathsf{G},\mathsf{B}\}}(\overline{\mathrm{T}})$ for our running example.

Example 9. For the filtering class $A = \mathcal{X}_{\mathrm{f}} = \{\mathsf{G}, \mathsf{B}\}$, $\overline{\mathrm{T}}_A$ is the operator $\overline{\mathrm{T}}_{\mathrm{f}}$ defined by (5), whose upper accessibility graph $\mathcal{G}(\overline{\mathrm{T}}_{\{\mathsf{G},\mathsf{B}\}})$ is depicted in Figure 4.

As an intermediary step towards determining $\mathcal{G}_{\{\mathsf{G},\mathsf{B}\}}(\overline{\mathrm{T}})$, we note that $\mathbb{I}_{A^{\mathrm{c}}} = \mathbb{I}_{\mathsf{R}}$ and evaluate $\overline{\mathrm{T}}$ in $-\alpha\mathbb{I}_{\mathsf{R}} + \mathbb{I}_y$ with $\alpha \in \mathbb{R}_{>0}$ and $y \in \{\mathsf{G}, \mathsf{B}\}$. The values of $\overline{\mathrm{T}}(-\alpha\mathbb{I}_{\mathsf{R}} + \mathbb{I}_y)(x)$ are easy to determine for $x = \mathsf{B}$: for all $\alpha \in \mathbb{R}_{>0}$,

$$\overline{\mathrm{T}}(-\alpha\mathbb{I}_{\mathsf{R}} + \mathbb{I}_{\mathsf{G}})(\mathsf{B}) = \max\{-\alpha, 1\} = 1$$

and

$$\overline{\mathrm{T}}(-\alpha\mathbb{I}_{\mathsf{R}} + \mathbb{I}_{\mathsf{B}})(\mathsf{B}) = \max\{-\alpha, 0\} = 0.$$

Those for $x = \mathsf{G}$ are a bit more involved: for all $\alpha \in \mathbb{R}_{>0}$,

$$\begin{aligned}
&\overline{\mathrm{T}}(-\alpha\mathbb{I}_{\mathsf{R}} + \mathbb{I}_{\mathsf{G}})(\mathsf{G}) \\
&\qquad = \max\{-\alpha\} \cup \left\{-\alpha\epsilon^2 + \epsilon \colon \epsilon \in \left[0, \frac{1}{2}\right]\right\} \\
&\qquad = \begin{cases} \frac{2-\alpha}{4} & \text{if } \alpha \leq 1 \\ \frac{1}{4\alpha} & \text{otherwise} \end{cases}
\end{aligned}$$

and

$$\begin{aligned}
&\overline{\mathrm{T}}(-\alpha\mathbb{I}_{\mathsf{R}} + \mathbb{I}_{\mathsf{B}})(\mathsf{G}) \\
&\qquad = \max\{-\alpha\} \cup \left\{-(1+\alpha)\epsilon^2 + 1 - \epsilon \colon \epsilon \in \left[0, \frac{1}{2}\right]\right\} \\
&\qquad = 1.
\end{aligned}$$

This leads to the graph $\mathcal{G}_{\{\mathsf{G},\mathsf{B}\}}(\overline{\mathrm{T}})$ depicted in Figure 5.

Compared to the graph $\mathcal{G}(\overline{\mathrm{T}}_{\{\mathsf{G},\mathsf{B}\}})$ in Figure 4, this new graph has an additional directed edge from G to itself, thereby capturing the intuition that G is 'almost accessible' from itself, as hinted at in the introduction of this section.

From here on, we will assume that all terminology regarding accessibility relations—'has access to', 'communicate', 'maximal', 'closed', 'regular', ...—are with respect to the relation $\cdot \xrightarrow{A,\cdot} \cdot$; to avoid confusion, we will often add the suffix 'A-' to indicate the class A from which the relation is derived.

4.2 STATING THE CONDITION

To state our necessary and sufficient condition for convergence, we iteratively partition the state space $\mathcal{X}$ using the accessibility and lower reachability relations we have introduced in the previous sections, in much the same way as De Bock et al. [2025, Section 4.4] do for their condition.

We start off in exactly the same way as them—see also Section 3.3. For reasons that will become clear further on, we let $A_1 \coloneqq \mathcal{X}$. First, we enumerate the maximal communication classes of $\cdot \overset{\cdot}{\rightsquigarrow} \cdot = \cdot \xrightarrow{A_1,\cdot} \cdot$ as $\mathcal{X}^1_{\mathrm{m},1}, \dots, \mathcal{X}^1_{\mathrm{m},M_1}$ and denote their union by $\mathcal{X}^1_{\mathrm{m}}$. Next, since $\mathcal{X}^1_{\mathrm{m}}$ is closed for $\cdot \overset{\cdot}{\rightsquigarrow} \cdot$, we can distinguish the remaining states in $\mathcal{X} \setminus \mathcal{X}^1_{\mathrm{m}}$ according to whether $\mathcal{X}^1_{\mathrm{m}}$ is lower reachable from them: $\mathcal{X}^1_{\mathrm{lr}}$ collects those from which $\mathcal{X}^1_{\mathrm{m}}$ is lower reachable and $\mathcal{X}^1_{\mathrm{f}}$ those from which it is not. Recall from Section 3.2 that thanks to Lemma 4, we can determine these two parts using the recursive scheme in Lemma 3.

If $\mathcal{X}^1_{\mathrm{f}} = \emptyset$, we are done. If $\mathcal{X}^1_{\mathrm{f}} \neq \emptyset$, Lemma 3 also tells us that

$$\mathcal{X}^1_{\mathrm{m}} \cup \mathcal{X}^1_{\mathrm{lr}} = \left\{x \in \mathcal{X} \colon \underline{\mathrm{T}}\mathbb{I}_{\mathcal{X}^1_{\mathrm{m}} \cup \mathcal{X}^1_{\mathrm{lr}}}(x) > 0\right\};$$

in other words, $\mathcal{X}^1_{\mathrm{f}} = (\mathcal{X}^1_{\mathrm{m}} \cup \mathcal{X}^1_{\mathrm{lr}})^{\mathrm{c}}$ is a filtering class! Consequently, we can partition $A_2 \coloneqq \mathcal{X}^1_{\mathrm{f}}$ as before, only this time using the accessibility relation $\cdot \xrightarrow{A_2,\cdot} \cdot$ rather than $\cdot \xrightarrow{A_1,\cdot} \cdot$. First we enumerate the A_2-maximal communication classes as $\mathcal{X}^2_{\mathrm{m},1}, \dots, \mathcal{X}^2_{\mathrm{m},M_2}$ and denote their union by $\mathcal{X}^2_{\mathrm{m}}$. Second, we collect those states in $A_2 \setminus \mathcal{X}^2_{\mathrm{m}}$ from which $B_2 \coloneqq A_2^{\mathrm{c}} \cup \mathcal{X}^2_{\mathrm{m}}$ is lower reachable in $\mathcal{X}^2_{\mathrm{lr}}$ and those from which it is not in $\mathcal{X}^2_{\mathrm{f}}$. Since $\mathcal{X}^2_{\mathrm{m}}$ is A_2-closed, the following result allows us to use the recursive scheme in Lemma 3 to do so.

Lemma 11. *Consider some filtering class A and some A-closed class $C \subseteq A$, and let $B \coloneqq A^{\mathrm{c}} \cup C$. Then $\underline{\mathrm{T}}^n \mathbb{I}_B(b) > 0$ for all $b \in B$ and $n \in \mathbb{N}$.*

If $\mathcal{X}^2_{\mathrm{f}} = \emptyset$, we are done. If $\mathcal{X}^2_{\mathrm{f}} \neq \emptyset$, the way forward should be clear: we again partition the remainder into 'maximal', 'lower reachable' and 'non-lower reachable' parts, now using the filtering class $A_3 \coloneqq \mathcal{X}^2_{\mathrm{f}} = (A_2^{\mathrm{c}} \cup \mathcal{X}^2_{\mathrm{m}} \cup \mathcal{X}^2_{\mathrm{lr}})^{\mathrm{c}}$. Since $\mathcal{X}$ is finite, after d iterations we eventually reach a point where $\mathcal{X}^d_{\mathrm{f}} = \emptyset$. See Figure 6 for a schematic depiction of the obtained partition.

With this decomposition in place, our necessary and sufficient condition for convergence now takes the following elegant form.

Theorem 12. *An upper transition operator $\overline{\mathrm{T}}$ is convergent if and only if for all $k \in \{1, \dots, d\}$, the A_k-maximal communication classes $\mathcal{X}^k_{\mathrm{m},1}, \dots, \mathcal{X}^k_{\mathrm{m},M_k}$ are A_k-regular.*

To illustrate this result, we apply it to two examples. We start with our running example, finally confirming that the upper transition operator that we have been considering is indeed convergent.

Example 10. Since $\cdot \overset{\cdot}{\rightsquigarrow} \cdot = \cdot \xrightarrow{A_1,\cdot} \cdot$ with $A_1 = \mathcal{X}$, it follows from before that $\mathcal{X}^1_{\mathrm{m}} = \mathcal{X}^1_{\mathrm{m},1} = \{\mathsf{R}\}$, $\mathcal{X}^1_{\mathrm{lr}} = \emptyset$ and $A_2 = \mathcal{X}^1_{\mathrm{f}} = \{\mathsf{G}, \mathsf{B}\}$. From the graph $\mathcal{G}_{\{\mathsf{G},\mathsf{B}\}}(\overline{\mathrm{T}})$ in Figure 5, we furthermore easily deduce that $\cdot \xrightarrow{A_2,\cdot} \cdot$ has a single

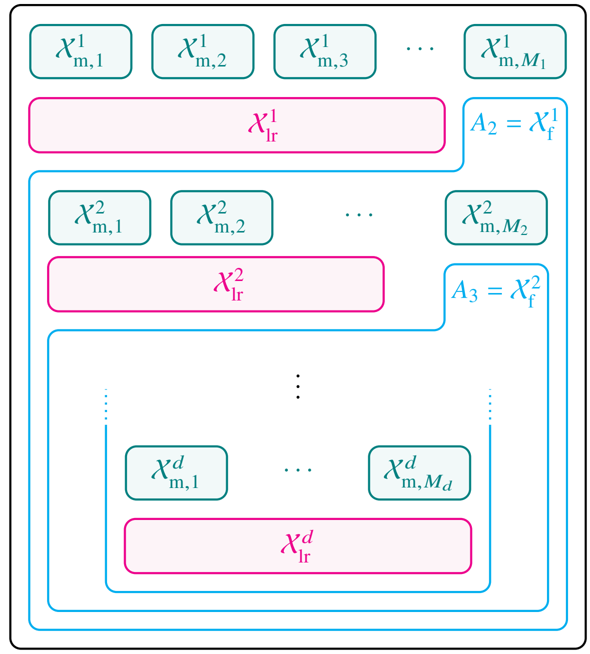


Figure 6: Schematic depiction of the complete partition.

communication class $\{\mathsf{G}, \mathsf{B}\}$, which is A_2-regular. Since the unique A_1-maximal communication class $\mathcal{X}^1_{\mathrm{m},1} = \{\mathsf{R}\}$ is A_1-regular and the unique A_2-maximal communication class $\mathcal{X}^2_{\mathrm{m},1} = \{\mathsf{G}, \mathsf{B}\}$ is A_2-regular, it follows from Theorem 12 that $\overline{\mathrm{T}}$ is indeed convergent.

Our final example, by which we conclude this contribution, is a bit more involved. It demonstrates how our notion of 'almost accessibility' might appear in practical scenarios and illustrates that our decomposition, as depicted in Figure 6, is different from the one of De Bock et al. [2025, Section 4.4].

Example 11. We consider a system that can be in one of four states: A, B, W and F, so $\mathcal{X} \coloneqq \{\mathsf{A}, \mathsf{B}, \mathsf{W}, \mathsf{F}\}$. The system typically switches back and forth between the two regime states A and B. From either of these states, there can however also be a small chance of switching to the warning state W. From this warning state, a restart can give us a chance of returning to state A, but this also comes with a small risk of the system failing and ending up in the failure state F, although this risk is negligible compared to the chance of returning to A. Once we are in the failure state F, the system remains in that state forever. This behaviour is captured by the following sets of transition mass functions:

$$\begin{aligned}
Q^{\mathsf{A}} &= \{(1-p)\mathbb{I}_{\mathsf{B}} + p\mathbb{I}_{\mathsf{W}} \colon p \in [0, \overline{p}_{\mathsf{A}}]\}, \\
Q^{\mathsf{B}} &= \{(1-p)\mathbb{I}_{\mathsf{A}} + p\mathbb{I}_{\mathsf{W}} \colon p \in [0, \overline{p}_{\mathsf{B}}]\}, \\
Q^{\mathsf{W}} &= \{p\mathbb{I}_{\mathsf{A}} + o(p)\mathbb{I}_{\mathsf{F}} + (1-p-o(p))\mathbb{I}_{\mathsf{W}} \colon p \in [0, \overline{p}_{\mathsf{W}}]\}, \\
Q^{\mathsf{F}} &= \{\mathbb{I}_{\mathsf{F}}\}.
\end{aligned}$$

In these expressions, $o \colon [0, \overline{p}_{\mathsf{W}}] \to [0, 1]$ is a continuous function such that $o(p) > 0$ for $p > 0$ but $\lim_{p \to 0} o(p)/p =$

0, which expresses that the probability $o(p)$ of going to the failure state F is positive (unless $p = 0$) but negligible compared to the probability p of returning to A from W. For example, we could take $o(p) = p^a$ for some $a > 1$, $o(p) = x\sin(x)$ or $o(p) = \sin^2(x)$. We also need to specify an upper probability $\overline{p}_{\mathsf{A}} > 0$ and $\overline{p}_{\mathsf{B}} > 0$ for switching to the warning state W from A and B, respectively, as well as an upper probability $\overline{p}_{\mathsf{W}} > 0$ for switching to A from W. These three upper probabilities should be small enough to ensure that Q^{A}, Q^{B} and Q^{W} are indeed sets of probability mass functions.

For all $f \in \mathbb{R}^{\mathcal{X}}$, the upper transition operator $\overline{\mathrm{T}}$ that corresponds to this system is given by $\overline{\mathrm{T}}f(\mathsf{F}) = f(\mathsf{F})$,

$$\overline{\mathrm{T}}f(\mathsf{A}) = \begin{cases} f(\mathsf{B}) & \text{if } f(\mathsf{B}) \geq f(\mathsf{W}), \\ f(\mathsf{B}) + \overline{p}_{\mathsf{A}}(f(\mathsf{W}) - f(\mathsf{B})) & \text{otherwise,} \end{cases}$$

$$\overline{\mathrm{T}}f(\mathsf{B}) = \begin{cases} f(\mathsf{A}) & \text{if } f(\mathsf{A}) \geq f(\mathsf{W}), \\ f(\mathsf{A}) + \overline{p}_{\mathsf{B}}(f(\mathsf{W}) - f(\mathsf{A})) & \text{otherwise,} \end{cases}$$

and

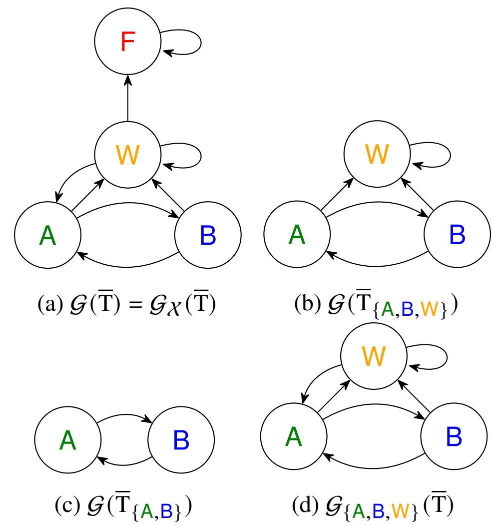
$$\begin{aligned}\overline{\mathrm{T}}f(\mathsf{W}) = \max\{&f(\mathsf{W}) + p(f(\mathsf{A}) - f(\mathsf{W})) \\ &+ o(p)(f(\mathsf{F}) - f(\mathsf{W})) : p \in [0, \overline{p}_{\mathsf{W}}]\}.\end{aligned}$$

To check whether $\overline{\mathrm{T}}$ is convergent, we first determine the upper accessibility graph $\mathcal{G}(\overline{\mathrm{T}})$; this is an easy task because of the expressions for $\overline{\mathrm{T}}$ above, which results in the graph depicted in Figure 7a. This graph has a single undominated communication class $\mathcal{X}^1_{\mathrm{m},1} = \{\mathsf{F}\}$, which is clearly regular. Since $\underline{\mathrm{T}}\mathbb{I}_{\mathsf{F}}(\mathsf{A}) = \underline{\mathrm{T}}\mathbb{I}_{\mathsf{F}}(\mathsf{B}) = \underline{\mathrm{T}}\mathbb{I}_{\mathsf{F}}(\mathsf{W}) = 0$, it follows that $\mathcal{X}^1_{\mathrm{lr}} = \emptyset$ and $\mathcal{X}^1_{\mathrm{f}} = \{\mathsf{A}, \mathsf{B}, \mathsf{W}\}$.

At this point, there is a distinction between our condition in Theorem 12 and the sufficient one of De Bock et al. [2025, Theorem 4.1]. We first look at the required steps for checking the sufficient one. On that approach, the next step is to determine the upper transition operator $\overline{\mathrm{T}}_{\{\mathsf{A},\mathsf{B},\mathsf{W}\}}$. We leave it to the reader to check that $\overline{\mathrm{T}}_{\{\mathsf{A},\mathsf{B},\mathsf{W}\}}$ is the upper transition operator corresponding to the closed sets of transition mass functions

$$\begin{aligned} Q^{\mathsf{A}}_{\{\mathsf{A},\mathsf{B},\mathsf{W}\}} &= \{(1-p)\mathbb{I}_{\mathsf{B}} + p\mathbb{I}_{\mathsf{W}} : p \in [0, \overline{p}_{\mathsf{A}}]\}, \\ Q^{\mathsf{B}}_{\{\mathsf{A},\mathsf{B},\mathsf{W}\}} &= \{(1-p)\mathbb{I}_{\mathsf{A}} + p\mathbb{I}_{\mathsf{W}} : p \in [0, \overline{p}_{\mathsf{B}}]\}, \\ Q^{\mathsf{W}}_{\{\mathsf{A},\mathsf{B},\mathsf{W}\}} &= \{\mathbb{I}_{\mathsf{W}}\}, \end{aligned}$$

whose accessibility graph $\mathcal{G}(\overline{\mathrm{T}}_{\{\mathsf{A},\mathsf{B},\mathsf{W}\}})$ is depicted in Figure 7b. Here too, this accessibility graph has a single undominated communication class—now $\{\mathsf{W}\}$—which is regular but not lower reachable from the other states—now A and B. The sufficient condition in [De Bock et al., 2025, Theorem 4.1] now requires us to consider the upper accessibility graph of the upper transition operator $\overline{\mathrm{T}}_{\{\mathsf{A},\mathsf{B}\}}$, which corresponds to the sets $Q^{\mathsf{A}}_{\{\mathsf{A},\mathsf{B}\}} = \{\mathbb{I}_{\mathsf{B}}\}$ and $Q^{\mathsf{B}}_{\{\mathsf{A},\mathsf{B}\}} = \{\mathbb{I}_{\mathsf{A}}\}$. This



(a) $\mathcal{G}(\overline{\mathrm{T}}) = \mathcal{G}_{\mathcal{X}}(\overline{\mathrm{T}})$ (b) $\mathcal{G}(\overline{\mathrm{T}}_{\{\mathsf{A},\mathsf{B},\mathsf{W}\}})$

(c) $\mathcal{G}(\overline{\mathrm{T}}_{\{\mathsf{A},\mathsf{B}\}})$ (d) $\mathcal{G}_{\{\mathsf{A},\mathsf{B},\mathsf{W}\}}(\overline{\mathrm{T}})$

Figure 7: Directed graphs for Example 11.

graph $\mathcal{G}(\overline{\mathrm{T}}_{\{\mathsf{A},\mathsf{B}\}})$, depicted in Figure 7c, consists of a single communication class $\{\mathsf{A}, \mathsf{B}\}$ which is automatically maximal but is however not regular: every closed directed path in this class has an even length. The sufficient condition [De Bock et al., 2025, Theorem 4.1] is therefore not satisfied, so cannot be used to establish convergence.

The operator $\overline{\mathrm{T}}$ is convergent, however, and we can show this by verifying the condition in Theorem 12. For $A_1 = \mathcal{X}$, we know that $\mathcal{G}_{A_1}(\overline{\mathrm{T}}) = \mathcal{G}(\overline{\mathrm{T}})$, so we already know from before that the accessibility graph $\mathcal{G}_{A_1}(\overline{\mathrm{T}}) = \mathcal{G}(\overline{\mathrm{T}})$ depicted in Figure 7a has a single maximal communication class $\mathcal{X}^1_{\mathrm{m},1} = \{\mathsf{F}\}$ that is regular, and that $A_2 = \{\mathsf{A}, \mathsf{B}, \mathsf{W}\}$. For the next step, rather than use $\mathcal{G}(\overline{\mathrm{T}}_{A_2})$, as De Bock et al. [2025, Section 4.4] do, we now need to consider the graph $\mathcal{G}_{A_2}(\overline{\mathrm{T}})$ instead. As we know from the discussion after Lemma 10 this will be a graph that includes at least the directed edges of $\mathcal{G}(\overline{\mathrm{T}}_{A_2})$, but possibly more: $\mathcal{G}(\overline{\mathrm{T}}_{A_2}) \subseteq \mathcal{G}_{A_2}(\overline{\mathrm{T}})$. Since $\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A_2^c} + \mathbb{I}_y)(x) = \overline{\mathrm{T}}_{A_2}\mathbb{I}_y(x)$ for all $\alpha \in \mathbb{R}_{>0}$, $x \in \{\mathsf{A}, \mathsf{B}\}$ and $y \in A_2$, the directed edges that leave either A or B are the same in both graphs, so it suffices to check the edges that leave W. The edge from W to W cannot disappear. Since $\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A_2^c} + \mathbb{I}_{\mathsf{B}})(\mathsf{W}) = 0$ for all $\alpha \in \mathbb{R}_{>0}$, no edge from W to B is added. Since

$$\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A_2^c} + \mathbb{I}_{\mathsf{A}})(\mathsf{W}) = \max\{p - \alpha o(p) : p \in [0, \overline{p}_{\mathsf{W}}]\} > 0$$

for all $\alpha \in \mathbb{R}_{>0}$, we do have to add an edge from W to A though. Consequently, and as can be seen in Figure 7d, the only addition to $\mathcal{G}(\overline{\mathrm{T}}_{A_2})$ is an extra edge from W to A. Crucially, the resulting graph $\mathcal{G}_{A_2}(\overline{\mathrm{T}})$ has a single communication class $\mathcal{X}^2_{\mathrm{m},1} = \{\mathsf{A}, \mathsf{B}, \mathsf{W}\}$, which is automatically maximal and furthermore easily seen to be regular. Consequently, $\mathcal{X}^2_{\mathrm{f}} = \emptyset$ and $d = 2$. Since $\mathcal{X}^1_{\mathrm{m},1}$ is A_1-regular and $\mathcal{X}^2_{\mathrm{m},1}$ is A_2-regular, we conclude from Theorem 12 that $\overline{\mathrm{T}}$ is convergent.

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 to the development of the ideas; the paper was written by the first two authors.

Acknowledgements

A lot of the ideas in this contribution trace back to Hanne Asselman's [2024] Master's thesis, 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 two additional research retreats. The authors thank Hanne and Natan for their valuable input, as well as for their blessing to continue this line of research without them.

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.

References

Marianne Akian and Stéphane Gaubert. Spectral theorem for convex monotone homogeneous maps, and ergodic control. *Nonlinear Analysis: Theory, Methods & Applications*, 52(2):637–679, 2003. doi: 10.1016/S0362-546X(02)00170-0.

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`.

Richard A. Brualdi and Herbert J. Ryser. *Combinatorial Matrix Theory*. Cambridge University Press, 1991. doi: 10.1017/CBO9781107325708.

Jasper De Bock and Natan T'Joens. Average behaviour of imprecise Markov chains: A single pointwise ergodic theorem for six different models. In *Proceedings of the Twelfth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA'21)*, volume 147 of *Proceedings of Machine Learning Research*, pages 90–99. PMLR, 2021.

Jasper De Bock, Alexander Erreygers, and Thomas Krak. Sum-product laws and efficient algorithms for imprecise Markov chains. In *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, volume 161 of *Proceedings of Machine Learning Research*, pages 1476–1485. PMLR, 2021.

Jasper De Bock, Alexander Erreygers, and Floris Persiau. A convenient characterisation of convergent upper transition operators. In *Proceedings of the Fourteenth International Symposium on Imprecise Probabilities: Theories and Applications*, volume 290 of *Proceedings of Machine Learning Research*, pages 115–125. PMLR, 2025.

Gert De Cooman and Filip Hermans. Imprecise probability trees: Bridging two theories of imprecise probability. *Artificial Intelligence*, 172(11):1400–1427, 2008. doi: 10.1016/j.artint.2008.03.001.

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. doi: 10.1017/S0269964809990039.

Gert De Cooman, Jasper De Bock, and Stavros Lopatatzidis. Imprecise stochastic processes in discrete time: global models, imprecise Markov chains, and ergodic theorems. *International Journal of Approximate Reasoning*, 76:18–46, 2016. doi: 10.1016/j.ijar.2016.04.009.

Robert Denk, Michael Kupper, and Max Nendel. Kolmogorov-type and general extension results for nonlinear expectations. *Banach Journal of Mathematical Analysis*, 12(3):515–540, 2018. doi: 10.1215/17358787-2017-0024.

Filip Hermans and Gert De Cooman. Characterisation of ergodic upper transition operators. *International Journal of Approximate Reasoning*, 53(4):573–583, 2012. doi: 10.1016/j.ijar.2011.12.008.

Filip Hermans and Damjan Škulj. *Stochastic Processes*, chapter 11 of Introduction to Imprecise Probabilities. Wiley, 2014.

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

Igor O. Kozine and Lev V. Utkin. Interval-valued finite Markov chains. *Reliable Computing*, 8(2):97–113, 2002. doi: 10.1023/A:1014745904458.

Thomas Krak. Hitting times for continuous-time imprecise-Markov chains. In *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence*, volume 180 of *Proceedings of Machine Learning Research*, pages 1031–1040. PMLR, 2022.

Bas Lemmens and Michael Scheutzow. On the dynamics of sup-norm non-expansive maps. *Ergodic Theory and Dynamical Systems*, 25(3):861–871, 2005. doi: 10.1017/s0143385704000665.

James R. Norris. *Markov Chains*. Cambridge University Press, 1997. doi: 10.1017/CBO9780511810633.

Marco Sangalli, Rick Reubsaet, Erik Quaeghebeur, and Thomas Krak. Computing lower and upper hitting probabilities for imprecise Markov chains. *International*

Journal of Approximate Reasoning, 194:109671, 2026. doi: 10.1016/j.ijar.2026.109671.

Robert Sine. A nonlinear Perron–Frobenius Theorem. *Proceedings of the American Mathematical Society*, 109(2): 331–336, 1990. doi: 10.2307/2047992.

Natan T'Joens and Jasper De Bock. Average behaviour in discrete-time imprecise Markov chains: A study of weak ergodicity. *International Journal of Approximate Reasoning*, 132:181–205, 2021. doi: 10.1016/j.ijar.2021.03.001.

Damjan Škulj. Discrete time Markov chains with interval probabilities. *International Journal of Approximate Reasoning*, 50(8):1314–1329, 2009. doi: 10.1016/j.ijar.2009.06.007.

Characterising the Convergence of Imprecise Markov Chains (Supplementary Material)

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

[1]Foundations Lab for imprecise probabilities, Ghent University, Ghent, Belgium

This appendix contains our proofs for the results in the main text, as well as some additional technical lemmas. We start off in Appendix A with proving everything related to lower reachability, continue in Appendix B with giving proofs for the results regarding accessibility, and end in Appendix C with our—rather long—proof for Theorem 12.

A EXTRA RESULTS AND RELEGATED PROOFS REGARDING LOWER REACHABILITY

First, we prove Lemma 3.

Proof of Lemma 3. That $B \subseteq B_n$ follows immediately from the assumption on B in the statement.

Observe that $(B_n)_{n\in\mathbb{Z}_{\geq 0}}$ is a non-decreasing sequence by the second equality in the statement. Since B^{c} is finite, this implies that there must be some smallest $k \in \mathbb{Z}_{\geq 0}$ such that $B_{k+1} = B_k$—and then $B_k = B_n$ for all $n \geq k$—and that $k \leq |B^{\mathrm{c}}|$.

That B_k is the set of states from which B is lower reachable follows immediately from the definition of lower reachability and the definition of B_n.

Hence, it remains for us to verify that for all $n \in \mathbb{Z}_{\geq 0}$,

$$B_{n+1} = \{x \in \mathcal{X} \colon \underline{\mathrm{T}}\mathbb{I}_{B_n}(x) > 0\} = B_n \cup \{x \in \mathcal{X} \setminus B_n \colon \underline{\mathrm{T}}\mathbb{I}_{B_n}(x) > 0\}.$$

To this end, note that for all $n \in \mathbb{Z}_{\geq 0}$, by (T5) and the definition of B_n there is some $\beta_n \in]0, 1]$ such that $\beta_n \mathbb{I}_{B_n} \leq \underline{\mathrm{T}}^n \mathbb{I}_B \leq \mathbb{I}_{B_n}$. Consequently, for all $n \in \mathbb{Z}_{\geq 0}$,

$$\beta_n \underline{\mathrm{T}}\mathbb{I}_{B_n} = \underline{\mathrm{T}}\beta_n \mathbb{I}_{B_n} \leq \underline{\mathrm{T}}\underline{\mathrm{T}}^n \mathbb{I}_B = \underline{\mathrm{T}}^{n+1}\mathbb{I}_B \leq \underline{\mathrm{T}}\mathbb{I}_{B_n}.$$

From this, we infer immediately that

$$B_{n+1} = \{x \in \mathcal{X} \colon \underline{\mathrm{T}}\mathbb{I}_{B_n}(x) > 0\} \quad \text{for all } n \in \mathbb{Z}_{\geq 0},$$

which is the first equality. For the second equality, observe that if $x \in B_n$, then

$$\underline{\mathrm{T}}^{n+1}\mathbb{I}_B(x) = \underline{\mathrm{T}}^n \underline{\mathrm{T}}\mathbb{I}_B(x) \geq \underline{\mathrm{T}}^n \beta_1 \mathbb{I}_{B_1}(x) = \beta_1 \underline{\mathrm{T}}^n \mathbb{I}_{B_1}(x) \geq \beta_1 \underline{\mathrm{T}}^n \mathbb{I}_B(x) > 0.$$ □

The next result we need to prove is Lemma 4.

Proof of Lemma 4. Assume towards contradiction that there are some $n \in \mathbb{N}$ and $x \in C$ such that $\underline{\mathrm{T}}^n \mathbb{I}_C(x) < 1$—the alternate case $\underline{\mathrm{T}}^n \mathbb{I}_C(x) > 1$ is of course impossible due to (T5). From this and (T8), it follows that $\overline{\mathrm{T}}^n \mathbb{I}_{C^{\mathrm{c}}}(x) = 1 - \underline{\mathrm{T}}^n \mathbb{I}_C(x) > 0$. Since

$$\overline{\mathrm{T}}^n \mathbb{I}_{C^{\mathrm{c}}}(x) = \overline{\mathrm{T}}^n \left(\sum_{y \in C^{\mathrm{c}}} \mathbb{I}_y \right)(x) \leq \sum_{y \in C^{\mathrm{c}}} \overline{\mathrm{T}}^n \mathbb{I}_y(x)$$

by (T2), we infer from this that there must be some $y \in C^{\mathrm{c}}$ such that $\overline{\mathrm{T}}^n \mathbb{I}_y(x) > 0$, or equivalently, $x \overset{n}{\rightsquigarrow} y$. Since $x \in C$, $y \in C^{\mathrm{c}}$ and C is closed, this is the contradiction we were after. □

B EXTRA RESULTS AND RELEGATED PROOFS REGARDING ACCESSIBILITY

First, we prove the equivalent characterisations of filtering classes in Lemma 6.

Proof of Lemma 6. It is not difficult to see that (ii) and (iii) imply (i): take $n = 1$ and use (T8) and (T5). Similarly, (iv) implies (ii) and (iii). Hence, it remains for us to verify that (i) implies (iv).

We'll prove by induction on n that $\beta_A \coloneqq \min\{\underline{\mathrm{T}}\mathbb{I}_{A^c}(x) \colon x \in A^c\} > 0$ satisfies the condition in (iv). For $n = 0$ and $n = 1$, since A is a filtering class our definition of β_A ensures that $\mathbb{I}_{A^c} \geq \underline{\mathrm{T}}^n\mathbb{I}_{A^c} \geq \beta_A^n\mathbb{I}_{A^c}$. For the inductive step, we assume that $\mathbb{I}_{A^c} \geq \underline{\mathrm{T}}^k\mathbb{I}_{A^c} \geq \beta_A^k\mathbb{I}_{A^c}$ for all $k \leq n$, and set out to verify this inequality for $n + 1$. Since $\underline{\mathrm{T}}^{n+1}\mathbb{I}_{A^c} = \underline{\mathrm{T}}(\underline{\mathrm{T}}^n\mathbb{I}_{A^c})$, it follows from the induction hypothesis for $k = 1$ and $k = n$, (T6) and (T3) that

$$\beta_A^{n+1}\mathbb{I}_{A^c} \leq \beta_A^n\underline{\mathrm{T}}\mathbb{I}_{A^c} = \underline{\mathrm{T}}\beta_A^n\mathbb{I}_{A^c} \leq \underline{\mathrm{T}}^{n+1}\mathbb{I}_{A^c} \leq \underline{\mathrm{T}}\mathbb{I}_{A^c} \leq \mathbb{I}_{A^c}. \qquad \square$$

Proof of Lemma 7. Since we know from (4) that $\overline{\mathrm{T}}\mathbb{I}_A(a) = 1$, it follows that for all $\alpha \in \mathbb{R}_{>0}$,

$$\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_A)(a) = \overline{\mathrm{T}}(-\alpha + (1+\alpha)\mathbb{I}_A)(a) = -\alpha + (1+\alpha)\overline{\mathrm{T}}\mathbb{I}_A(a) = 1.$$

From this, it follows that for all $\alpha \in \mathbb{R}_{>0}$,

$$1 = \overline{\mathrm{T}}(-\alpha|A|\mathbb{I}_{A^c} + \mathbb{I}_A)(a) = \overline{\mathrm{T}}\left(\sum_{x \in A} -\alpha\mathbb{I}_{A^c} + \mathbb{I}_x\right)(a) \leq \sum_{x \in A}\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_x)(a). \tag{6}$$

Hence, for all $\alpha \in \mathbb{R}_{>0}$, the set $C_\alpha \coloneqq \{x \in A \colon \overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_x)(a) > 0\}$ is non-empty, and it follows from (T6) that $\alpha \mapsto C_\alpha$ decreases for increasing α. Since $\bigcap_{\alpha \in \mathbb{R}_{>0}} C_\alpha$ cannot be empty—otherwise we'd reach a contradiction with (6)—the statement follows. $\square$

In our proof for Lemma 8, we'll rely on the following intermediary result.

Lemma 13. *For any filtering class A,*

$$\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) \geq 0 \quad \textit{for all } n \in \mathbb{N}, \alpha \in \mathbb{R}_{>0}, x, y \in A.$$

Proof. This follows immediately from (T6) and Lemma 6:

$$\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) \geq \overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c})(x) = -\alpha\underline{\mathrm{T}}\mathbb{I}_{A^c}(x) = 0. \qquad \square$$

Proof of Lemma 8. Our proof will be one by induction on n. The base case $n = 1$ holds trivially. For the inductive step, we assume that the result holds for some $n \in \mathbb{N}$ and consider any $x, y \in \mathcal{X}$. Now suppose there is some sequence $x = z_0, \ldots, z_{n+1} = y$ in $\mathcal{X}$ that satisfies the condition in the statement. Then by the induction hypothesis, for all $\alpha, \beta \in \mathbb{R}_{>0}$, $\overline{\mathrm{T}}^n(-\beta\mathbb{I}_{A^c} + \mathbb{I}_{z_n})(x) > 0$ and $\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(z_n) > 0$. Since furthermore $\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(z) \geq -\alpha$ for all $z \in A^c$ and $\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(z) \geq 0$ for all $z \in A$ by Lemma 13, it follows that, with $\gamma \coloneqq \overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(z_n) > 0$,

$$\overline{\mathrm{T}}^{n+1}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) = \overline{\mathrm{T}}^n\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) \geq \overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \gamma\mathbb{I}_{z_n})(x) = \gamma\overline{\mathrm{T}}^n\left(-\frac{\alpha}{\gamma}\mathbb{I}_{A^c} + \mathbb{I}_{z_n}\right)(x).$$

From this and the induction hypothesis, we infer that for all $\alpha \in \mathbb{R}_{>0}$, $\overline{\mathrm{T}}^{n+1}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) > 0$.

For the converse implication, we suppose that there is no such sequence from x to y and show that then there is some $\alpha \in \mathbb{R}_{>0}$ for which $\overline{\mathrm{T}}^{n+1}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) \leq 0$. Since there is no such sequence, there is no $z \in A$ such that for all $\alpha \in \mathbb{R}_{>0}$, $\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_z)(x) > 0$ and—using the induction hypothesis—$\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(z) > 0$. Since A is a filtering class, we know from Lemma 6 that there is some $\beta_A \in]0, 1]$ for which $\underline{\mathrm{T}}^n\mathbb{I}_{A^c} \geq \beta_A^n\mathbb{I}_{A^c}$. Consequently, for all $\alpha \in \mathbb{R}_{>0}$ and $z \in A^c$,

$$\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(z) \leq \alpha\overline{\mathrm{T}}^n(-\mathbb{I}_{A^c})(z) + \overline{\mathrm{T}}\mathbb{I}_y(z) = -\alpha\underline{\mathrm{T}}^n\mathbb{I}_{A^c}(z) + \overline{\mathrm{T}}\mathbb{I}_y(z) \leq -\alpha\beta_A^n + 1.$$

So for all $\alpha \in \mathbb{R}_{>0}$ for which $\alpha\beta_A^n - 1 > 0$, and with $\mathcal{Z}_\alpha \coloneqq \{z \in A^c : \overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(z) > 0\}$,

$$\begin{aligned}\overline{\mathrm{T}}^{n+1}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) &= \overline{\mathrm{T}}\left(\sum_{z\in\mathcal{X}} \overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(z)\mathbb{I}_z\right)(x)\\ &\leq \overline{\mathrm{T}}(-(\alpha\beta_A^n - 1)\mathbb{I}_{A^c} + \mathbb{I}_{\mathcal{Z}_\alpha})(x)\\ &\leq \sum_{z\in\mathcal{Z}_\alpha} \overline{\mathrm{T}}\left(-\frac{\alpha\beta_A^n - 1}{|\mathcal{X}|}\mathbb{I}_{A^c} + \mathbb{I}_z\right)(x).\end{aligned}$$

Note that by (T6), the map $\alpha \mapsto \mathcal{Z}_\alpha$ decreases with increasing α. If it decreases to the empty set, we're done. If it doesn't, then let $\mathcal{Z}^* \coloneqq \bigcap_{\alpha\in\mathbb{R}_{>0}} \mathcal{Z}_\alpha$. Then there must be some $\alpha \in \mathbb{R}_{>0}$ such that $\mathcal{Z}_\alpha = \mathcal{Z}^*$ and $\overline{\mathrm{T}}(-\frac{\alpha\beta_A^n - 1}{|\mathcal{X}|}\mathbb{I}_{A^c} + \mathbb{I}_z)(x) \leq 0$ for all $z \in \mathcal{Z}^*$, otherwise we contradict our initial assumption. For this α, we infer from the previous inequality that $\overline{\mathrm{T}}^{n+1}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) \leq 0$, which is what we needed to prove. □

Proof of Proposition 9. Property (A1) follows from Lemma 8 and (A2) follows from Lemma 7. □

The last result related to accessibility in the main text that we still need to prove is Lemma 10.

Proof of Lemma 10. It clearly suffices to prove that for all $x, y \in A$,

$$\overline{\mathrm{T}}_A\mathbb{I}_y(x) > 0 \Leftrightarrow (\forall\alpha \in \mathbb{R}_{>0})\ \overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) > 0.$$

So fix any $x, y \in A$. Since $\overline{\mathrm{T}}$ is finitely generated, there is some *finite* set $Q^x \subseteq \Sigma_{\mathcal{X}}$ such that

$$\overline{\mathrm{T}}f(x) = \max\{\mathrm{E}_q(f) : q \in Q^x\} \quad \text{for all } f \in \mathbb{R}^{\mathcal{X}}. \tag{7}$$

Consider any such set Q^x.

If $\overline{\mathrm{T}}_A\mathbb{I}_y(x) > 0$, then since Q^x is closed because it is finite, it follows from the definition of $\overline{\mathrm{T}}_A$ that there is some probability mass function q in Q^x such that $\mathrm{E}_q(\mathbb{I}_{A^c}) = 0$ and $\mathrm{E}_q(\mathbb{I}_y) > 0$. Consequently, for all $\alpha \in \mathbb{R}_{>0}$,

$$\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) \geq \mathrm{E}_q(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y) = -\alpha\mathrm{E}_q(\mathbb{I}_{A^c}) + \mathrm{E}_q(\mathbb{I}_y) > 0.$$

For the implication to the right, assume that $\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) > 0$ for all $\alpha \in \mathbb{R}_{>0}$. For all $n \in \mathbb{N}$, it then follows from Equation (7) that there is some probability mass function $q_n \in Q^x$ such that $\mathrm{E}_{q_n}(-n\mathbb{I}_{A^c} + \mathbb{I}_y) > 0$. Since Q^x is finite, there is at least one mass function $q \in Q^x$ that appears infinitely often in the sequence $(q_n)_{n\in\mathbb{N}}$. It must be that $q(y) > 0$, otherwise $\mathrm{E}_q(-n\mathbb{I}_{A^c} + \mathbb{I}_y) \leq 0$ for all $n \in \mathbb{N}$. We now claim that $\mathrm{E}_q(\mathbb{I}_{A^c}) = 0$, from which it follows that indeed

$$\overline{\mathrm{T}}_A\mathbb{I}_y(x) \geq \mathrm{E}_{q\downarrow_A}(\mathbb{I}_y) = q(y) > 0.$$

To verify this claim, we assume towards contradiction that $c \coloneqq \mathrm{E}_q(\mathbb{I}_{A^c}) > 0$. Then for all $\alpha \in \,]1/c, +\infty[$,

$$\mathrm{E}_q(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y) = -\alpha\mathrm{E}_q(\mathbb{I}_{A^c}) + \mathrm{E}_q(\mathbb{I}_y) = -\alpha c + q(y) < -1 + q(y) < 0.$$

On the other hand, for all $\alpha \in \,]1/c, +\infty[$, and with n_α the smallest natural number $n \geq \alpha$ such that $q_n = q$,

$$\mathrm{E}_q(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y) \geq \mathrm{E}_q(-n_\alpha\mathbb{I}_{A^c} + \mathbb{I}_y) = \mathrm{E}_{q_n}(-n_\alpha\mathbb{I}_{A^c} + \mathbb{I}_y) > 0.$$

We've arrived at the contradiction which verifies our claim, and this finalises our proof. □

Finally, we establish two technical results which will come in handy further down the line.

Lemma 14. *Consider some filtering class A and some closed class $C \subseteq A$. Then for all $x \in C$ and $n \in \mathbb{N}$, $\overline{\mathrm{T}}^n\mathbb{I}_C(x) > 0$.*

Proof. Since C is closed and $\cdot \xrightarrow{A,\cdot} \cdot$ is an accessibility relation [Proposition 9], it follows from repeated application of (A2) and (A1) that there is some $z \in C$ such that $x \xrightarrow{A,n} z$. It follows from (T6) and this that

$$\overline{\mathrm{T}}^n\mathbb{I}_C(x) \geq \overline{\mathrm{T}}^n(-\mathbb{I}_{A^c} + \mathbb{I}_z)(x) > 0.$$

□

Lemma 15. *Consider some filtering class A and some closed class $C \subseteq A$. Then for all $x \in C$ and $n \in \mathbb{N}$, there is some $\alpha \in \mathbb{R}_{>0}$ such that*

$$\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_{C^c})(x) \leq 0.$$

Proof. If $C^c = \emptyset$, this follows immediately from (T1); hence, we may assume without loss of generality that $C^c \neq \emptyset$, or equivalently, that $|C^c| > 0$. Assume towards contradiction that there are some $x \in C$ and $n \in \mathbb{N}$ such that

$$0 < \overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_{C^c}) \geq x \quad \text{for all } \alpha \in \mathbb{R}_{>0}.$$

From this and (T2), it follows that

$$0 < \overline{\mathrm{T}}^n(-|C^c|\alpha\mathbb{I}_{A^c} + \mathbb{I}_{C^c}) = \overline{\mathrm{T}}^n\left(\sum_{y\in C^c} -\alpha\mathbb{I}_{A^c} + \mathbb{I}_y\right)(x) \leq \sum_{y\in C^c} \overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) \quad \text{for all } \alpha \in \mathbb{R}_{>0}.$$

From this, it follows that the set $\mathcal{Y}_\alpha \coloneqq \{y \in C^c \colon \overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) > 0\}$ is non-empty for all $\alpha \in \mathbb{R}_{>0}$. Since the map $\alpha \mapsto \mathcal{Y}_\alpha$ is furthermore decreasing for increasing α due to (T6), we conclude that $\bigcap_{\alpha\in\mathbb{R}_{>0}} \mathcal{Y}_\alpha$ is non-empty. But this tells us that there is some $y \in C^c$ such that $x \xrightarrow{A,n} y$, which contradicts the assumption in the statement that C is closed. □

C PROOF FOR THEOREM 12

The remainder of this appendix will be devoted to the proof of our main result Theorem 12. Before doing so, we still need to prove Lemma 11 though.

Proof of Lemma 11. For $b \in A^c$, it follows from (T6) and Lemma 6 that for all $n \in \mathbb{N}$,

$$\underline{\mathrm{T}}^n\mathbb{I}_B(b) \geq \underline{\mathrm{T}}^n\mathbb{I}_{A^c}(b) > 0.$$

The alternative case $b \in C$ is a bit more involved. Assume towards contradiction that there are some $b \in C$ and $n \in \mathbb{N}$ for which $\underline{\mathrm{T}}^n\mathbb{I}_B(b) = 0$. Then $\overline{\mathrm{T}}^n\mathbb{I}_{B^c}(b) = 1$ by (T8); consequently, for all $\alpha \in \mathbb{R}_{>0}$,

$$\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_{B^c})(b) \geq \overline{\mathrm{T}}^n(-\alpha\mathbb{I}_B + \mathbb{I}_{B^c})(b) = \overline{\mathrm{T}}^n(-\alpha + (1+\alpha)\mathbb{I}_{B^c})(b) = -\alpha + (1+\alpha)\overline{\mathrm{T}}^n\mathbb{I}_{B^c}(b) = 1,$$

where we used (T6) for the inequality and (T7) for the penultimate equality. From this, it follows that for all $k \in \mathbb{N}$,

$$1 \leq \overline{\mathrm{T}}^n(-k|B^c|\mathbb{I}_{A^c} + \mathbb{I}_{B^c})(b) \leq \sum_{z\in B^c} \overline{\mathrm{T}}^n(-k\mathbb{I}_{A^c} + \mathbb{I}_z)(b),$$

so there is at least one $z_k \in B^c = A \cap C^c$ for which $\overline{\mathrm{T}}(-k\mathbb{I}_{A^c} + \mathbb{I}_z)(b) > 0$. Because B^c is finite, there is at least one $z^\star \in B^c = A \cap C^c$ that appears infinitely often in the sequence $(z_k)_{k\in\mathbb{N}}$. Then for all $\alpha \in \mathbb{R}_{>0}$, and with k_α the smallest natural number $n \geq \alpha$ such that $z_{k_\alpha} = z^\star$,

$$\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_{z^\star})(b) \geq \overline{\mathrm{T}}^n(-k_\alpha\mathbb{I}_{A^c} + \mathbb{I}_{z^\star})(b) = \overline{\mathrm{T}}^n(-k_\alpha\mathbb{I}_{A^c} + \mathbb{I}_{z_{k_\alpha}})(b) > 0.$$

This means that $b \xrightarrow{A,n} z^\star$, which is a contradiction because $b \in C$, $z^\star \in C^c$ and C is closed. □

C.1 SUFFICIENCY

Let us establish first that the condition in Theorem 12 is sufficient for convergence.

Proposition 16. *If for all $k \in \{1, \ldots, d\}$, the A_k-maximal communication classes $\mathcal{X}^k_{\mathrm{m},1}, \ldots, \mathcal{X}^k_{\mathrm{m},M_k}$ are A_k-regular, then $\overline{\mathrm{T}}$ is convergent.*

We start our proof for Proposition 16 with an intermediary technical result, which will also turn out to be useful in our proof for the necessity further on.

Lemma 17. *Consider some filtering class A and some A-closed class $C \subseteq A$. Then for all $x, y \in C$, $f \in \mathbb{R}^{\mathcal{X}}$ and $n \in \mathbb{N}$, there is some $\alpha \in \mathbb{R}_{>0}$ such that*

$$\begin{aligned}\overline{\mathrm{T}}^n f(x) &\geq \min f{\downarrow_C} + (f(y) - \min f{\downarrow_C})\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x)\\ &\geq \min f{\downarrow_C}.\end{aligned}$$

Proof. Since $\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) \geq 0$ for all $\alpha \in \mathbb{R}_{>0}$ [Lemma 13], the second inequality follows immediately from the first one.

To prove the first inequality, let $c \coloneqq f(y) - \min f{\downarrow_C}$. To deal with the case $c = 0$, we let $\delta \coloneqq \min f{\downarrow_C} - \min f$. For $\delta = 0$, the inequality follows from (T1). For $\delta > 0$,

$$\overline{\mathrm{T}}^n f(x) - \min f = \overline{\mathrm{T}}^n(f - \min f)(x) \geq \overline{\mathrm{T}}^n(\delta\mathbb{I}_C)(x) = \delta\overline{\mathrm{T}}^n\mathbb{I}_C(x) > 0,$$

where the last inequality follows from Lemma 14. Since $c = 0$, this inequality implies the one in the statement.

For the case $c > 0$, we choose $\beta \in \mathbb{R}_{>0}$ such that $f \geq \min f{\downarrow_C} + c\mathbb{I}_y - \beta\mathbb{I}_{C^c}$. By Lemma 15, there is some $\gamma \in \mathbb{R}_{>0}$ such that $\overline{\mathrm{T}}^n(-\gamma\mathbb{I}_{A^c} + \mathbb{I}_{C^c})(x) \leq 0$. Then with $\alpha \coloneqq \gamma\beta/c > 0$,

$$f \geq \min f{\downarrow_C} + c\mathbb{I}_y - \beta\mathbb{I}_{C^c} = \min f{\downarrow_C} - \gamma\beta\mathbb{I}_{A^c} + c\mathbb{I}_y - \beta(-\gamma\mathbb{I}_{A^c} + \mathbb{I}_{C^c}) = \min f{\downarrow_C} + c(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y) - \beta(-\gamma\mathbb{I}_{A^c} + \mathbb{I}_{C^c}),$$

and therefore,

$$\begin{aligned}\overline{\mathrm{T}}^n f(x) &\geq \overline{\mathrm{T}}^n(\min f{\downarrow_C} + c(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y) - \beta(-\gamma\mathbb{I}_{A^c} + \mathbb{I}_{C^c}))(x)\\ &= \min f{\downarrow_C} + \overline{\mathrm{T}}^n(c(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y) - \beta(-\gamma\mathbb{I}_{A^c} + \mathbb{I}_{C^c}))(x)\\ &\geq \min f{\downarrow_C} + \overline{\mathrm{T}}^n(c(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y))(x) - \overline{\mathrm{T}}^n(\beta(-\gamma\mathbb{I}_{A^c} + \mathbb{I}_{C^c}))(x)\\ &= \min f{\downarrow_C} + c\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) - \beta\overline{\mathrm{T}}^n(-\gamma\mathbb{I}_{A^c} + \mathbb{I}_{C^c})(x)\\ &\geq \min f{\downarrow_C} + c\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x),\end{aligned}$$

where the last inequality follows from our choice of γ. □

For the following result, it'll help to recall Definition 4.1 from [De Bock et al., 2025].

Definition 18. Given some class C, $\overline{\mathrm{T}}$ is *convergent (ergodic) on C* if for all $f \in \mathbb{R}^{\mathcal{X}}$, the sequence $((\overline{\mathrm{T}}^n f){\downarrow_C})_{n\in\mathbb{N}}$ converges (to a constant function).

Lemma 19. *Consider some filtering class A and some maximal communication class $C \subseteq A$. If C is A-regular, then $\overline{\mathrm{T}}$ is ergodic on C.*

Proof. Fix any $f \in \mathbb{R}^{\mathcal{X}}$, and consider its limit set $\Omega_f = \{\omega_1, \ldots, \omega_{p_f}\}$. Let

$$c_{\min} \coloneqq \min\{\omega_r(x) \colon r \in \{1, \ldots, p_f\}, x \in C\} \quad \text{and} \quad c_{\max} \coloneqq \max\{\omega_r(x) \colon r \in \{1, \ldots, p_f\}, x \in C\}.$$

To prove the statement, it suffices to show that $c_{\min} = c_{\max}$.

Assume towards contradiction that $c_{\min} < c_{\max}$. Then there are some $k, \ell \in \{1, \ldots, p_f\}$ and $x, y \in C$ such that $c_{\min} = \omega_k(x) < \omega_\ell(y) = c_{\max}$. Since ω_k and ω_ℓ are part of a p_f-cycle, there is some $m \in \mathbb{N}$ such that $\omega_k(x) = \overline{\mathrm{T}}^{m+np_f}\omega_\ell(x)$ for all $n \in \mathbb{N}$. Since C is A-regular and $x, y \in C$, we furthermore know that there is some $\tilde{n} \in \mathbb{N}$ such that x is A-accessible from y in $m + \tilde{n}p_f$ steps, which implies that $c_\alpha \coloneqq \overline{\mathrm{T}}^{m+\tilde{n}p_f}(-\alpha\mathbb{I}_{A^c} + \mathbb{I}_y)(x) > 0$ for all $\alpha \in \mathbb{R}_{>0}$. Finally, it follows from Lemma 17 that there is some $\tilde{\alpha} \in \mathbb{R}_{>0}$ such that

$$\begin{aligned}\overline{\mathrm{T}}^{m+\tilde{n}p_f}\omega_\ell(x) &\geq \min \omega_\ell{\downarrow_C} + (\omega_\ell(y) - \min \omega_\ell{\downarrow_C})\overline{\mathrm{T}}^{m+\tilde{n}p_f}(-\tilde{\alpha}\mathbb{I}_{A^c} + \mathbb{I}_y)(x)\\ &\geq c_{\min} + (c_{\max} - c_{\min})c_{\tilde{\alpha}} > c_{\min}.\end{aligned}$$

This contradicts the fact that $c_{\min} = \omega_k(x) = \overline{\mathrm{T}}^{m+\tilde{n}p_f}\omega_\ell(x)$. □

Lemma 20. *With $D \coloneqq \bigcup_{k=1}^{d} \mathcal{X}_{\mathrm{lr}}^{k}$, the map $\overline{\mathrm{S}}\colon \mathbb{R}^{\mathcal{X}} \to \mathbb{R}^{\mathcal{X}}$ defined by*

$$\overline{\mathrm{S}}g(x) = \begin{cases} \overline{\mathrm{T}}g(x) & \text{if } x \in D \\ g(x) & \text{otherwise} \end{cases} \quad \text{for all } g \in \mathbb{R}^{\mathcal{X}}, x \in \mathcal{X}$$

is an upper transition operator. Furthermore, $D^{\mathrm{c}} = \bigcup_{k=1}^{k} \mathcal{X}_{\mathrm{m}}^{k}$ is a $\mathcal{X}$-closed class for $\overline{\mathrm{S}}$ that is lower reachable from every $x \in \mathcal{X}$ for $\overline{\mathrm{S}}$.

Proof. That $\overline{\mathrm{S}}$ is an upper transition operator follows immediately from the definition of $\overline{\mathrm{S}}$ and the fact that $\overline{\mathrm{T}}$ is an upper transition operator. Similarly, that D^{c} is closed for $\overline{\mathrm{S}}$ follows immediately from the definition of $\overline{\mathrm{S}}$. Hence, it remains for us to show that D^{c} is lower reachable (for $\overline{\mathrm{S}}$) from every $x \in \mathcal{X}$.

Applying Lemma 3 with $B^1 = \mathcal{X}_{\mathrm{m}}^{1}$ [which is allowed due to Lemma 11 with $A = \mathcal{X}$], we get the non-decreasing sequence $(B_n^1)_{n\in\mathbb{Z}_{\geq 0}}$ which is constant from k_1 onward with $B_{k_1}^1 = \mathcal{X}_{\mathrm{m}}^{1} \cup \mathcal{X}_{\mathrm{lr}}^{1} = A_2^{\mathrm{c}}$. Observe now that for all $n \in \mathbb{Z}_{\geq 0}$,

$$B_{n+1}^1 = B_n^1 \cup \{x \in \mathcal{X} \setminus B_n^1 \colon \underline{\mathrm{T}}\mathbb{I}_{B_n^1}(x) > 0\} = B_n^1 \cup \{x \in \mathcal{X} \setminus B_n^1 \colon \underline{\mathrm{S}}\mathbb{I}_{B_n^1}(x) > 0\},$$

where the equality follows from the definition of $\overline{\mathrm{S}}$. Since $\mathcal{X}_{\mathrm{m}}^{1}$ is trivially $\mathcal{X}$-closed for $\overline{\mathrm{S}}$, it follows from the equality above that $(B_n^1)_{n\in\mathbb{Z}_{\geq 0}}$ is also the sequence we get by applying Lemma 3 with $A = \mathcal{X}$ and $C = \mathcal{X}_{\mathrm{m}}^{1}$ for $\overline{\mathrm{S}}$. Consequently, for all $k \geq k_1$, $\beta_k^1 \coloneqq \min\{\underline{\mathrm{S}}^k\mathbb{I}_{\mathcal{X}_{\mathrm{m}}^{1}}(x) \colon x \in A_2^{\mathrm{c}}\} > 0$ and $\beta_k^1\mathbb{I}_{A_2^{\mathrm{c}}} \leq \underline{\mathrm{S}}^k\mathbb{I}_{\mathcal{X}_{\mathrm{m}}^{1}} \leq \mathbb{I}_{A_2^{\mathrm{c}}}$. In particular, for all $x \in \mathcal{X}_{\mathrm{m}}^{1} \cup \mathcal{X}_{\mathrm{lr}}^{1} = A_2^{\mathrm{c}}$ it follows from this, (T6) and (T3) that $\underline{\mathrm{S}}^{k_1}\mathbb{I}_{D^{\mathrm{c}}}(x) \geq \underline{\mathrm{S}}^{k_1}\mathbb{I}_{\mathcal{X}_{\mathrm{m}}^{1}}(x) > 0$, so D^{c} is lower reachable from x.

If $d = 1$, we're done. If $d > 1$, we repeat the same argument. Applying Lemma 3 with $B^2 = A_2^{\mathrm{c}} \cup \mathcal{X}_{\mathrm{m}}^{2}$ [which is allowed by Lemma 11 for $A = A_2$ and $C = \mathcal{X}_{\mathrm{m}}^{2}$], we get the non-decreasing sequence $(B_n^2)_{n\in\mathbb{Z}_{\geq 0}}$ which is constant from k_2 onward with $B_{k_2}^2 = A_2^{\mathrm{c}} \cup \mathcal{X}_{\mathrm{m}}^{2} \cup \mathcal{X}_{\mathrm{lr}}^{2} = A_3^{\mathrm{c}}$. Here too, it follows from the definition of $\overline{\mathrm{S}}$ that, for all $n \in \mathbb{Z}_{\geq 0}$,

$$B_{n+1}^2 = B_n^2 \cup \{x \in \mathcal{X} \setminus B_n^2 \colon \underline{\mathrm{T}}\mathbb{I}_{B_n^2}(x) > 0\} = B_n^2 \cup \{x \in \mathcal{X} \setminus B_n^2 \colon \underline{\mathrm{S}}\mathbb{I}_{B_n^2}(x) > 0\}.$$

Since $\mathcal{X}_{\mathrm{m}}^{2}$ is also A_2-closed for $\overline{\mathrm{S}}$, it follows from Lemma 11 and the equality above that $(B_n^2)_{n\in\mathbb{Z}_{\geq 0}}$ is also the sequence we get by applying Lemma 3 with $B = B^2 = A_2^{\mathrm{c}} \cup \mathcal{X}_{\mathrm{m}}^{2}$ for $\overline{\mathrm{S}}$. Consequently, for all $k \geq k_2$, $\beta_k^2 \coloneqq \min\{\underline{\mathrm{S}}^k\mathbb{I}_{A_2^{\mathrm{c}}\cup\mathcal{X}_{\mathrm{m}}^{2}}(x) \colon x \in A_3^{\mathrm{c}}\} > 0$ and $\beta_k^2\mathbb{I}_{A_3^{\mathrm{c}}} \leq \underline{\mathrm{S}}^k\mathbb{I}_{A_2^{\mathrm{c}}\cup\mathcal{X}_{\mathrm{m}}^{2}} \leq \mathbb{I}_{A_3^{\mathrm{c}}}$. Furthermore, it follows from (repeated application of) the definition of $\overline{\mathrm{S}}$ and (T5) that $\underline{\mathrm{S}}^k\mathbb{I}_{\mathcal{X}_{\mathrm{m}}^{1}\cup\mathcal{X}_{\mathrm{m}}^{2}} \geq \mathbb{I}_{\mathcal{X}_{\mathrm{m}}^{1}\cup\mathcal{X}_{\mathrm{m}}^{2}}$ for all $k \in \mathbb{N}$. For any $x \in \mathcal{X}_{\mathrm{m}}^{2} \cup \mathcal{X}_{\mathrm{lr}}^{2} = A_3^{\mathrm{c}} \setminus A_2^{\mathrm{c}}$, it follows from all of this that

$$\underline{\mathrm{S}}^{k_1+k_2}\mathbb{I}_{D^{\mathrm{c}}}(x) \geq \underline{\mathrm{S}}^{k_1+k_2}\mathbb{I}_{\mathcal{X}_{\mathrm{m}}^{1}\cup\mathcal{X}_{\mathrm{m}}^{2}}(x) \geq \underline{\mathrm{S}}^{k_2}(\mathbb{I}_{\mathcal{X}_{\mathrm{m}}^{2}} + \beta_{k_1}^1\mathbb{I}_{A_2^{\mathrm{c}}})(x) \geq \beta_{k_1}^1\underline{\mathrm{S}}^{k_2}\mathbb{I}_{A_2^{\mathrm{c}}\cup\mathcal{X}_{\mathrm{m}}^{2}}(x) \geq \beta_{k_1}^1\beta_{k_2}^2 > 0,$$

so D^{c} is lower reachable from x.

If $d = 2$, we're done. If $d > 3$, we simply repeat the same argument $d-2$ more times. In the end, we'll have concluded that D is lower reachable from every $x \in \mathcal{X}$. ◻

Proof of Proposition 16. Fix any $f \in \mathbb{R}^{\mathcal{X}}$ and consider its limit set $\Omega_f = \{\omega_1, \ldots, \omega_{p_f}\}$. Thanks to the assumptions in the statement, we know from Lemma 19 that for all $k \in \{1, \ldots, d\}$ and $\ell \in \{1, \ldots, M_k\}$, $\omega_1{\downarrow}_{\mathcal{X}_{\mathrm{m},\ell}^{k}}, \ldots, \omega_{p_f}{\downarrow}_{\mathcal{X}_{\mathrm{m},\ell}^{k}}$ are all constant and equal.

Let $D \coloneqq \bigcup_{k=1}^{d} \mathcal{X}_{\mathrm{lr}}^{k}$, and consider the upper transition operator $\overline{\mathrm{S}}$ as defined in Lemma 20. This way, $\overline{\mathrm{S}}\omega_i = \overline{\mathrm{T}}\omega_i = \omega_{i+1}$ for all $i \in \{1, \ldots, p_f\}$. Consequently, the limit set of $(\overline{\mathrm{S}}^n\omega_1)_{n\in\mathbb{N}}$ is $\{\omega_1, \ldots, \omega_{p_f}\}$. However, since $D^{\mathrm{c}} = \mathcal{X}_{\mathrm{m}}^{1} \cup \cdots \cup \mathcal{X}_{\mathrm{m}}^{d}$ is closed and lower reachable from any $x \in \mathcal{X}$ for $\overline{\mathrm{S}}$ [Lemma 20], it follows from [De Bock et al., 2025, Lemma 4.4] that $\omega_1 = \cdots = \omega_{p_f}$, and therefore $p_f = 1$. ◻

C.2 NECESSITY

Finally, we set out to prove that the condition in Theorem 12 is not only sufficient for $\overline{\mathrm{T}}$ to be convergent, but also necessary.

Proposition 21. *If an upper transition operator $\overline{\mathrm{T}}$ is convergent, then for all $k \in \{1, \ldots, d\}$ the A_k-maximal communication classes $\mathcal{X}_{\mathrm{m},1}^{k}, \ldots, \mathcal{X}_{\mathrm{m},M_k}^{k}$ are A_k-regular.*

Our proof is quite long, which is why we split it up into a couple of intermediary results. The first one of these is a basic result regarding $\{0,1\}$-valued matrices, which should be obvious; in essence, it's related to Lemma 8.

Lemma 22. *Consider any non-empty finite set $\mathcal{Z}$ and some matrix $M\colon \mathcal{Z}\times\mathcal{Z}\to\{0,1\}$. Then for all $x,y\in\mathcal{Z}$ and $n\in\mathbb{N}$, $M^n(x,y)>0$ if and only if there is some sequence $x=z_0,z_1,\dots,z_n=y$ in $\mathcal{Z}$ such that $M(z_{k-1},z_k)=1$ for all $k\in\{1,\dots,n\}$.*

The following two intermediary results are a bit more involved.

Lemma 23. *Consider some filtering class A and some maximal communication class $C\subseteq A$. If $\overline{\mathrm{T}}$ is convergent on C, then for all $x,y\in C$ and $\alpha\in\mathbb{R}_{>0}$, $\lim_{n\to+\infty}\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c}+\mathbb{I}_y)(x)>0$.*

Proof. Fix any $x,y\in C$ and $\alpha\in\mathbb{R}_{>0}$. For all $z\in C$, since C is a communication class, y is A-accessible from z, so there is some $m_z\in\mathbb{N}$ such that $c_z:=\overline{\mathrm{T}}^{m_z}(-\alpha\mathbb{I}_{A^c}+\mathbb{I}_y)(z)>0$. Let $c:=\min_{z\in C}c_z>0$.

Since $\overline{\mathrm{T}}$ is convergent on C, the limit $L:=\lim_{n\to+\infty}\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c}+\mathbb{I}_y)(x)$ exists. We need to prove that $L>0$. Since the limit exists, there is some $N\in\mathbb{N}$ such that

$$\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A^c}+\mathbb{I}_y)(x)\leq L+\frac{c}{2|C|}\quad\text{for all } n\in\mathbb{N}, n\geq N. \tag{8}$$

Since C is A-closed, it follows from Lemma 17 that $\overline{\mathrm{T}}^N(-|C|\frac{\alpha}{c}\mathbb{I}_{A^c}+\mathbb{I}_C)(x)\geq 1$, which, since

$$\overline{\mathrm{T}}^N\Big(-|C|\frac{\alpha}{c}\mathbb{I}_{A^c}+\mathbb{I}_C\Big)(x)=\overline{\mathrm{T}}^N\Bigg(\sum_{z\in C}-\frac{\alpha}{c}\mathbb{I}_{A^c}+\mathbb{I}_z\Bigg)(x)\leq\sum_{z\in C}\overline{\mathrm{T}}^N\Big(-\frac{\alpha}{c}\mathbb{I}_{A^c}+\mathbb{I}_z\Big)(x),$$

implies that there is some $z\in C$ such that $\overline{\mathrm{T}}^N(-\frac{\alpha}{c}\mathbb{I}_{A^c}+\mathbb{I}_z)(x)\geq\frac{1}{|C|}$.

By (T6), Lemma 13 and the definition of c_z, we have that $\overline{\mathrm{T}}^{m_z}(-\alpha\mathbb{I}_{A^c}+\mathbb{I}_y)\geq-\alpha\mathbb{I}_{A^c}+c_z\mathbb{I}_z\geq-\alpha\mathbb{I}_{A^c}+c\mathbb{I}_z$, and therefore

$$\overline{\mathrm{T}}^{N+m_z}(-\alpha\mathbb{I}_{A^c}+\mathbb{I}_y)(x)=\overline{\mathrm{T}}^N\overline{\mathrm{T}}^{m_z}(-\alpha\mathbb{I}_{A^c}+\mathbb{I}_y)(x)\geq\overline{\mathrm{T}}^N(-\alpha\mathbb{I}_{A^c}+c\mathbb{I}_z)(x)=c\overline{\mathrm{T}}^N(-\tfrac{\alpha}{c}\mathbb{I}_{A^c}+\mathbb{I}_z)(x)\geq\frac{c}{|C|}.$$

Combining this inequality with the one in (8) for $n=N+m_z\geq N$, we find that $\frac{c}{|C|}\leq L+\frac{c}{2|C|}$, from which we infer that $L\geq\frac{c}{2|C|}>0$, as desired. □

Lemma 24. *For all $k\in\{1,\dots,d\}$, there is some $\gamma_k\in\,]0,1]$ such that, for all $x\in A_k^c$ and $n\in\mathbb{Z}_{\geq0}$, $\underline{\mathrm{T}}^n\mathbb{I}_{A_k^c}(x)\geq\gamma_k$.*

Proof. We provide a proof by induction on k. For the base case $k=1$, we have $A_1=\mathcal{X}$ and therefore $A_1^c=\emptyset$. The statement therefore holds trivially for any γ_1. We choose $\gamma_1:=1$. Now assume that the statement holds for some $k\in\{1,\dots,d-1\}$; we will show that it also holds for $k+1$.

For all $z\in\mathcal{X}_{\mathrm{m}}^k$, since $\mathcal{X}_{\mathrm{m}}^k$ is A_k-closed, we know from Lemma 15 that there is some $\alpha_z>0$ such that $\overline{\mathrm{T}}(-\alpha_z\mathbb{I}_{A_k^c}+\mathbb{I}_{(\mathcal{X}_{\mathrm{m}}^k)^c})(z)\leq 0$. Fix any $\alpha>1$ such that $\alpha\geq\alpha_z$ for all $z\in\mathcal{X}_{\mathrm{m}}^k$. Let $B:=A_k^c\cup\mathcal{X}_{\mathrm{m}}^k$. Then by Lemmas 3 and 11, there is some $m\in\mathbb{N}$ such that $\underline{\mathrm{T}}^m\mathbb{I}_B(z)>0$ for all $z\in B\cup\mathcal{X}_{\mathrm{lr}}^k$; let

$$\delta:=\min\big\{\underline{\mathrm{T}}^m\mathbb{I}_B(z)\colon z\in\mathcal{X}_{\mathrm{lr}}^k\big\}\in\,]0,1].$$

Since A_{k+1} is a filtering class, it follows from Lemma 6 that there is some $\beta\in\,]0,1]$ such that $\underline{\mathrm{T}}^n\mathbb{I}_{A_{k+1}^c}\geq\beta^n\mathbb{I}_{A_{k+1}^c}$ for all $n\in\mathbb{Z}_{\geq0}$. Let $\gamma_{k+1}:=\min\big\{\delta\frac{\gamma_k}{\alpha},\beta^{m-1}\big\}>0$.

Consider now any $x\in A_{k+1}^c$ and $n\in\mathbb{Z}_{\geq0}$. We need to prove that $\underline{\mathrm{T}}^n\mathbb{I}_{A_k^c}(x)\geq\gamma_k$. If $n<m$, this follows from the definition of β: $\underline{\mathrm{T}}^n\mathbb{I}_{A_{k+1}^c}(x)\geq\beta^n\geq\beta^{m-1}\geq\gamma_{k+1}$. If $n\geq m$, the argument is a bit more involved.

We first prove by induction on $\ell\in\mathbb{Z}_{\geq0}$ that $\underline{\mathrm{T}}^\ell\mathbb{I}_B\geq\frac{\gamma_k}{\alpha}\mathbb{I}_B$ for all $\ell\in\mathbb{Z}_{\geq0}$. The base case $\ell=0$ holds trivially since $\underline{\mathrm{T}}^0\mathbb{I}_B=\mathbb{I}_B$ and $\frac{\gamma_k}{\alpha}\leq1$. For the inductive step, we assume that the statement holds for some $\ell\in\mathbb{Z}_{\geq0}$ and show that it also holds for $\ell+1$.

Consider any $z\in\mathcal{X}_{\mathrm{m}}^k$. By the hypothesis of our induction on k and (T6), we have that $\underline{\mathrm{T}}^\ell\mathbb{I}_B(y)\geq\underline{\mathrm{T}}\mathbb{I}_{A_k^c}(y)\geq\gamma_k$ for all $y\in A_k^c$. By the hypothesis of our induction on ℓ, we have $\underline{\mathrm{T}}^\ell\mathbb{I}_B(y)\geq\frac{\gamma_k}{\alpha}$ for all $y\in\mathcal{X}_{\mathrm{m}}^k$. Since $\underline{\mathrm{T}}^\ell\mathbb{I}_B\geq0$ by (T5), it follows that $\underline{\mathrm{T}}^\ell\mathbb{I}_B\geq\gamma_k\mathbb{I}_{A_k^c}+\frac{\gamma_k}{\alpha}\mathbb{I}_{\mathcal{X}_{\mathrm{m}}^k}$, from which we infer that

$$\underline{\mathrm{T}}^{\ell+1}\mathbb{I}_B(z)=\underline{\mathrm{T}}\underline{\mathrm{T}}^\ell\mathbb{I}_B(z)\geq\underline{\mathrm{T}}\Big(\gamma_k\mathbb{I}_{A_k^c}+\frac{\gamma_k}{\alpha}\mathbb{I}_{\mathcal{X}_{\mathrm{m}}^k}\Big)(z)=\frac{\gamma_k}{\alpha}\underline{\mathrm{T}}(\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_{\mathcal{X}_{\mathrm{m}}^k})(z). \tag{9}$$

Since $\overline{\mathrm{T}}(-\alpha_x\mathbb{I}_{A_k^c}+\mathbb{I}_{(\mathcal{X}_{\mathrm{m}}^k)^c})(z)\leq 0$ and $\alpha\geq\alpha_z$, it follows from (T6) that

$$\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_{(\mathcal{X}_{\mathrm{m}}^k)^c})(z)\leq\overline{\mathrm{T}}(-\alpha_z\mathbb{I}_{A_k^c}+\mathbb{I}_{(\mathcal{X}_{\mathrm{m}}^k)^c})(z)\leq 0,$$

implying that

$$\underline{\mathrm{T}}(\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_{\mathcal{X}_{\mathrm{m}}^k})(z)=\underline{\mathrm{T}}(\alpha\mathbb{I}_{A_k^c}+1-\mathbb{I}_{(\mathcal{X}_{\mathrm{m}}^k)^c})(z)=1+\underline{\mathrm{T}}(\alpha\mathbb{I}_{A_k^c}-\mathbb{I}_{(\mathcal{X}_{\mathrm{m}}^k)^c})(z)=1-\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_{(\mathcal{X}_{\mathrm{m}}^k)^c})(z)\geq 1.$$

Substituting this back into Equation (9), we find that $\underline{\mathrm{T}}^{\ell+1}\mathbb{I}_B(z)\geq\frac{\gamma_k}{\alpha}$. On the other hand, for all $z\in A_k^c$, it follows from the hypothesis of our induction on k that $\underline{\mathrm{T}}^{\ell+1}\mathbb{I}_B(z)\geq\gamma_k\geq\frac{\gamma_k}{\alpha}$. Hence, $\underline{\mathrm{T}}^{\ell+1}\mathbb{I}_B(z)\geq\frac{\gamma_k}{\alpha}$ for all $z\in A_k^c\cup\mathcal{X}_{\mathrm{m}}^k=B$. Since $\underline{\mathrm{T}}^{\ell+1}\mathbb{I}_B\geq 0$, it follows that $\underline{\mathrm{T}}^{\ell+1}\mathbb{I}_B\geq\frac{\gamma_k}{\alpha}\mathbb{I}_B$. This concludes the induction on ℓ.

Recall from Section 4.2 that $A_{k+1}^c=A_k^c\cup\mathcal{X}_{\mathrm{m}}^k\cup\mathcal{X}_{\mathrm{lr}}^k$. If $x\in A_k^c\cup\mathcal{X}_{\mathrm{m}}^k$, then (T6) and the induction on ℓ above immediately yield that

$$\underline{\mathrm{T}}^n\mathbb{I}_{A_{k+1}^c}(x)\geq\underline{\mathrm{T}}^n\mathbb{I}_B(x)\geq\frac{\gamma_k}{\alpha}\mathbb{I}_B(x)=\frac{\gamma_k}{\alpha}\geq\delta\frac{\gamma_k}{\alpha}\geq\gamma_{k+1}.$$

If instead $x\in\mathcal{X}_{\mathrm{lr}}^k$, then the induction on ℓ above implies that $\underline{\mathrm{T}}^{n-m}\mathbb{I}_B\geq\frac{\gamma_k}{\alpha}\mathbb{I}_B$. It therefore follows from (T6) that, once more,

$$\underline{\mathrm{T}}^n\mathbb{I}_{A_{k+1}^c}(x)\geq\underline{\mathrm{T}}^n\mathbb{I}_B(x)=\underline{\mathrm{T}}^m\underline{\mathrm{T}}^{n-m}\mathbb{I}_B(x)\geq\underline{\mathrm{T}}^m\Big(\frac{\gamma_k}{\alpha}\mathbb{I}_B\Big)(x)=\frac{\gamma_k}{\alpha}\underline{\mathrm{T}}^m\mathbb{I}_B(x)\geq\delta\frac{\gamma_k}{\alpha}\geq\gamma_{k+1}.$$

This concludes the induction on k and therefore also the proof. □

Finally, then, we are ready to prove the necessity of our condition.

Proof of Proposition 21. Consider any $k\in\{1,\dots,d\}$ and $\ell\in\{1,\dots,M_k\}$. We need to prove that $C:=\mathcal{X}_{\mathrm{m},\ell}^k$ is A_k-regular. Let $\gamma_k\in\,]0,1]$ be as in Lemma 24 and let $\beta_k:=\gamma_k/(2|A_k|)\in\,]0,1]$. Since C is A_k-closed, we know that for all $x\in C$ and $y\in C^c$, there is some $\alpha_{x,y}\in\mathbb{R}_{>0}$ such that $\overline{\mathrm{T}}(-\alpha_{x,y}\beta_k\mathbb{I}_{A_k^c}+\mathbb{I}_y)(x)\leq 0$. Let $\underline{\alpha}:=\max\{\alpha_{x,y}\colon x\in C,y\in C^c\}>0$. Then clearly, for any $\alpha\geq\underline{\alpha}$, $x\in C$ and $y\in C^c$, $\overline{\mathrm{T}}(-\alpha\beta_k\mathbb{I}_{A_k^c}+\mathbb{I}_y)(x)\leq 0$.

Fix any $\alpha\geq\min\{2/\gamma_k,\underline{\alpha}\}>0$ and consider the matrix $M_\alpha\colon C\times C\to\{0,1\}$ defined for all $x,y\in C$ by

$$M_\alpha(x,y):=\begin{cases}1 & \text{if }\overline{\mathrm{T}}(-\alpha\beta_k\mathbb{I}_{A_k^c}+\mathbb{I}_y)(x)>0,\\ 0 & \text{otherwise.}\end{cases}$$

Since $\overline{\mathrm{T}}$ is convergent, it is in particular convergent on C. It therefore follows from Lemma 23 that, for all $x,y\in C$, $\lim_{n\to+\infty}\overline{\mathrm{T}}^n(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_y)(x)>0$. This implies that there is some $n_\alpha\in\mathbb{N}$ such that $\overline{\mathrm{T}}^{n_\alpha}(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_y)(x)>0$ for all $x,y\in C$.

Fix any $x,y\in C$ and let $z_0:=x\in C$. For all $z\in A_k^c$,

$$\overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_y)(z)\leq\overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c}+1)(z)=1-\alpha\underline{\mathrm{T}}^{n_\alpha-1}\mathbb{I}_{A_k^c}(z)\leq 1-\alpha\gamma_k,$$

using Lemma 24 for the last inequality. Since $\alpha\geq 2/\gamma_k$, this implies that for all $z\in A_k^c$,

$$\overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_y)(z)\leq 1-\frac{\alpha\gamma_k}{2}-\frac{\alpha\gamma_k}{2}\leq-\frac{\alpha\gamma_k}{2}=-|A_k|\alpha\beta_k=-\sum_{z\in A_k}\alpha\beta_k.$$

Hence,

$$\overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_y)\leq\sum_{z\in A_k}\alpha\beta_k\mathbb{I}_{A_k^c}+\overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_y)(z)\mathbb{I}_z.\tag{10}$$

Since $\overline{\mathrm{T}}^{n_\alpha}(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_y)(x)>0$, it follows that

$$\begin{aligned}0<\overline{\mathrm{T}}^{n_\alpha}(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_y)(z_0)&=\overline{\mathrm{T}}\,\overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_y)(z_0)\\&\leq\overline{\mathrm{T}}\left(\sum_{z\in A_k}-\alpha\beta_k\mathbb{I}_{A_k^c}+\overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_y)(z)\mathbb{I}_z\right)(z_0)\\&\leq\sum_{z\in A_k}\overline{\mathrm{T}}\big(-\alpha\beta_k\mathbb{I}_{A_k^c}+\overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c}+\mathbb{I}_y)(z)\mathbb{I}_z\big)(z_0).\end{aligned}$$

This inequality implies that there is some $z_1 \in A_k$ such that

$$0 < \overline{\mathrm{T}}\big(-\alpha\beta_k \mathbb{I}_{A_k^c} + \overline{\mathrm{T}}^{n_\alpha - 1}(-\alpha \mathbb{I}_{A_k^c} + \mathbb{I}_y)(z_1)\mathbb{I}_{z_1}\big)(z_0). \tag{11}$$

On the one hand, since $\overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c} + \mathbb{I}_y)(z_1) \leq 1$ by (T1), it follows from (11) that

$$\overline{\mathrm{T}}(-\alpha\beta_k\mathbb{I}_{A_k^c} + \mathbb{I}_{z_1})(z_0) \geq \overline{\mathrm{T}}\big(-\alpha\beta_k\mathbb{I}_{A_k^c} + \overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c} + \mathbb{I}_y)(z_1)\mathbb{I}_{z_1}\big)(z_0) > 0.$$

Since it's not possible that $z_1 \in C^c$—recall that $\overline{\mathrm{T}}(-\alpha\beta_k\mathbb{I}_{A_k^c} + \mathbb{I}_{z_1})(z_0) \leq 0$ as $\alpha \geq \underline{\alpha}$—it must be that $z_1 \in C$.

On the other hand, (11) also implies that $\overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c} + \mathbb{I}_y)(z_1) > 0$, since otherwise it would follow from (T1) that

$$\overline{\mathrm{T}}\big(-\alpha\beta_k\mathbb{I}_{A_k^c} + \overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c} + \mathbb{I}_y)(z_1)\mathbb{I}_{z_1}\big)(z_0) \leq 0.$$

So we conclude that there is some $z_1 \in C$ such that

$$M_\alpha(z_0, z_1) = 1 \quad \text{and} \quad \overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c} + \mathbb{I}_y)(z_1) > 0.$$

Using a similar argument, this time starting from $\overline{\mathrm{T}}^{n_\alpha-1}(-\alpha\mathbb{I}_{A_k^c} + \mathbb{I}_y)(z_1) > 0$ instead of $\overline{\mathrm{T}}^{n_\alpha}(-\alpha\mathbb{I}_{A_k^c} + \mathbb{I}_y)(z_0) > 0$, we find some $z_2 \in C$ such that

$$M_\alpha(z_1, z_2) = 1 \quad \text{and} \quad \overline{\mathrm{T}}^{n_\alpha-2}(-\alpha\mathbb{I}_{A_k^c} + \mathbb{I}_y)(z_2) > 0.$$

Repeating this argument $n_\alpha - 3$ times more, we find a sequence $z_0, z_1, \ldots, z_{n_\alpha-1} \in C$ such that

$$M_\alpha(z_{i-1}, z_i) = 1 \quad \text{and} \quad \overline{\mathrm{T}}^{n_\alpha-i}(-\alpha\mathbb{I}_{A_k^c} + \mathbb{I}_y)(z_i) > 0 \quad \text{for all } i \in \{1, \ldots, n_\alpha - 1\}.$$

In particular, for $i = n_\alpha - 1$, we find that $\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A_k^c} + \mathbb{I}_y)(z_{n_\alpha-1}) > 0$ and therefore also, with $z_{n_\alpha} \coloneqq y \in C$, that

$$\overline{\mathrm{T}}\big(-\alpha\beta_k\mathbb{I}_{A_k^c} + \mathbb{I}_{z_{n_\alpha}}\big)(z_{n_\alpha-1}) \geq \overline{\mathrm{T}}(-\alpha\mathbb{I}_{A_k^c} + \mathbb{I}_{z_{n_\alpha}})(z_{n_\alpha-1}) > 0.$$

We have thus found that there is a sequence $x = z_0, z_1, \ldots, z_{n_\alpha} = y$ in C such that, for all $i \in \{1, \ldots, n_\alpha\}$, $M_\alpha(z_{i-1}, z_i) = 1$. Using Lemma 22, we conclude that $M_\alpha^{n_\alpha}(x, y) > 1$.

Since our choices of $x, y \in C$ were arbitrary, we have shown that $M_\alpha^{n_\alpha} \geq 1$. By Theorems 3.4.4 and 3.5.6 in [Brualdi and Ryser, 1991], this implies that $M_\alpha^n > 0$ for all $n \geq N \coloneqq (|C| - 1)^2 + 1$. Applying Lemma 22 again, we find that for all $x, y \in C$ and $n \geq N$, there is some sequence $x = z_0, z_1, \ldots, z_n = y$ in C such that, for all $i \in \{1, \ldots, n\}$, $M_\alpha(z_{i-1}, z_i) = 1$, or equivalently, $\overline{\mathrm{T}}(-\alpha\beta_k\mathbb{I}_{A_k^c} + \mathbb{I}_{z_i})(z_{i-1}) > 0$.

Since our choice of $\alpha \geq \min\{2/\gamma_k, \underline{\alpha}\}$ was arbitrary, we conclude that for all $x, y \in C$ and $n \geq N$, there is for all $\alpha \geq \min\{2/\gamma_k, \underline{\alpha}\}$ some sequence $x = z_0, z_1, \ldots, z_n = y$ in C such that, for all $i \in \{1, \ldots, n\}$, $\overline{\mathrm{T}}(-\alpha\beta_k\mathbb{I}_{A_k^c} + \mathbb{I}_{z_i})(z_{i-1}) > 0$. Since there are only finitely many possible sequences $x = z_0, z_1, \ldots, z_n = y$ in C, there is therefore at least one such sequence that appears for arbitrarily high values of α. For that particular sequence $x = z_0, z_1, \ldots, z_n = y$ in C, it follows that for all $i \in \{1, \ldots, n\}$ and $\alpha \in \mathbb{R}_{>0}$, $\overline{\mathrm{T}}(-\alpha\mathbb{I}_{A_k^c} + \mathbb{I}_{z_i})(z_{i-1}) > 0$. From this and Lemma 8, we conclude that $x \xrightarrow{A_k, n} y$. Since our choices of $x, y \in C$ and $n \geq N$ were arbitrary, we conclude that C is A_k-regular. □