The Limit Behaviour of Imprecise Continuous-Time Markov Chains

Jasper De Bock[1]

Abstract We study the limit behaviour of a nonlinear differential equation whose solution is a superadditive generalisation of a stochastic matrix, prove convergence, and provide necessary and sufficient conditions for ergodicity. In the linear case, the solution of our differential equation is equal to the matrix exponential of an intensity matrix and can then be interpreted as the transition operator of a homogeneous continuous-time Markov chain. Similarly, in the generalised nonlinear case that we consider, the solution can be interpreted as the lower transition operator of a specific set of non-homogeneous continuous-time Markov chains, called an imprecise continuous-time Markov chain. In this context, our convergence result shows that for a fixed initial state, an imprecise continuous-time Markov chain always converges to a limiting distribution, and our ergodicity result provides a necessary and sufficient condition for this limiting distribution to be independent of the initial state.

Keywords Markov chain · Continuous time · Imprecise · Convergence · Limiting distribution · Ergodicity · Matrix exponential · Lower transition operator · Lower transition rate operator

Mathematics Subject Classification Primary 34K25 · Secondary 34K50 · 60J27 · 60J35

Communicated by Charles Doering.

✉ Jasper De Bock
jasper.debock@ugent.be

1 Data Science Lab, Ghent University, Technologiepark 914, 9052 Zwijnaarde, Belgium

Springer

1 Introduction

Consider a real-valued $n \times n$ matrix Q, and let T_t be a real-valued time-dependent $n \times n$ matrix such that

$$\frac{\mathrm{d}}{\mathrm{d}t} T_t = Q T_t \quad \text{for all } t \geq 0$$

and $T_0 = I$, with I the n-dimensional unit matrix. The unique solution of this differential equation is then well known to be given by the matrix exponential e^{Qt} of Q. If Q is furthermore an intensity matrix—has non-negative off-diagonal elements and rows that sum to zero—then $T_t = e^{Qt}$ will be a stochastic matrix. In that case, T_t can be interpreted as the transition operator of a homogeneous continuous-time Markov chain. Indeed, if we identify $\{1, \ldots, n\}$ with the state space $\mathcal{X}$ of such a Markov chain and let Q be its transition rate matrix, then for any two states $x, y \in \mathcal{X}$, $T_t(x, y)$ is the probability $P(X_t = y | X_0 = x)$ of ending up in state y at time t, conditional on starting in state x at time zero.

Rather remarkably, for any transition rate matrix Q, the conditional probability $P(X_t = y | X_0 = x)$ will always converge ([Chung 1967](), Theorem II.10.1) as t approaches infinity. However, in general, this limiting value may depend on the initial state x. If this is not the case, that is, if there is a probability mass function P_∞ on $\mathcal{X}$ such that

$$\lim_{t \to +\infty} P(X_t = y | X_0 = x) = P_\infty(y) \quad \text{for all } y \in \mathcal{X}, \tag{1}$$

then the homogeneous continuous-time Markov chain under consideration—or, equivalently, the transition rate matrix Q—is said to have a unique limiting distribution P_∞. From an applied point of view, the existence of such a unique limiting distribution is very important, because it implies that for large enough values of t, predicting the current value of X_t does not require any knowledge about its initial values. Hence, we are led to the following question: what conditions does Q need to satisfy in order for P_∞ to exist? As it turns out, this question has an elegant answer: the required conditions are relatively easy and are fully determined by the signs of the components of Q; see, for example, [Anderson (1991)]().

Our main goal here is to extend these results to a nonlinear context, where the intensity matrix Q is replaced by a lower transition rate operator $\underline{Q}$, which is a nonlinear—superadditive—generalisation of an intensity matrix. Much as in the original case, this lower transition rate operator gives rise to a corresponding lower transition operator $\underline{T}_t$, which is a nonlinear—superadditive—generalisation of a stochastic matrix. Specifically, for every real-valued function f on $\mathcal{X}$, $\underline{T}_t f$ is completely determined by the nonlinear differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} \underline{T}_t f = \underline{Q} \underline{T}_t f \quad \text{for all } t \geq 0, \tag{2}$$

with boundary condition $\underline{T}_0 f = f$ ([Škulj 2015]()). The aim of this paper is to study the properties of the operator $\underline{T}_t$ and, in particular, its limit behaviour as t approaches infinity. Our first main contribution—see Theorem [12]()—is a proof that $\underline{T}_t$ will always converge to a limiting operator $\underline{T}_\infty$. Our second main contribution—see Theorem [19]()—

is a simple necessary and sufficient condition for $\underline{Q}$ to be ergodic, in the sense that for all real-valued functions f on $\mathcal{X}$, $\underline{T}_\infty f$ is constant.

Our motivation for developing these results, and the reason for this paper's title, is that $\underline{T}_t f(x)$ can be interpreted as the conditional lower expectation $\underline{\mathbb{E}}(f(X_t)|X_0 = x)$ of an imprecise continuous-time Markov chain, which, basically, is a set of continuous-time Markov chains whose possibly time-dependent transition rate matrix Q_t is partially specified, in the sense that all that we known about it is that it takes values in some given set of transition rate matrices $\mathcal{Q}$.[1] Indeed, as recently shown in Škulj (2015) and Krak and De Bock (submitted), for the largest such set of Markov chains, and under relatively mild conditions on $\mathcal{Q}$,[2] the tightest possible lower bound on the conditional expectation $\mathbb{E}(f(X_t)|X_0 = x)$—the conditional lower expectation $\underline{\mathbb{E}}(f(X_t)|X_0 = x)$—is equal to the solution $\underline{T}_t f$ of the differential equation (2), with $\underline{Q}$ the lower envelope of $\mathcal{Q}$.

Therefore, our results can be interpreted as statements about the limit behaviour of the conditional lower expectation $\underline{\mathbb{E}}(f(X_t)|X_0 = x)$. For example, our first main result then states that $\underline{\mathbb{E}}(f(X_t)|X_0 = x)$ is guaranteed to converge to a limiting value as t approaches infinity. Similarly, the notion of ergodicity that we consider can then be interpreted to mean that this limiting value does not depend on the initial state x, or equivalently, that the imprecise continuous-time Markov chain under study has a unique limiting lower expectation functional $\underline{\mathbb{E}}_\infty$, in the sense that

$$\underline{\mathbb{E}}_\infty(f) = \lim_{t \to +\infty} \underline{\mathbb{E}}(f(X_t)|X_0 = x) \quad \text{for all } x \in \mathcal{X} \text{ and all real functions } f \text{ on } \mathcal{X}.$$
(3)

In this way, our second main result—a necessary and sufficient condition for $\underline{Q}$ to be ergodic—turns into a practical tool: it provides a simple criterion for checking whether or not a given imprecise continuous-time Markov chain has a unique limiting lower expectation functional $\underline{\mathbb{E}}_\infty$. In the special case where the lower transition rate operator $\underline{Q}$ is actually a transition rate matrix Q, our notion of ergodicity coincides with the usual one and, in that case, our results can be used to check whether the continuous-time Markov chain that corresponds to Q has a unique limiting distribution P_∞, whose expectation functional $\mathbb{E}_\infty$ will then be equal to $\underline{\mathbb{E}}_\infty$.

That being said, this paper does not adopt any specific interpretation, but takes a purely mathematical point of view. Our object of study here is the solution $\underline{T}_t$ of the differential equation (2), and our main results are a proof that $\underline{T}_t$ converges to a limit $\underline{T}_\infty$, and a necessary and sufficient condition for $\underline{Q}$ to be ergodic, in the sense that $\underline{T}_\infty$ maps functions to constants. As explained above, these results are directly applicable

---

[1] Alternatively, an imprecise continuous-time Markov chain can also be identified with an even larger set of stochastic processes, not all of which are continuous-time Markov chains. Loosely speaking, the partially specified transition rate matrix $Q \in \mathcal{Q}$ is then allowed to be time-dependent *and* history-dependent (in the sense that it may depend on the value of the state at previous time points); see Krak and De Bock (submitted) for more information. Here too, $\underline{T}_t f(x)$ can be interpreted as the conditional lower expectation $\underline{\mathbb{E}}(f(X_t)|X_0 = x)$ (Krak and De Bock, submitted).

[2] It should have separately specified rows. Loosely speaking, this means that the rows of the intensity matrices $Q \in \mathcal{Q}$ are allowed to vary independently of each other. More formally, it means that if Q is an intensity matrix such that, for all $x \in \mathcal{X}$, there is some $Q_x \in \mathcal{Q}$ of which the x-row is equal to the x-row of Q—in the sense that $Q_x(x, y) = Q(x, y)$ for all $y \in \mathcal{X}$—then Q should also be an element of $\mathcal{Q}$.

to—and inspired by—the theory of imprecise continuous-time Markov chains; more information about this field of study can be found in Škulj (2012, 2015), Krak and De Bock (submitted) and Troffaes et al. (2015). However, our results should also be of interest to other fields whose aim it is to robustify the theory of continuous-time Markov chains, such as continuous-time Markov decision processes (Guo and Hernández-Lerma 2009), continuous-time-controlled Markov chains (Guo and Hernández-Lerma 2003), and interval continuous-time Markov chains (Galdino 2013). More generally, we think that our ideas and results are relevant to any theory that studies—or requires—a robust generalisation of the matrix exponential of an intensity matrix.

We end this introduction with a brief overview of the structure of this paper. After Sect. 2, in which we introduce some basic preliminary concepts, the rest of this paper is structured as follows.

We start in Sect. 3 by introducing the concept of a lower transition operator $\underline{T}$, which is a nonlinear—superadditive—generalisation of a stochastic matrix; the operator $\underline{T}_t$ that is studied in this paper is a special case. We provide a definition, explain the connection with coherent lower previsions (Troffaes and de Cooman 2014; Walley 1991), and use this connection to establish a number of technical properties.

Section 4 then goes on to define ergodicity for lower transition operators, which is a discrete-time version of the notion of ergodicity that we study in this paper, and recalls that a lower transition operator $\underline{T}$ will exhibit this type of ergodicity if and only if it is regularly absorbing (Hermans and Cooman 2012). We also introduce a new property, called being 1-step absorbing, and show that it is a sufficient condition for $\underline{T}$ to be ergodic.

Next, in Sect. 5, we introduce the concept of a lower transition rate operator $\underline{Q}$, which, as already mentioned before, is a nonlinear—superadditive—generalisation of an intensity matrix. We provide a definition, prove a number of properties, and establish a connection with lower transition operators.

Having introduced all of these related concepts and their properties, the rest of this paper focusses on our main object of interest, which is the time-dependent lower transition operator $\underline{T}_t$ that corresponds to a given lower transition rate operator $\underline{Q}$. Section 6 defines this operator as the unique solution to Eq. (2), shows that it is indeed a lower transition operator, and then proves that it also satisfies another—closely related—differential equation, which applies directly to $\underline{T}_t$ rather than $\underline{T}_t f$. We end this section by establishing a limit expression for $\underline{T}_t$, which resembles—and generalises—the well-known limit definition of a matrix exponential.

With these characterisations of $\underline{T}_t$ in hand, Sect. 7 then moves on to study its limit behaviour and, in particular, its convergence. The main result of this section is that $\underline{T}_t$ is guaranteed to converge as t goes to infinity, and that this limit $\underline{T}_\infty$ is again a lower transition operator.

Section 8 is again concerned with ergodicity, but now in a continuous-time sense. First of all, we show that $\underline{Q}$ is ergodic—$\underline{T}_\infty(f)(x)$ does not depend on x—if and only if, for any $t > 0$, $\underline{T}_t$ is ergodic in the discrete-time sense of Sect. 4. Secondly, for any $t > 0$, we show that $\underline{T}_t$ is regularly absorbing if and only if it is 1-step absorbing. Thirdly, we establish a simple qualitative method for checking whether $\underline{T}_t$ is 1-step absorbing; this method does not depend on t and is expressed directly in terms of the lower transition rate operator $\underline{Q}$. Finally, we explain how these three results, when

combined, lead to a simple necessary and sufficient condition for Q to be ergodic. All that is needed in order to check this condition is the sign of a limited number of evaluations of Q.

Section 9 concludes this paper. It briefly discusses our main result and then goes on to suggest some ideas for future research, including a number of open questions that we consider to be important. The proofs of all our results are gathered in the Appendix; they are organised per section and in order of appearance. The appendix also contains some additional technical lemmas.

2 Preliminaries

Consider some finite *state space* $\mathcal{X}$, and let $\mathcal{L}(\mathcal{X})$ be the set of all real-valued functions on $\mathcal{X}$. A map A from $\mathcal{L}(\mathcal{X})$ to $\mathcal{L}(\mathcal{X})$ is called an operator, and we use I to denote the identity operator that maps any $f \in \mathcal{L}(\mathcal{X})$ to itself. An operator A is called non-negatively homogeneous if

$$A(\lambda f) = \lambda A(f) \quad \text{for all } f \in \mathcal{L}(\mathcal{X}) \text{ and all } \lambda \geq 0.$$

For any $S \subseteq \mathcal{X}$, we will let $\mathbb{I}_S \in \mathcal{L}(\mathcal{X})$ be the indicator of S, defined by $\mathbb{I}_S(x) := 1$ if $x \in S$ and $\mathbb{I}_S(x) := 0$ otherwise. If S is a singleton $\{x\}$, we also write $\mathbb{I}_x$ instead of $\mathbb{I}_{\{x\}}$. We will not distinguish between real numbers and constant functions in $\mathcal{L}(\mathcal{X})$, in the sense that whenever it is convenient to do so, we will interpret $a \in \mathbb{R}$ as a constant function in $\mathcal{L}(\mathcal{X})$, trivially defined by $a(x) := a$ for all $x \in \mathcal{X}$. We use $\mathbb{N}$ to refer to the set of natural numbers without zero and let $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.

For any $f \in \mathcal{L}(\mathcal{X})$, we let $\|f\| := \|f\|_\infty := \max\{|f(x)| : x \in \mathcal{X}\}$ be the maximum norm, and for any non-negatively homogeneous operator A, we consider the induced operator norm

$$\|A\| := \sup\{\|Af\| : f \in \mathcal{L}(\mathcal{X}), \|f\| = 1\}. \tag{4}$$

Not only do these norms satisfy the usual defining properties of a norm, they also satisfy the following additional properties; see the Appendix for a proof. For all $f \in \mathcal{L}(\mathcal{X})$ and any two non-negatively homogeneous operators A and B:

N1: $\|Af\| \leq \|A\| \|f\|$
N2: $\|AB\| \leq \|A\| \|B\|$

The limit of a sequence will always be taken with respect to the norm that corresponds to its elements. For example, we say that a sequence $\{A_n\}_{n \in \mathbb{N}}$ of non-negatively homogeneous operators converges to an operator A and write $\lim_{n \to +\infty} A_n = A$, if $\lim_{n \to +\infty} \|A_n - A\| = 0$, where $\|\cdot\|$ is the operator norm. As another example, for any $f \in \mathcal{L}(\mathcal{X})$, we say that $A_n f$ converges to Af and write $\lim_{n \to +\infty} A_n f = Af$, if $\lim_{n \to +\infty} \|Af - f\| = 0$, where $\|\cdot\|$ is now the maximum norm. If A_n converges to A, then for all $f \in \mathcal{L}(\mathcal{X})$, $A_n f$ converges to Af. However, in general, the converse may not be true. Nevertheless, for the specific operators that we will consider in this paper, we will see that these two types of convergence are equivalent; see Propositions 2 and 7. Since derivatives are defined using limits, similar comments apply to

derivatives as well. If A_t is a time-dependent non-negatively homogeneous operator, then $\frac{d}{dt} A_t = A$ implies that $\frac{d}{dt} A_t f = Af$ for all $f \in \mathcal{L}(\mathcal{X})$, but the converse may not be true. However, here too, we will show that for the specific type of operator that we consider, these statements are nevertheless equivalent; see Proposition 9.

Finally, we would like to remark that our results do not crucially depend on our use of the maximum norm. For example, if we were to replace the maximum norm by the Euclidean norm—or any other L^p-norm with $p \geq 1$—and consider the corresponding operator norm, then the resulting limit notion—and hence any result about convergence or derivatives—would be identical, simply because these norms are equivalent to ours, in the sense that they induce the same topologies. The reason why we prefer the maximum norm is simply because of it is practical convenience in the proofs.

3 Lower Transition Operators

The first important type of non-negatively homogeneous operator that we will consider in this paper is a lower transition operator $\underline{T}$. As we will show in Sect. 6, the solution $\underline{T}_t$ of the differential equation that we study in this paper is of this type.

Definition 1 (*Lower transition operator*) A lower transition operator $\underline{T}$ is an operator $\underline{T}$ such that for all $f, g \in \mathcal{L}(\mathcal{X})$ and $\lambda \geq 0$:

L1: $\underline{T} f \geq \min f$;
L2: $\underline{T}(f + g) \geq \underline{T}(f) + \underline{T}(g)$; [superadditivity]
L3: $\underline{T}(\lambda f) = \lambda \underline{T}(f)$. [non-negative homogeneity]

The corresponding upper transition operator $\overline{T}$ is defined by

$$\overline{T} f := -\underline{T}(-f) \quad \text{for all } f \in \mathcal{L}(\mathcal{X}). \tag{5}$$

Basically, a lower transition operator is just a superadditive generalisation of a stochastic matrix. If the superadditivity axiom is replaced by an additivity axiom, a lower transition operator will coincide with its upper transition operator and can then be identified with a stochastic matrix T.

Example 1 For all $x \in \mathcal{X}$, consider some $\alpha_x \in [0, 1]$, and let

$$\underline{T}(f)(x) := \alpha_x \min f + (1 - \alpha_x) f(x) \quad \text{for all } f \in \mathcal{L}(\mathcal{X}). \tag{6}$$

The operator $\underline{T}$ that is defined in this way is then clearly a lower transition operator. In order to see that, it suffices to check each of the defining properties L1–L3; we leave this as an exercise. If every α_x is 0, then $\underline{T}$ is equal to the identity operator I. If every α_x is 1, then $\underline{T}$ is equal to the min-operator—using our convention that the real number min f can be identified with a constant function in $\mathcal{L}(\mathcal{X})$.

For every lower transition operator $\underline{T}$ and any $x \in \mathcal{X}$, the functional $\underline{T}(\cdot)(x)$ is a coherent lower prevision (Troffaes and de Cooman 2014; Walley 1991): a superadditive, non-negatively homogeneous map from $\mathcal{L}(\mathcal{X})$ to $\mathbb{R}$ that dominates the minimum. Therefore, lower transition operators are basically just finite vectors of coherent lower

previsions. As a direct consequence, the following properties are implied by the corresponding versions for coherent lower previsions; see Walley (1991, 2.6.1). For any $f, g \in \mathcal{L}(\mathcal{X})$ and $\mu \in \mathbb{R}$ and all sequences $\{f_n\}_{n \in \mathbb{N}} \subseteq \mathcal{L}(\mathcal{X})$:

L4: $\min f \leq \underline{T}f \leq \overline{T}f \leq \max f$;
L5: $\underline{T}(f + \mu) = \underline{T}(f) + \mu$;
L6: $f \geq g \Rightarrow \underline{T}(f) \geq \underline{T}(g)$ and $\overline{T}(f) \geq \overline{T}(g)$;
L7: $\left| \underline{T}f - \underline{T}g \right| \leq \overline{T}(|f - g|)$;
L8: $\lim_{n \in \mathbb{N}} f_n = f \Rightarrow \lim_{n \to +\infty} \underline{T}f_n = \underline{T}f$.

As a rather straightforward consequence of L4 and L7, we also find that

L9: $\left\| \underline{T} \right\| \leq 1$;
L10: $\left\| \underline{T}f - \underline{T}g \right\| \leq \|f - g\|$; [non-expansiveness]
L11: $\left\| \underline{T}A - \underline{T}B \right\| \leq \|A - B\|$,

where A and B are non-negatively homogeneous operators from $\mathcal{L}(\mathcal{X})$ to $\mathcal{L}(\mathcal{X})$; see the Appendix for a proof.

We end this section with some basic results about sequences and powers of lower transition operators. First of all, if an operator is the pointwise limit of a sequence of lower transition operators, then this limit is again a lower transition operator.

Proposition 1 *Consider an operator $\underline{T}$ and a sequence $\{\underline{T}_n\}_{n \in \mathbb{N}}$ of lower transition operators such that $\lim_{n \to +\infty} \underline{T}_n f = \underline{T}f$ for all $f \in \mathcal{L}(\mathcal{X})$. Then, $\underline{T}$ is a lower transition operator.*

Furthermore, as our next result establishes, it does not matter whether this limit is taken pointwise or with respect to the operator norm.

Proposition 2 *For any lower transition operator $\underline{T}$ and any sequence $\{\underline{T}_n\}_{n \in \mathbb{N}}$ of lower transition operators:*

$$\lim_{n \to +\infty} \underline{T}_n = \underline{T} \Leftrightarrow (\forall f \in \mathcal{L}(\mathcal{X})) \quad \lim_{n \to +\infty} \underline{T}_n f = \underline{T}f.$$

Finally, since a lower transition operator $\underline{T}$ is always non-expansive—see L10—it follows from the results of Sine (1990) that in the limit, for large n, the powers $\underline{T}^n$ will start to follow a cyclic pattern, the period of which has a universal upper bound.

Proposition 3 *Let $\mathcal{X}$ be some fixed finite state space. Then, there is a natural number $p \in \mathbb{N}$ such that, for any lower transition operator $\underline{T}$ from $\mathcal{L}(\mathcal{X})$ to $\mathcal{L}(\mathcal{X})$:*

$$(\forall f \in \mathcal{L}(\mathcal{X}))(\exists g \in \mathcal{L}(\mathcal{X})) \quad \lim_{k \to +\infty} \underline{T}^{pk} f = g.$$

4 Ergodicity for Lower Transition Operators

In the linear case, that is, if the lower transition operator $\underline{T}$ is actually a stochastic matrix T, then under rather weak assumptions, T^n converges to a limit matrix that has identical rows, or equivalently, for all $f \in \mathcal{L}(\mathcal{X})$, $\lim_{n \to +\infty} T^n f$ exists and is a

constant function. This property of T is called ergodicity,[3] and the conditions under which it happens are well studied; see, for example, Seneta ([2006](#), Section 4.2).

For our present purposes, we are interested in a generalised version of this concept of ergodicity, which applies to lower transition operators.

Definition 2 (*Ergodic lower transition operator*) A lower transition operator $\underline{T}$ is ergodic if, for all $f \in \mathcal{L}(\mathcal{X})$, $\lim_{n \to +\infty} \underline{T}^n f$ exists and is a constant function.

Similarly, the corresponding upper transition operator $\overline{T}$ is said to be ergodic if, for all $f \in \mathcal{L}(\mathcal{X})$, $\lim_{n \to +\infty} \overline{T}^n f$ exists and is a constant function. It follows from Eq. ([5](#)) that both notions are equivalent: $\underline{T}$ is ergodic if and only if $\overline{T}$ is.

Hermans and Cooman ([2012](#)) characterised this notion of ergodicity, showing that a lower transition operator is ergodic if and only if it is regularly absorbing; see Proposition [4](#) further on. The following definition of a regularly absorbing lower transition operator is an equivalent but slightly simplified version of theirs; Lemma [21](#) in the Appendix establishes the equivalence.

Definition 3 (*Regularly absorbing lower transition operator*) A lower transition operator $\underline{T}$ is regularly absorbing if it satisfies the following two conditions:

$$\mathcal{X}_{\mathrm{RA}} := \{x \in \mathcal{X} \colon (\exists n \in \mathbb{N}) \ \min \overline{T}^n \mathbb{I}_x > 0\} \neq \emptyset$$

and

$$(\forall x \in \mathcal{X} \setminus \mathcal{X}_{\mathrm{RA}})(\exists n \in \mathbb{N}) \ \underline{T}^n \mathbb{I}_{\mathcal{X}_{\mathrm{RA}}}(x) > 0.$$

The first condition is called top class regularity, and the second condition is called top class absorption.

Proposition 4 *A lower transition operator $\underline{T}$ is ergodic if and only if it is regularly absorbing.*

If a lower transition operator satisfies Definition [3](#) with $n := 1$, we call this lower transition operator 1-step absorbing.

Definition 4 (*1-step absorbing lower transition operator*) A lower transition operator $\underline{T}$ is 1-step absorbing if it satisfies the following two conditions:

$$\mathcal{X}_{\mathrm{1A}} := \{x \in \mathcal{X} \colon \ \min \overline{T} \mathbb{I}_x > 0\} \neq \emptyset$$

and

$$(\forall x \in \mathcal{X} \setminus \mathcal{X}_{\mathrm{1A}}) \ \underline{T} \mathbb{I}_{\mathcal{X}_{\mathrm{1A}}}(x) > 0.$$

---

[3] This terminology is not universally adopted; we follow Senata ([2006](#), p. 128). Some authors use ergodicity to refer to a stronger property, which additionally requires that the identical rows of $\lim_{n \to +\infty} T^n$ consist of strictly positive elements, and which can be shown to be equivalent to the existence of some $n \in \mathbb{N}$ such that T^n consists of strictly positive elements only.

Since $\mathcal{X}_{1A}$ is clearly a subset of $\mathcal{X}_{RA}$, it follows from L6 that $\underline{T}\mathbb{I}_{\mathcal{X}_{RA}} \geq \underline{T}\mathbb{I}_{\mathcal{X}_{1A}}$, and therefore, every 1-step absorbing lower transition operator is guaranteed to be regularly absorbing as well. By combining this observation with Proposition 4, it follows that being 1-step absorbing is a sufficient condition for ergodicity. However, in general, this stronger condition of being 1-step absorbing is not necessary for ergodicity. The reason why we are nevertheless interested in this stronger property is because, as we will show further on in Sect. 8, for the particular lower transition operators $\underline{T}_t$ that are the focus of this paper, both of these properties—Definitions 3 and 4—are equivalent; see Proposition 15.

5 Lower Transition Rate Operators

Having introduced a nonlinear generalisation of a stochastic matrix, we now move on to introduce a similar generalisation of an intensity matrix—a matrix that has non-negative off-diagonal elements and rows that sum to zero. Again, the only difference is the additivity axiom, which we relax by replacing it with a superadditivity axiom.

Definition 5 (*Lower transition rate operator*) A lower transition rate operator $\underline{Q}$ is an operator $\underline{Q}$ such that for all $f, g \in \mathcal{L}(\mathcal{X})$, $\lambda \geq 0$, $\mu \in \mathbb{R}$ and $x, y \in \mathcal{X}$:

R1: $\underline{Q}(\mu) = 0$;
R2: $\underline{Q}(f + g) \geq \underline{Q}(f) + \underline{Q}(g)$;   [superadditivity]
R3: $\underline{Q}(\lambda f) = \lambda \underline{Q}(f)$;   [non-negative homogeneity]
R4: $x \neq y \Rightarrow \underline{Q}(\mathbb{I}_y)(x) \geq 0$.

The corresponding upper transition operator $\overline{Q}$ is defined by

$$\overline{Q}f := -\underline{Q}(-f) \quad \text{for all } f \in \mathcal{L}(\mathcal{X}). \tag{7}$$

As a rather straightforward consequence of this definition, a lower transition rate operator also satisfies the following properties; see the Appendix for a proof. For all $f \in \mathcal{L}(\mathcal{X})$, $\mu \in \mathbb{R}$ and $x \in \mathcal{X}$:

R5: $\underline{Q}(f) \leq \overline{Q}(f)$;
R6: $\underline{Q}(f + \mu) = \underline{Q}(f)$;
R7: $\overline{Q}(\mathbb{I}_x)(x) \leq 0$;
R8: $2\|f\| \underline{Q}(\mathbb{I}_x)(x) \leq (f(x) - \min f)\underline{Q}(\mathbb{I}_x)(x) \leq \underline{Q}(f)(x)$;
R9: $\|\underline{Q}\| \leq 2 \max_{x \in \mathcal{X}} |\underline{Q}(\mathbb{I}_x)(x)|$.

Example 2 Consider a binary state space $\mathcal{X} = \{a, b\}$, and let us adopt the notational convention that x^c denotes the complement of x, in the sense that $a^c = b$ and $b^c = a$. Also, for all $x \in \mathcal{X}$, let λ_x be some non-negative real number. As the reader may verify by checking the defining properties, the operator $\underline{Q}$ that is given by

$$\underline{Q}(f)(x) := \min\left\{0, \lambda_x\left(f(x^c) - f(x)\right)\right\} \quad \text{for all } f \in \mathcal{L}(\mathcal{X}) \text{ and } x \in \mathcal{X}$$

is then guaranteed to be a lower transition rate operator. Furthermore, by applying Eq. (7), it follows that the corresponding upper transition operator is equal to

$$\overline{Q}(f)(x) = \max\{0, \lambda_x(f(x^c) - f(x))\} \quad \text{for all } f \in \mathcal{L}(\mathcal{X}) \text{ and } x \in \mathcal{X}.$$

If every λ_x is 0, we obtain the degenerate case where $\underline{Q}$ and $\overline{Q}$ are both equal to the zero operator, in the sense that $\underline{Q}(f)(x) = \overline{Q}(f)(x) = 0$.

Lower transition rate operators are very closely related to lower transition operators: they can be derived from each other. The following two results make this explicit.

Proposition 5 *Let $\underline{Q}$ Zbe a lower transition rate operator. Then, for all $\Delta \geq 0$ such that $\Delta \|\underline{Q}\| \leq 1$, $I + \Delta\underline{Q}$ is a lower transition operator.*

Proposition 6 *Let $\underline{T}$ be a lower transition operator. Then, for all $\Delta > 0$, $\underline{Q} := 1/\Delta(\underline{T} - I)$ is a lower transition rate operator.*

Because of this connection, we can use results for lower transition operators to obtain analogous results for lower transition rate operators. The following properties can, for example, be derived from L8, L10 and L11 respectively; see the Appendix for a proof. For any sequence $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{L}(\mathcal{X})$ and all $f, g \in \mathcal{L}(\mathcal{X})$:

R10: $\lim_{n\to+\infty} f_n = f \Rightarrow \lim_{n\to+\infty} \underline{Q}f_n = \underline{Q}f$;

R11: $\|\underline{Q}f - \underline{Q}g\| \leq 2\|\underline{Q}\|\|f - g\|$;

R12: $\|\underline{Q}A - \underline{Q}B\| \leq 2\|\underline{Q}\|\|A - B\|$,

where A and B are non-negatively homogeneous operators from $\mathcal{L}(\mathcal{X})$ to $\mathcal{L}(\mathcal{X})$. Similarly, as another example, our next result follows rather easily from Proposition 2.

Proposition 7 *For any lower transition rate operator $\underline{Q}$ and any sequence $\{\underline{Q}_n\}_{n\in\mathbb{N}}$ of lower transition rate operators:*

$$\lim_{n\to+\infty} \underline{Q}_n = \underline{Q} \Leftrightarrow (\forall f \in \mathcal{L}(\mathcal{X})) \quad \lim_{n\to+\infty} \underline{Q}_n f = \underline{Q}f.$$

6 The Differential Equation of Interest

With all of the above material in place, we are now ready to introduce our main object of study: the time-dependent operator $\underline{T}_t$ that corresponds to a given lower transition rate operator.

Let $\underline{Q}$ be an arbitrary lower transition rate operator. Then, for any $t \geq 0$, we let $\underline{T}_t$ be the operator that is defined for all $f \in \mathcal{L}(\mathcal{X})$ by the differential equation

$$\frac{d}{dt}\underline{T}_t f = \underline{Q}\underline{T}_t f \quad \text{for all } t \geq 0 \tag{8}$$

and the boundary condition $\underline{T}_0 f := f$, where, at the boundary point $t = 0$, we only consider the right derivative. The validity of this definition is justified by a recent result of Škulj (2015), who showed that the above differential equation has a unique solution for all $t \geq 0$.

If $\underline{Q}$ is additive or, equivalently, if $\underline{Q}$ can be identified with an intensity matrix Q, then $\underline{T}_t$ is equal to its matrix exponential e^{Qt}. In the general case, the operator $\underline{T}_t$ can be regarded as a superadditive generalisation of the matrix exponential.

Example 3 Consider a binary state space $\mathcal{X} = \{a, b\}$ and let $\underline{Q}$ be the lower transition rate operator that was introduced in Example 2. Then, for all $f \in \mathcal{L}(\mathcal{X})$, as we explained above, the differential equation (8), with $\underline{T}_0 f = f$ as its boundary condition, has a unique solution $\underline{T}_t f$ for all $t \geq 0$. In fact, in this simple binary example, the solution even has a closed-form expression. For any $t \geq 0$ and any $x \in \mathcal{X}$, it is given by

$$\underline{T}_t(f)(x) = \begin{cases} \min f & \text{if } f(x) = \min f \\ \min f + (\max f - \min f)e^{-\lambda_x t} & \text{if } f(x) = \max f. \end{cases} \quad (9)$$

Proving that this expression is indeed a—necessarily unique—solution of Eq. (8) is a matter of straightforward verification; we leave this as an exercise.

If every λ_x is 0, or equivalently, if $\underline{Q}$ is equal to the zero operator, then $\underline{T}_t(f) = f$, which means that in this degenerate case, $\underline{T}_t$ is equal to the identity operator I.

The rest of this section presents a number of basic properties of the operator $\underline{T}_t$ and establishes some alternative characterisations for it.

First of all, as a direct consequence of its definition, we find that $\underline{T}_t$ satisfies the following semigroup property:

$$\underline{T}_{t_1 + t_2} = \underline{T}_{t_1} \underline{T}_{t_2} \quad \text{for all } t_1, t_2 \geq 0. \quad (10)$$

Secondly, as already suggested by our notation, $\underline{T}_t$ is a lower transition operator.

Proposition 8 *Let $\underline{Q}$ be a lower transition rate operator. Then, for all $t \geq 0$, $\underline{T}_t$ is a lower transition operator.*

Thirdly, as our next result establishes, we do not need to consider Eq. (8) for every $f \in \mathcal{L}(\mathcal{X})$ separately. Instead, we can apply a similar differential equation to the operator $\underline{T}_t$ itself.

Proposition 9 *Let $\underline{Q}$ be a lower transition rate operator. Then, $\underline{T}_0 = I$ and*

$$\frac{\mathrm{d}}{\mathrm{d}t}\underline{T}_t = \underline{Q}\,\underline{T}_t \quad \text{for all } t \geq 0, \quad (11)$$

where for $t = 0$, we only consider the right derivative.

At first sight, it might seem as if Eq. (11) is an immediate consequence of Eq. (8). However, as explained in Sect. 2, this is not the case. The issue is that in general, convergence with respect to the operator norm is not equivalent to pointwise convergence. For this reason, despite the fact that we know from Eq. (8) that $\frac{\mathrm{d}}{\mathrm{d}t}\underline{T}_t f$ exists for every $f \in \mathcal{L}(\mathcal{X})$, the existence of $\frac{\mathrm{d}}{\mathrm{d}t}\underline{T}_t$ still remains non-trivial and therefore requires the explicit use of Proposition 9.

Finally, $\underline{T}_t$ can also be defined directly, without any reference to a differential equation. The following simple limit expression resembles—and generalises—the well-known limit definition of a matrix exponential.

Proposition 10 *Let $\underline{Q}$ be a lower transition rate operator. Then,*

$$\underline{T}_t = \lim_{n \to +\infty} \left(I + \frac{t}{n} \underline{Q} \right)^n$$

for all $t \geq 0$.

The operator $\underline{T}_t$ also satisfies some additional properties, some of which are stated and proved in the Appendix. However, since these properties are rather technical, and because we only need them in our proofs, we have chosen not to include them in the main text. Nevertheless, some of these properties—especially those that are stated in Proposition 31 and Corollary 32—may be of independent interest to the reader.

7 Convergence of the Differential Equation's Solution

Having introduced our main object of study in the previous section, we now move on to study its limit behaviour.

As mentioned in the introduction, if $\underline{Q}$ is an intensity matrix Q, then rather remarkably, regardless of the specific intensity matrix Q that is considered, the operator $\underline{T}_t = e^{Qt}$ is guaranteed to converge to a limit (Chung 1967, Theorem II.10.1) as t goes to infinity. By analogy, for any—possibly nonlinear—lower transition rate operator $\underline{Q}$, one might suspect that the corresponding lower transition operator $\underline{T}_t$ will also converge to a limit. The aim of this section is to prove that this conjecture is indeed true.

In order to do this, our starting point is Proposition 3, which established that there is a natural number $p \in \mathbb{N}$ such that, for every lower transition operator $\underline{T}$, the powers $\underline{T}^{pk}$ converge pointwise as k goes to infinity. In general, this result can only provide us with a proof of cyclic behaviour, because p could be high. However, for our specific lower transition operator, we can loose the dependency on p. In particular, for any $\Delta \geq 0$, since the semi-group property (10) implies that $\underline{T}_{\Delta k} = \underline{T}_{\Delta}^k = \underline{T}_{\Delta/p}^{pk}$, it follows trivially from Proposition 3 that $\underline{T}_{\Delta k} = \underline{T}_{\Delta}^k$ will converge pointwise as k goes to infinity. By combining this result with Propositions 1 and 2, we obtain the following result, a formal proof of which can be found in the Appendix.

Proposition 11 *Let $\underline{Q}$ be a lower transition rate operator. Then, for any $\Delta \geq 0$, the lower transition operator $\underline{T}_{\Delta k}$ converges to a limit $\underline{T} = \lim_{k \to +\infty} \underline{T}_{\Delta k}$, and this limit is again a lower transition operator.*

For $\Delta = 0$, this result is rather trivial, because in that case, $\underline{T}_{\Delta k} = \underline{T}_0 = I$. However, for $\Delta > 0$, the result is not trivial at all. Basically, it states that along any equally spaced discretisation of the positive real line, the operator $\underline{T}_t$ is guaranteed to converge to some limiting operator $\underline{T}_\infty$. Since the discretisation step Δ can be arbitrarily small, and since $\underline{T}_t$ is continuous—because it is the solution of a differential equation—one would intuitively expect that the discretisation aspect of this result can be removed, and that the limit $\underline{T}_\infty$ does not depend on the specific way in which t approaches infinity. The following theorem confirms that this is indeed the case.

Theorem 12 *For any lower transition rate operator $\underline{Q}$, the corresponding lower transition operator $\underline{T}_t$ converges to a limit $\underline{T}_\infty = \lim_{t \to +\infty} \underline{T}_t$, and this limit is again a lower transition operator.*

Example 4 Consider a binary state space $\mathcal{X} = \{a, b\}$, let $\underline{Q}$ be the lower transition rate operator from Example 2 and, for all $f \in \mathcal{L}(\mathcal{X})$, let $\underline{T}_t f$ be the corresponding solution of Eq. (8) that was provided in Example 3. In this particular case, the limiting operator $\underline{T}_\infty$ whose existence is guaranteed by Theorem 12 can be derived explicitly. Fix any $f \in \mathcal{L}(\mathcal{X})$. If $f(x) = \min f$, then it follows from Eq. (9) that

$$\lim_{t \to +\infty} \underline{T}_t(f)(x) = \lim_{t \to +\infty} \min f = \min f = f(x).$$

Similarly, if $f(x) = \max f$, Eq. (9) implies that

$$\lim_{t \to +\infty} \underline{T}_t(f)(x) = \min f + (\max f - \min f) \lim_{t \to +\infty} e^{-\lambda_x t}$$

$$= \begin{cases} \min f & \text{if } \lambda_x > 0 \\ \max f = f(x) & \text{if } \lambda_x = 0. \end{cases}$$

Hence, in all cases, we find that

$$\underline{T}_\infty(f)(x) = \lim_{t \to +\infty} \underline{T}_t(f)(x) = \begin{cases} f(x) & \text{if } \lambda_x = 0 \\ \min f & \text{if } \lambda_x > 0 \end{cases} \quad \text{for all } x \in \mathcal{X}.$$

By comparing this expression with Eq. (6) in Example 1, we find that it corresponds to the special case that is obtained by letting $\alpha_x = 0$ if $\lambda_x = 0$ and $\alpha_x = 1$ if $\lambda_x > 0$. Hence, as predicted by Theorem 12, the limiting operator $\underline{T}_\infty$ is indeed a lower transition operator.

8 Ergodicity for Lower Transition Rate Operators

As explained in the introduction, the convergence result in Theorem 12 is particularly important in the context of imprecise continuous-time Markov chains, as it can then be interpreted to mean that every imprecise continuous-time Markov chain is guaranteed to converge to a limiting distribution. However, in general, this limiting distribution is not guaranteed to be unique, in the sense that it may depend on the initial state x of the chain. Or to rephrase it in our more abstract setting: $\underline{T}_\infty(f)(x) = \lim_{t \to +\infty} \underline{T}_t(f)(x)$ may depend on x. Whenever this is not the case, we say that $\underline{Q}$ is ergodic.

Definition 6 (*Ergodic lower transition rate operator*) A lower transition rate operator $\underline{Q}$ is ergodic if, for all $f \in \mathcal{L}(\mathcal{X})$, $\lim_{t \to +\infty} \underline{T}_t f$ exists and is a constant function.

In the linear case, that is, if $\underline{Q}$ can be identified with an intensity matrix Q, this definition requires that $\underline{T}_t f = e^{Qt} f$ converges to constant function for every $f \in \mathcal{L}(\mathcal{X})$, which is equivalent to requiring that e^{Qt} converges to a matrix that has identical rows.

Hence, for the linear case, our notion of ergodicity coincides with existing concepts of ergodicity for intensity matrices; see, for example, Tornambè ([1995], Definition 4.17).[4]

The following example illustrates our notion of ergodicity, by providing examples of ergodic as well as non-ergodic lower transition rate operators.

Example 5 Consider a binary state space $\mathcal{X} = \{a, b\}$, and let $\underline{Q}$ be the binary lower transition rate operator that we introduced in Example 2, for which we have shown in Example 4 that, for all $f \in \mathcal{L}(\mathcal{X})$,

$$\lim_{t \to +\infty} \underline{T}_t(f)(x) = \begin{cases} f(x) & \text{if } \lambda_x = 0 \\ \min f & \text{if } \lambda_x > 0 \end{cases} \quad \text{for all } all\, x \in \mathcal{X}.$$

For this example, the ergodicity of $\underline{Q}$ depends on the value of the non-negative parameters λ_x. If $\lambda_a > 0$ and $\lambda_b > 0$, then for all $f \in \mathcal{L}(\mathcal{X})$, we find that

$$\lim_{t \to +\infty} \underline{T}_t(f)(a) = \min f = \lim_{t \to +\infty} \underline{T}_t(f)(b).$$

Hence, in that case, it follows from Definition 6 that $\underline{Q}$ is ergodic. However, this is no longer true if $\lambda_a = 0$ or $\lambda_b = 0$. For example, if $\lambda_a = 0$, we find that

$$\lim_{t \to +\infty} \underline{T}_t(\mathbb{I}_a)(a) = \mathbb{I}_a(a) = 1 \neq 0 = \mathbb{I}_a(b) = \min \mathbb{I}_a = \lim_{t \to +\infty} \underline{T}_t(\mathbb{I}_a)(b),$$

and therefore, in that case, we indeed find that $\underline{Q}$ is not ergodic, because the condition in Definition 6 fails for $f = \mathbb{I}_a$. The argument for $\lambda_b = 0$ is completely analogous.

In these examples, checking whether $\underline{Q}$ is ergodic is easy because we have a simple closed-form expression for $\underline{T}_t$ and its limit $\underline{T}_\infty$, which we can then use to verify Definition 6 directly. However, for more complicated lower transition rate operators, such a direct approach will not be practical, because closed-form expressions for $\underline{T}_t$ or $\underline{T}_\infty$ will usually be difficult—if not impossible—to obtain. In order to avoid this issue, we will now develop an alternative ergodicity check, in the form of a simple necessary and sufficient condition for a lower transition rate operator $\underline{Q}$ to be ergodic.

Our first step towards finding this condition is to link the continuous-time type of ergodicity that is considered in Definition 6 to the discrete-time version that we discussed in Sect. 4. Our next result establishes that $\underline{Q}$ is ergodic in the sense of Definition 6 if and only if, for some arbitrary but fixed time $t > 0$, the operator $\underline{T}_t$ is ergodic in the sense of Definition 2.

Proposition 13 *Let $\underline{Q}$ be a lower transition rate operator and consider any $t > 0$. Then, $\underline{Q}$ is ergodic if and only if $\underline{T}_t$ is ergodic.*

---

[4] Again, as was the case for the discrete-time version that we discussed in Sect. 4, the meaning of ergodicity is not universally adopted. For example, there are also authors who use ergodicity to refer to a stronger property, which additionally requires that the identical rows of $\lim_{t \to +\infty} e^{Qt}$ consist of strictly positive elements.

At first sight—at least to us—this result is rather surprising. Since the ergodicity of $\underline{Q}$ is a property that depends on the evolution of $\underline{T}_t$ as t approaches infinity, one would not suspect such a property to be completely determined by the features of a single operator $\underline{T}_t$, on an arbitrary time point $t > 0$. Nevertheless, as the above result shows, this is indeed the case.

By combining this result with Proposition 4, we immediately obtain the following alternative characterisation of ergodicity.

Corollary 14 *Let $\underline{Q}$ be a lower transition rate operator and consider any $t > 0$. Then, $\underline{Q}$ is ergodic if and only if $\underline{T}_t$ is regularly absorbing.*

This result is clearly a good first step in obtaining a simple characterisation of ergodicity. Indeed, due to this result, instead of having to study the limit behaviour of $\underline{T}_t$ as t approaches infinity, it now suffices to restrict attention to a single time point $t > 0$ that we can choose ourselves, and to check whether for this time point t, the operator $\underline{T}_t$ is regularly absorbing. Furthermore, as our next result establishes, checking whether this particular type of lower transition operator is regularly absorbing is easier than it is for general lower transition operators: in this special case, being regularly absorbing is equivalent to being 1-step absorbing.

Proposition 15 *Let $\underline{Q}$ be a lower transition rate operator and consider any $t > 0$. Then, $\underline{T}_t$ is regularly absorbing if and only if it is 1-step absorbing.*

By combining this result with Corollary 14, we immediately obtain yet another necessary and sufficient condition for $\underline{Q}$ to be ergodic.

Corollary 16 *Let $\underline{Q}$ be a lower transition rate operator and consider any $t > 0$. Then, $\underline{Q}$ is ergodic if and only if $\underline{T}_t$ is 1-step absorbing.*

Because of this result, checking whether $\underline{Q}$ is ergodic is now reduced to checking whether $\underline{T}_t$ is 1-step absorbing, for some arbitrary—freely chosen—time point $t > 0$. Although this is already easier than studying the limit behaviour of $\underline{T}_t$ directly, it is still non-trivial. As can be seen from Definition 4, it requires us to evaluate the strict positivity of numbers that are of the form $\overline{T}_t \mathbb{I}_x(y)$ and $\underline{T}_t \mathbb{I}_A(x)$, with $x, y \in \mathcal{X}$ and $A \subseteq \mathcal{X}$. At first sight, this still seems to be a rather cumbersome task that will involve either solving the differential equation (8) or applying the limit expression in Proposition 10. However, as it turns out, this is not the case.

Indeed, as we are about to show, the strict positivity of $\overline{T}_t \mathbb{I}_x(y)$ and $\underline{T}_t \mathbb{I}_A(x)$ does not depend on the specific value of t, but only on the lower transition operator $\underline{Q}$. In order to make this specific, we introduce the following notions of upper and lower reachability.

Definition 7 (*Upper reachability*) For any $x, y \in \mathcal{X}$, we say that x is upper reachable from y, and denote this by $y \dashrightarrow x$, if there is some sequence $y = x_0, \ldots, x_n = x$ such that, for all $k \in \{1, \ldots, n\}$:

$$x_k \neq x_{k-1} \text{ and } \overline{Q}(\mathbb{I}_{x_k})(x_{k-1}) > 0.$$

Definition 8 (*Lower reachability*) For any $x \in \mathcal{X}$ and $A \subseteq \mathcal{X}$, we say that A is lower reachable from x, and denote this by $x \rightarrowtail A$, if $x \in A_n$, where $\{A_k\}_{k \in \mathbb{N}_0}$ is a non-decreasing sequence that is defined by $A_0 := A$ and

$$A_{k+1} := A_k \cup \{y \in \mathcal{X} \setminus A_k : \underline{Q}(\mathbb{I}_{A_k})(y) > 0\} \quad \text{for all } k \in \mathbb{N}_0, \tag{12}$$

and where n is the first index such that $A_n = A_{n+1}$.

Example 6 Consider a binary state space $\mathcal{X} = \{a, b\}$, and let $\underline{Q}$ and $\overline{Q}$ be the lower and upper transition rate operators of Example 2. Then, for upper reachability, for any $x \in \mathcal{X}$, it follows from Definition 7 that $x \dashrightarrow x$ and that

$$x^c \dashrightarrow x \Leftrightarrow \overline{Q}(\mathbb{I}_x)(x^c) > 0 \Leftrightarrow \max\{0, \lambda_x\} > 0 \Leftrightarrow \lambda_{x^c} > 0.$$

For lower reachability, we consider three cases. If $A = \emptyset$, then $\underline{Q}(\mathbb{I}_A)(y) = \underline{Q}(0)(y) = 0$ for all $y \in \mathcal{X}$, and therefore, we infer from Definition 8 that $A_0 = A_1 = \emptyset$, which implies that there is no $x \in \mathcal{X}$ such that $x \rightarrowtail A$. If $A = \mathcal{X}$, then since the initial sequence $A_0, \ldots, A_n$ must be increasing, we infer from $A_0 = A = \mathcal{X}$ that $n = 0$, and therefore, we find that in this case, $x \rightarrowtail A$ for every $x \in \mathcal{X}$. The final case is when $A = \{z\}$ for some $z \in \mathcal{X}$. In that case, since $\underline{Q}(\mathbb{I}_{A_0})(z^c) = \underline{Q}(\mathbb{I}_z)(z^c) = \min\{0, \lambda_{z^c}\} = 0$, we find that $A_0 = A_1 = \{z\}$, which implies that $x \rightarrowtail A$ if and only if $x = z$. Hence, in all cases, and for any $x \in \mathcal{X}$, we find that

$$x \rightarrowtail A \Leftrightarrow x \in A.$$

An important property of lower and upper reachability is that they are both easy to check. For simple lower transition rate operators, as illustrated in Example 6, some basic argumentation allows us to do this manually. For more complicated transition rate operators, we can use the automated methods that we are about to introduce; see Algorithms 1 and 2 below.

In order to check upper reachability, it suffices to draw a directed graph that has the elements of $\mathcal{X}$ as its nodes and that features an arrow from y to x if and only if $\overline{Q}(\mathbb{I}_x)(y) > 0$. Checking whether y is upper reachable from x is then clearly equivalent to checking whether it is possible to start in x and follow the arrows in the graph to reach y. This is a standard reachability problem that can be solved by means of techniques from graph theory, such as the Roy–Warshall algorithm (Roy 1959; Warshall 1962). Algorithm 1 provides an implementation of this classic algorithm, which returns a Boolean-valued function UR on $\mathcal{X} \times \mathcal{X}$ such that, for any given $x, y \in \mathcal{X}$, UR$(x, y) = \texttt{true}$ if x is upper reachable from y and UR$(x, y) = \texttt{false}$ otherwise. The time complexity is $\mathcal{O}(|\mathcal{X}|^3)$.

Lower reachability essentially requires us to construct the sequence $\{A_k\}_{k \in \mathbb{N}_0}$ up to the index n. Since it follows from the increasing nature of this initial sequence that n is bounded above by $|\mathcal{X} \setminus A|$, this too is a straightforward task. Algorithm 2 provides an implementation that, for any given set of states $A \subseteq \mathcal{X}$, returns the set of all states from which A is lower reachable. The time complexity of this algorithm is again $\mathcal{O}(|\mathcal{X}|^3)$.

The reason why we are interested in these notions of lower and upper reachability is the following two equivalences.

Algorithm 1: Checking upper reachability

Data: a lower transition rate operator $\underline{Q}$ and its upper transition rate operator $\overline{Q}$
Result: a Boolean-valued function UR on $\mathcal{X} \times \mathcal{X}$ such that, for all $x, y \in \mathcal{X}$, $\text{UR}(x, y) = \texttt{true}$ if $y \dashrightarrow x$ and $\text{UR}(x, y) = \texttt{false}$ otherwise.

1 for $x \in \mathcal{X}$ **do**
2   **for** $y \in \mathcal{X}$ **do**
3   **if** $x = y$ *or* $\overline{Q}(\mathbb{I}_x)(y) > 0$ **then** $\text{UR}(x, y) \leftarrow \texttt{true}$
4   **else** $\text{UR}(x, y) \leftarrow \texttt{false}$
5 for $z \in \mathcal{X}$ **do**
6   **for** $x \in \mathcal{X}$ **do**
7   **for** $y \in \mathcal{X}$ **do**
8   **if** $\text{UR}(x, z) = \texttt{true}$ *and* $\text{UR}(z, y) = \texttt{true}$ **then** $\text{UR}(x, y) \leftarrow \texttt{true}$
9 return UR

Algorithm 2: Checking lower reachability

Data: a lower transition rate operator $\underline{Q}$ and a set of states $A \subseteq \mathcal{X}$
Result: $B := \{y \in \mathcal{X} : y \twoheadrightarrow A\}$

1 $B \leftarrow A$
2 do
3 $S \leftarrow \emptyset$
4   **for** $y \in \mathcal{X} \setminus B$ **do**
5   **if** $\underline{Q}(\mathbb{I}_B)(y) > 0$ **then** $S \leftarrow S \cup \{y\}$
6 $B \leftarrow B \cup S$
7 while $S \neq \emptyset$
8 return B

Proposition 17 *Let $\underline{Q}$ be a lower transition rate operator. Then, for any $t > 0$ and any $x, y \in \mathcal{X}$:*

$$\overline{T}_t \mathbb{I}_x(y) > 0 \Leftrightarrow y \dashrightarrow x.$$

Proposition 18 *Let $\underline{Q}$ be a lower transition rate operator. Then, for any $t > 0$, any $x \in \mathcal{X}$ and any $A \subseteq \mathcal{X}$:*

$$\underline{T}_t \mathbb{I}_A(x) > 0 \Leftrightarrow x \twoheadrightarrow A.$$

By combining these equivalences with Definition 4 and Corollary 16, we easily obtain the following result, which is the characterisation of ergodicity that we have been after all along.

Theorem 19 *A lower transition rate operator $\underline{Q}$ is ergodic if and only if*

$$\mathcal{X}_{1A} := \{x \in \mathcal{X} : (\forall y \in \mathcal{X}) \, y \dashrightarrow x\} \neq \emptyset$$

and

$$(\forall x \in \mathcal{X} \setminus \mathcal{X}_{1A}) \, x \twoheadrightarrow \mathcal{X}_{1A}.$$

We consider this necessary and sufficient condition for the ergodicity of $\underline{Q}$ to be one of the main contributions of this paper. The reason why it is to be preferred over other necessary and sufficient conditions, such as those that are given in Corollaries 14 and 16, is because it does not require us to evaluate the operator $\underline{T}_t$. Instead, all we have to do is to solve a limited number of lower and upper reachability problems, which, as can be seen from Definitions 7 and 8, only requires us to evaluate the operator $\underline{Q}$. This is obviously preferable, because $\underline{Q}$ is directly available, whereas $\underline{T}_t$ is known only indirectly through the differential equation (8) or the limit expression in Proposition 10.

The following example illustrates how Theorem 19 can be used to check ergodicity for the simple running example in this paper.

Example 7 Consider a binary state space $\mathcal{X} = \{a, b\}$, and let $\underline{Q}$ be the lower transition rate operator of Example 2. It then follows from Example 6 that, for all $x \in \mathcal{X}$,

$$\lambda_{x^c} > 0 \Leftrightarrow (\forall y \in \mathcal{X}) \; y \xrightarrow{\cdot} x$$

and, for all $x \in \mathcal{X}$ and $A \subseteq \mathcal{X}$,

$$A = \mathcal{X} \Leftrightarrow (\forall x \in \mathcal{X} \setminus A) \; x \xrightarrow{\cdot} A.$$

Due to this last equivalence, it follows from Theorem 19 that $\underline{Q}$ is ergodic if and only if $\mathcal{X}_{1A} = \mathcal{X}$. Hence, since the first equivalence implies that $\overline{\mathcal{X}}_{1A} = \mathcal{X}$ if and only if $\lambda_a > 0$ and $\lambda_b > 0$, we conclude that for this simple example, $\underline{Q}$ is ergodic if and only if the parameters λ_a and λ_b are both strictly positive. This conclusion is identical to the conclusion of Example 5. However, in our current example, it is based directly on $\underline{Q}$, whereas in Example 5, we required explicit information about the limit behaviour of $\underline{T}_t$ as t goes to infinity.

For complicated lower transition rate operators, it can be quite cumbersome to manually check the conditions in Theorem 19. In those cases, we can use Algorithm 3

Algorithm 3: Checking ergodicity

Data: a lower transition rate operator $\underline{Q}$ and its upper transition rate operator $\overline{Q}$
Result: a Boolean ergodic, which is `true` if $\underline{Q}$ is ergodic, and `false` otherwise.

1 $\mathcal{X}_{1A} \leftarrow \emptyset$
2 run Algorithm 1 to obtain UR
3 **for** $x \in \mathcal{X}$ **do**
4 | **if** UR$(x, y) = $ `true` *for all* $y \in \mathcal{X}$ **then** $\mathcal{X}_{1A} \leftarrow \mathcal{X}_{1A} \cup \{x\}$
5 **if** $\mathcal{X}_{1A} = \emptyset$ **then** Ergodic $\leftarrow$ `false`
6 **else**
7 | run Algorithm 2 with $A = \mathcal{X}_{1A}$ to obtain B
8 | **if** $B \neq \mathcal{X}$ **then** Ergodic $\leftarrow$ `false`
9 | **else** Ergodic $\leftarrow$ `true`
10 **return** Ergodic

---

instead, which provides a simple automated method for checking these conditions. Since the time complexity of Algorithms 1 and 2 are both $\mathcal{O}(|\mathcal{X}|^3)$, the time complexity of this final algorithm is clearly $\mathcal{O}(|\mathcal{X}|^3)$ as well.

9 Conclusions

The first main conclusion of this paper is that the time-dependent lower transition operator $\underline{T}_t$ that corresponds to a lower transition rate operator $\underline{Q}$ will always converge to a limiting lower transition operator $\underline{T}_\infty$. The second main conclusion is that checking whether $\underline{Q}$ is ergodic—in the sense that $\underline{T}_\infty f$ is constant for every $f \in \mathcal{L}(\mathcal{X})$—is very simple, because it does not require any knowledge of $\underline{T}_\infty$. As can be seen from Theorem 19, it is necessary and sufficient for at least one state x to be upper reachable from every other state y, and for the set $\mathcal{X}_{1A}$ of all the states x that satisfy this condition to be lower reachable from each of the states that does not.

The reason why we consider these conclusions to be important is because, as explained in the introduction, in the context of imprecise continuous-time Markov chains (Škulj 2012, 2015; Krak and De Bock, submitted; Troffaes et al. 2015), the convergence of $\underline{T}_t$ can be interpreted to mean that every imprecise continuous-time Markov chain converges to a limiting distribution, and the ergodicity of $\underline{Q}$ can be interpreted to mean that this limiting distribution is unique, in the sense that it does not depend on the initial state of the chain.

However, of course, the existence and uniqueness of such a limiting distribution is not the only aspect of the limit behaviour of an imprecise continuous-time Markov chains that is important, and quite a lot of problems still remain unsolved. For example, in the linear case, if $\underline{Q}$ is a transition rate matrix Q that is ergodic, then the limiting distribution P_∞—see Eq. (1)—is the unique probability mass function that satisfies $P_\infty Q = 0$, and it can then be computed by solving this linear system of equations (Tornambè 1995, Theorem 4.12).[5] By analogy, it seems reasonable to expect that for a lower transition rate operator $\underline{Q}$ that is ergodic, the limiting lower expectation functional $\underline{\mathbb{E}}_\infty$ in Eq. (3) is the unique solution of a nonlinear system of equations $\underline{\mathbb{E}}_\infty \underline{Q} = 0$. However, it remains an open question whether this conjecture is indeed true. Furthermore, even if this would be the case, then still, it remains to be seen whether it is possible to solve this system efficiently.

Another interesting line of future research would be to study ergodicity from a quantitative rather than just qualitative point of view, by also taking into account the rate of convergence. For the discrete-time type of ergodicity that we discussed in Sect. 4, such a study has already been conducted in Hermans and Cooman (2012) and Škulj and Hable (2011), leading to the development of a coefficient of ergodicity that simultaneously captures both the qualitative aspect of convergence—"does it converge or not?"—and the quantitative aspect—"at which rate does it converge?". We think

---

[5] At first sight, it might seem as if Tornambè (1995, Theorem 4.12) states that P_∞ satisfies $Q P_\infty = 0$ instead of $P_\infty Q = 0$. However, this is not the case. The confusion arises because the matrix Q in Tornambè (1995) is defined as the transpose of the intensity matrix Q that is considered by us, in the sense that it has columns that sum to zero rather than rows that sum to zero.

that similar coefficients of ergodicity can also be developed for the continuous-time models that we have considered in this paper.

Finally, we would like to point out that these suggestions for future research are just the tip of the iceberg, because they focus solely on the limit behaviour of imprecise continuous-time Markov chains. Ultimately, we hope that our contributions will serve as a first step towards a further theoretic development of the general field of imprecise continuous-time Markov chains. The reason why we consider such developments to be important is because, given the success of precise continuous-time Markov chains in various fields of application (Anderson 1991), and the ever increasing demand for features such as reliability and robustness in these applications, we are convinced that imprecise continuous-time Markov chains have plenty of applied potential. Nevertheless, almost no applications have been developed so far. It seems to us that one of the main reasons for this lack of applications is a severe lack of available theoretical tools. We hope that a further theoretical development of the field of imprecise continuous-time Markov chains will allow this field to flourish, and will turn it into a full-fledged robust extension of the field of continuous-time Markov chains.

Appendix: Proofs

Proofs of Results in Sect. 2

Let A and B be two non-negatively homogeneous operators from $\mathcal{L}(\mathcal{X})$ to $\mathcal{L}(\mathcal{X})$ and consider any $f, g \in \mathcal{L}(\mathcal{X})$ and $\lambda \in \mathbb{R}$.

It is well known that the maximum norm on $\mathcal{L}(\mathcal{X})$ satisfies the defining properties of a norm: it is absolutely homogeneous ($\|\lambda f\| = |\lambda| \, \|f\|$), it is subadditive ($\|f + g\| \leq \|f\| + \|g\|$), and it separates points ($\|f\| = 0 \Rightarrow f = 0$). The induced operator norm also satisfies these properties. Firstly, it is absolutely homogeneous because the maximum norm is:

$$
\begin{aligned}
\|\lambda A\| &= \sup\{\|\lambda A f\| : f \in \mathcal{L}(\mathcal{X}), \|f\| = 1\} \\
&= \sup\{|\lambda| \, \|A f\| : f \in \mathcal{L}(\mathcal{X}), \|f\| = 1\} \\
&= |\lambda| \sup\{\|A f\| : f \in \mathcal{L}(\mathcal{X}), \|f\| = 1\} = |\lambda| \, \|A\| \, .
\end{aligned}
$$

Secondly, it is subadditive because the maximum norm is:

$$
\begin{aligned}
\|A + B\| &= \sup\{\|(A + B) f\| : f \in \mathcal{L}(\mathcal{X}), \|f\| = 1\} \\
&= \sup\{\|A f + B f\| : f \in \mathcal{L}(\mathcal{X}), \|f\| = 1\} \\
&\leq \sup\{\|A f\| + \|B f\| : f \in \mathcal{L}(\mathcal{X}), \|f\| = 1\} \\
&\leq \sup\{\|A\| + \|B\| : f \in \mathcal{L}(\mathcal{X}), \|f\| = 1\} = \|A\| + \|B\| \, .
\end{aligned}
$$

Thirdly, it separates points because the maximum norm does: if $\|A\| = 0$, then $A = 0$ because, for all $f \in \mathcal{L}(\mathcal{X})$, it follows from N1—which we will prove next—that

$$0 \leq \|Af\| \leq \|A\|\,\|f\| = 0$$

and therefore, since the maximum norm separates points, that $Af = 0$.

In order to prove N1, we consider two cases: $f = 0$ and $f \neq 0$. If $f \neq 0$, or equivalently, if $\|f\| \neq 0$, we let $g := f/\|f\|$. If $f = 0$, or equivalently, if $\|f\| = 0$, we let $g := 1$. In both cases, this guarantees that $f = \|f\|\,g$ and $\|g\| = 1$ and therefore, we find that

$$\|Af\| = \|A(\|f\|\,g)\| = \|\,\|f\|\,Ag\| = \|f\|\,\|Ag\| \leq \|f\|\,\|A\|\,,$$

where the inequality holds because $\|f\| \geq 0$ and $\|Ag\| \leq \|A\|$.

Finally, N2 follows rather easily from N1:

$$\begin{aligned}
\|AB\| &= \sup\{\|ABf\| : f \in \mathcal{L}(\mathcal{X}), \|f\| = 1\} \\
&\leq \sup\{\|A\|\,\|Bf\| : f \in \mathcal{L}(\mathcal{X}), \|f\| = 1\} \\
&= \|A\| \sup\{\|Bf\| : f \in \mathcal{L}(\mathcal{X}), \|f\| = 1\} = \|A\|\,\|B\|\,.
\end{aligned}$$

Proofs of Results in Sect. 3

Proof of L9, L10 and L11 L9 follows from Eq. (4) because we know from L4 that $\|\underline{T}f\| \leq \|f\|$ for all $f \in \mathcal{L}(\mathcal{X})$. L10 follows from L7 and L4 (in that order). L11 follows from Eq. (4) and L10. □

Proof of Proposition 1 For any $f, g \in \mathcal{L}(\mathcal{X})$, we find that

$$\underline{T}(f + g) = \lim_{n \to +\infty} \underline{T}_n(f + g) \geq \lim_{n \to +\infty} \underline{T}_n(f) + \lim_{n \to +\infty} \underline{T}_n(g) = \underline{T}(f) + \underline{T}(g),$$

where the inequality holds because L2 implies that $\underline{T}_n(f + g) \geq \underline{T}_n(f) + \underline{T}_n(g)$ for all $n \in \mathbb{N}$. Since this is true for any $f, g \in \mathcal{L}(\mathcal{X})$, $\underline{T}_\infty$ satisfies L2. The proof for L1 and L3 is completely analogous. Hence, $\underline{T}$ is a lower transition operator. □

Proof of Proposition 2 The direct implication follows trivially from N1. For the converse implication, we provide a proof by contradiction. Assume that $\lim_{n \to +\infty} \underline{T}_n f = \underline{T}f$ for all $f \in \mathcal{L}(\mathcal{X})$. Assume *ex absurdo* that the sequence $\{\underline{T}_n\}_{n \in \mathbb{N}}$ does not converge to $\underline{T}$, or equivalently, that $\limsup_{n \to +\infty} \|\underline{T}_n - \underline{T}\| > 0$. Then clearly, there is some $\epsilon > 0$ and an increasing sequence n_k, $k \in \mathbb{N}$, of natural numbers such that $\|\underline{T}_{n_k} - \underline{T}\| > \epsilon$ for all $k \in \mathbb{N}$. Furthermore, for all $k \in \mathbb{N}$, it follows from $\|\underline{T}_{n_k} - \underline{T}\| > \epsilon$ and Eq. (4) that there is some $f_k \in \mathcal{L}(\mathcal{X})$ such that $\|f_k\| = 1$ and $\|\underline{T}_{n_k} f_k - \underline{T}f_k\| > \epsilon$. Since the sequence f_k, $k \in \mathbb{N}$, is clearly bounded—because $\|f_k\| = 1$—it follows from the Bolzano–Weierstrass theorem that it has a convergent subsequence, which implies that there is some $f \in \mathcal{L}(\mathcal{X})$ and an increasing sequence

k_i, $i \in \mathbb{N}$, of natural numbers such that $\lim_{i \to +\infty} \left\| f_{k_i} - f \right\| = 0$. Furthermore, since we have assumed that $\lim_{n \to +\infty} \underline{T}_n f = \underline{T} f$, it follows that

$$\lim_{i \to +\infty} \left\| \underline{T}_{n_{k_i}} f - \underline{T} f \right\| = \lim_{n \to +\infty} \left\| \underline{T}_n f - \underline{T} f \right\| = 0.$$

Hence, since it follows from L10 that

$$
\begin{aligned}
\left\| \underline{T}_{n_{k_i}} f_{k_i} - \underline{T} f_{k_i} \right\| &= \left\| \left(\underline{T}_{n_{k_i}} f_{k_i} - \underline{T}_{n_{k_i}} f \right) + \left(\underline{T}_{n_{k_i}} f - \underline{T} f \right) + \left(\underline{T} f - \underline{T} f_{k_i} \right) \right\| \\
&\leq \left\| \underline{T}_{n_{k_i}} f_{k_i} - \underline{T}_{n_{k_i}} f \right\| + \left\| \underline{T}_{n_{k_i}} f - \underline{T} f \right\| + \left\| \underline{T} f - \underline{T} f_{k_i} \right\| \\
&\leq \left\| f_{k_i} - f \right\| + \left\| \underline{T}_{n_{k_i}} f - \underline{T} f \right\| + \left\| f_{k_i} - f \right\|,
\end{aligned}
$$

we find that

$$\lim_{i \to +\infty} \left\| \underline{T}_{n_{k_i}} f_{k_i} - \underline{T} f_{k_i} \right\| = 0.$$

Since $\left\| \underline{T}_{n_{k_i}} f_{k_i} - \underline{T} f_{k_i} \right\| > \epsilon > 0$ for all $i \in \mathbb{N}$, this is a contradiction. $\qquad \square$

Proof of Proposition 3 Since $\underline{T}$ is non-expansive with respect to the maximum norm (it satisfies L10) and has at least one recurrent point (for $f = 1$, L4 implies that $\underline{T}^k f = f$ for all $k \in \mathbb{N}$), this result is an immediate consequence of Theorem 20, with $D = \mathcal{L}(\mathcal{X})$ and $A = \underline{T}$. $\qquad \square$

Theorem 20 *For any finite-dimensional space D that is equipped with the maximum norm, there is a natural number $p \in \mathbb{N}$ such that, for any non-expansive map A from D to D that has at least one recurrent point, and any $f \in D$, the sequence $\{A^{pn} f\}_{n \in \mathbb{N}}$ converges.*

Proof This result corresponds to the second part of Sine (1990, Theorem 1). $\qquad \square$

Proofs of Results in Sect. 4

Lemma 21 *A lower transition operator $\underline{T}$ is regularly absorbing if and only if*

$$\mathcal{X}'_{\text{RA}} := \left\{ x \in \mathcal{X} : (\exists n \in \mathbb{N})(\forall k \geq n) \ \min \overline{T}^k \mathbb{I}_x > 0 \right\} \neq \emptyset$$

and

$$\left(\forall x \in \mathcal{X} \setminus \mathcal{X}'_{\text{RA}} \right) (\exists n \in \mathbb{N}) \ \overline{T}^n \mathbb{I}_{\mathcal{X} \setminus \mathcal{X}'_{\text{RA}}}(x) < 1.$$

Furthermore, the set $\mathcal{X}'_{\text{RA}}$ is equal to the set $\mathcal{X}_{\text{RA}}$ that was used in Definition 3.

Proof of Lemma 21 Consider any $x \in \mathcal{X}_{\text{RA}}$. Definition 3 then implies that there is some $n \in \mathbb{N}$ such that $\min \overline{T}^n \mathbb{I}_x > 0$, and therefore, because of L4, we know that $\overline{T}^{n+1} \mathbb{I}_x = \overline{T}(\overline{T}^n \mathbb{I}_x) \geq \min \overline{T}^n \mathbb{I}_x > 0$, which implies that $\min \overline{T}^{n+1} \mathbb{I}_x > 0$. In the same way, we also find that $\min \overline{T}^{n+2} \mathbb{I}_x > 0$ and, by continuing in this way,

that $\min \overline{T}^k \mathbb{I}_x > 0$ for all $k \geq n$. Since this holds for all $x \in \mathcal{X}_{RA}$, it follows that $\mathcal{X}_{RA} \subseteq \mathcal{X}'_{RA}$. Since $\mathcal{X}'_{RA}$ is clearly a subset of $\mathcal{X}_{RA}$, this implies that $\mathcal{X}'_{RA} = \mathcal{X}_{RA}$. Hence, trivially, $\mathcal{X}_{RA} \neq \emptyset$ if and only if $\mathcal{X}'_{RA} \neq \emptyset$. The result now follows because it holds for all $x \in \mathcal{X} \setminus \mathcal{X}'_{RA} = \mathcal{X} \setminus \mathcal{X}_{RA}$ and all $n \in \mathbb{N}$ that

$$\overline{T}^n \mathbb{I}_{\mathcal{X} \setminus \mathcal{X}'_{RA}}(x) = \overline{T}^n (1 - \mathbb{I}_{\mathcal{X}'_{RA}})(x) = 1 - \underline{T}^n(\mathbb{I}_{\mathcal{X}'_{RA}})(x) = 1 - \underline{T}^n(\mathbb{I}_{\mathcal{X}_{RA}})(x),$$

where the second equality follows from L5 and Eq. (5). $\qquad\square$

Proof of Proposition 4 Since we know from Lemma 21 that our definition of a regularly absorbing lower transition operator is equivalent to the definition in Hermans and Cooman (2012), this result is identical to Hermans and Cooman (2012, Proposition 3). $\qquad\square$

Proofs of Results in Sect. 5

Proof of R5, R6, R7, R8, and R9 R5 holds because it follows from Eq. (7), R2 and R1 that

$$\underline{Q}(f) - \overline{Q}(f) = \underline{Q}(f) + \underline{Q}(-f) \leq \underline{Q}(f - f) = \underline{Q}(0) = 0.$$

R6 holds because it follows from R1 and R2 that

$$\underline{Q}(f) = \underline{Q}(f) + \underline{Q}(\mu) \leq \underline{Q}(f + \mu) = \underline{Q}(f + \mu) + \underline{Q}(-\mu) \leq \underline{Q}(f).$$

R7 holds because it follows from Eq. (7), R6, R2, and R4—in that order—that

$$\begin{aligned}
\overline{Q}(\mathbb{I}_x)(x) &= -\underline{Q}(-\mathbb{I}_x)(x) = -\underline{Q}(1 - \mathbb{I}_x)(x) \\
&= -\underline{Q}(\textstyle\sum_{y \in \mathcal{X} \setminus \{x\}} \mathbb{I}_y)(x) \leq -\sum_{y \in \mathcal{X} \setminus \{x\}} \underline{Q}(\mathbb{I}_y)(x) \leq 0.
\end{aligned}$$

R8 holds because it follows from R6, R2, R3, R4, R7, and R5—in that order—that

$$\begin{aligned}
\underline{Q}(f)(x) &= \underline{Q}(f - \min f)(x) \geq \textstyle\sum_{y \in \mathcal{X}} \underline{Q}\big((f(y) - \min f)\mathbb{I}_y\big)(x) \\
&= \textstyle\sum_{y \in \mathcal{X}} (f(y) - \min f)\underline{Q}(\mathbb{I}_y)(x) \\
&\geq (f(x) - \min f)\underline{Q}(\mathbb{I}_x)(x) \\
&\geq (\max f - \min f)\underline{Q}(\mathbb{I}_x)(x) \geq 2\,\|f\|\,\underline{Q}(\mathbb{I}_x)(x).
\end{aligned}$$

We end by proving R9. Consider any $g \in \mathcal{L}(\mathcal{X})$ such that $\|g\| = 1$. It then follows from R8 and R7 that

$$\underline{Q}(g) \geq 2\,\|g\| \min_{x \in \mathcal{X}} \underline{Q}(\mathbb{I}_x)(x) \geq -2 \max_{x \in \mathcal{X}} \big|\underline{Q}(\mathbb{I}_x)(x)\big|.$$

Similarly, since $\|-g\| = \|g\| = 1$, we also find that $\underline{Q}(-g) \geq -2 \max_{x \in \mathcal{X}} \big|\underline{Q}(\mathbb{I}_x)(x)\big|$. By combining these two inequalities with R5 and Eq. (7), it follows that

$$-2 \max_{x \in \mathcal{X}} \big|\underline{Q}(\mathbb{I}_x)(x)\big| \leq \underline{Q}(g) \leq \overline{Q}(g) = -\underline{Q}(-g) \leq 2 \max_{x \in \mathcal{X}} \big|\underline{Q}(\mathbb{I}_x)(x)\big|,$$

which implies that $\left\| \underline{Q}(g) \right\| \leq 2 \max_{x \in \mathcal{X}} \left| \underline{Q}(\mathbb{I}_x)(x) \right|$. Since this is true for all $g \in \mathcal{L}(\mathcal{X})$ such that $\|g\| = 1$, R9 now follows from Eq. (4). $\qquad \square$

Proof of Proposition 5 L2 and L3 follow trivially from R2 and R3. We only prove L1. Consider any $f \in \mathcal{L}(\mathcal{X})$. Then,

$$
\begin{aligned}
(I + \Delta \underline{Q}) f = f + \Delta \underline{Q} f &= f + \Delta \sum_{x \in \mathcal{X}} \mathbb{I}_x \underline{Q}(f)(x) \\
&\geq f + \Delta \sum_{x \in \mathcal{X}} (f(x) - \min f) \mathbb{I}_x \underline{Q}(\mathbb{I}_x)(x) \\
&\geq f - \Delta \sum_{x \in \mathcal{X}} (f(x) - \min f) \mathbb{I}_x \left\| \underline{Q}(\mathbb{I}_x) \right\| \\
&\geq f - \Delta \sum_{x \in \mathcal{X}} (f(x) - \min f) \mathbb{I}_x \left\| \underline{Q} \right\| \\
&= f - \Delta \left\| \underline{Q} \right\| (f - \min f) \\
&= (f - \min f) \left(1 - \Delta \left\| \underline{Q} \right\| \right) + \min f \geq \min f,
\end{aligned}
$$

where the first inequality follows from R8 and the third inequality follows from Eq. (4) and the fact that $\|\mathbb{I}_x\| = 1$. $\qquad \square$

Proof of Proposition 6 Simply check each of the defining properties: R1 holds because L4 implies that $\underline{T}(\mu) = \mu$ for all $\mu \in \mathbb{R}$, R2 follows from L2, R3 follows from L3, and R4 follows from L1. $\qquad \square$

Proof of R10, R11, and R12 R10, R11, and R12 are trivial if $\underline{Q} = 0$. Therefore, we may assume that $\underline{Q} \neq 0$, which implies that $\left\| \underline{Q} \right\| > 0$. Now, let $\underline{T} := I + 1/\|\underline{Q}\| \underline{Q}$. It then follows from Proposition 5 that $\underline{T}$ is a lower transition operator. We first prove R10. If $\lim_{n \to +\infty} f_n = f$, then $\lim_{n \to +\infty} \underline{T} f_n = \underline{T} f$ because of L8. Since $\underline{Q} = \left\| \underline{Q} \right\| (\underline{T} - I)$, this implies that $\lim_{n \to +\infty} \underline{Q} f_n = \underline{Q}$. R11 holds because

$$
\begin{aligned}
\left\| \underline{Q} f - \underline{Q} g \right\| = \left\| \left\| \underline{Q} \right\| (\underline{T} f - f) - \left\| \underline{Q} \right\| (\underline{T} g - g) \right\| \\
\leq \left\| \underline{Q} \right\| \left\| \underline{T} f - \underline{T} g \right\| + \left\| \underline{Q} \right\| \| f - g \| \leq 2 \left\| \underline{Q} \right\| \| f - g \|,
\end{aligned}
$$

where the last inequality follows from L10. Similarly, R12 holds because

$$
\begin{aligned}
\left\| \underline{Q} A - \underline{Q} B \right\| = \left\| \left\| \underline{Q} \right\| (\underline{T} A - A) - \left\| \underline{Q} \right\| (\underline{T} B - B) \right\| \\
\leq \left\| \underline{Q} \right\| \left\| \underline{T} A - \underline{T} B \right\| + \left\| \underline{Q} \right\| \| A - B \| \leq 2 \left\| \underline{Q} \right\| \| A - B \|,
\end{aligned}
$$

where the last inequality follows from L11. $\qquad \square$

Proof of Proposition 7 The direct implication follows trivially from N1. We only prove the converse implication. Assume that $\lim_{n \to +\infty} \underline{Q}_n f = \underline{Q} f$ for all $f \in \mathcal{L}(\mathcal{X})$. For all $x \in \mathcal{X}$, this implies that $\lim_{n \to +\infty} \underline{Q}_n(\mathbb{I}_x)(x) = \underline{Q}(\mathbb{I}_x)(x)$, which in turn

implies that there is some $c_x > 0$ such that $|\underline{Q}(\mathbb{I}_x)(x)| < c_x$ and $|\underline{Q}_n(\mathbb{I}_x)(x)| < c_x$ for all $n \in \mathbb{N}$. Let $c := \max_{x \in \mathcal{X}} c_x$. It then follows from R9 that $||\underline{Q}|| \leq 2c$ and $||\underline{Q}_n|| \leq 2c$ for all $n \in \mathbb{N}$. Choose any $0 < \Delta \leq 1/2c$. It then follows from Proposition 5 that $\underline{T} := I + \Delta \underline{Q}$ and $\underline{T}_n := I + \Delta \underline{Q}_n$, $n \in \mathbb{N}$, are lower transition operators. Furthermore, since $\lim_{n \to +\infty} \underline{Q}_n f = \underline{Q} f$ for all $f \in \mathcal{L}(\mathcal{X})$, it follows that $\lim_{n \to +\infty} \underline{T}_n f = \underline{T} f$ for all $f \in \mathcal{L}(\mathcal{X})$. By applying Proposition 2, we now find that $\lim_{n \to +\infty} \underline{T}_n = \underline{T}$, which implies that $\lim_{n \to +\infty} \underline{Q}_n = \underline{Q}$ because

$$\left\| \underline{Q}_n - \underline{Q} \right\| = \frac{1}{\Delta} \left\| \Delta \underline{Q}_n - \Delta \underline{Q} \right\| = \frac{1}{\Delta} \left\| (I + \Delta \underline{Q}_n) - (I + \Delta \underline{Q}) \right\| = \frac{1}{\Delta} \left\| \underline{T}_n - \underline{T} \right\|.$$

$\square$

Proofs of Results in Sect. 6

Lemma 22 *Let $\underline{Q}$ be a lower transition rate operator. Then, for all $f \in \mathcal{L}(\mathcal{X})$, $\underline{T}_s f$ is continuously differentiable on $[0, \infty)$.*

Proof It follows from Eq. (8) that $\underline{T}_s f$ is continuous on $[0, \infty)$. Therefore, since $\underline{Q}$ is a continuous operator [R10], $\underline{Q} \underline{T}_s f$ is also continuous on $[0, \infty)$. Because of Eq. (8), this implies that $\underline{T}_s f$ is continuously differentiable on $[0, \infty)$. $\square$

Lemma 23 *Let $\underline{Q}$ be a lower transition rate operator, and let $\Gamma(s)$ be a continuously differentiable map from $[0, t]$ to $\mathcal{L}(\mathcal{X})$ for which $\frac{d}{ds} \Gamma(s) \geq \underline{Q} \Gamma(s)$ for all $s \in [0, t]$. Then, $\min \Gamma(t) \geq \min \Gamma(0)$.*

Proof Since $\Gamma(s)$ is continuously differentiable on $[0, t]$, it follows that for every $x \in \mathcal{X}$, $\Gamma(s)(x)$ is also continuously differentiable on $[0, t]$, which implies that it is absolutely continuous on $[0, t]$. Hence, since a minimum of a finite number of absolutely continuous functions is again absolutely continuous, we find that $\min \Gamma(s)$ is absolutely continuous on $[0, t]$, which implies—see Royden and Fitzpatrick (2010, Theorem 10, Section 6.5)—that $\min \Gamma(s)$ has a derivative $\frac{d}{ds} \min \Gamma(s)$ almost everywhere on $(0, t)$, that this derivative is Lebesgue integrable over $[0, t]$, and that

$$\min \Gamma(t) = \min \Gamma(0) + \int_0^t \left(\frac{d}{ds} \min \Gamma(s) \right) ds. \tag{13}$$

Consider now any $t^* \in (0, t)$ for which $\min \Gamma(s)$ has a derivative and consider any $x \in \mathcal{X}$ for which $\Gamma(t^*)(x) = \min \Gamma(t^*)$ [clearly, there is at least one such x]. Since $\Gamma(s)(x)$ is differentiable, $\frac{d}{ds} \Gamma(s)(x)$ exists in t^*. Assume *ex absurdo* that $\frac{d}{ds} \Gamma(s)(x)|_{s=t^*}$ is not equal to $\frac{d}{ds} \min \Gamma(s)|_{s=t^*}$ or, equivalently, that $\frac{d}{ds}(\Gamma(s)(x) - \min \Gamma(s))|_{s=t^*} \neq 0$. Then, because $\Gamma(s)(x) - \min \Gamma(s)$ is continuous [since $\Gamma(s)(x)$ and $\min \Gamma(s)$ are both (absolutely) continuous] and because $t^* \in (0, t)$ and $\Gamma(t^*)(x) - \min \Gamma(t^*) = 0$, it follows that there is some $t' \in (0, t)$ such that $\Gamma(t')(x) - \min \Gamma(t') < 0$ or, equivalently, such that $\Gamma(t')(x) < \min \Gamma(t')$. Since this is clearly a contradiction, it follows that

$$\frac{d}{ds} \Gamma(s)(x) \Big|_{s=t^*} = \frac{d}{ds} \min \Gamma(s) \Big|_{s=t^*}. \tag{14}$$

We also have that

$$\frac{d}{ds}\Gamma(s)(x)\Big|_{s=t^*} \geq \underline{Q}\big(\Gamma(t^*)\big)(x) \geq \big(\Gamma(t^*)(x) - \min \Gamma(t^*)\big)\underline{Q}(\mathbb{I}_x)(x) = 0,$$

where the second inequality follows from R8 and the last equality follows because $\Gamma(t^*)(x) = \min \Gamma(t^*)$. By combining this result with Eq. (14), we find that, for all $t^* \in (0, t)$ for which $\min \Gamma(s)$ has a derivative, $\frac{d}{ds}\min \Gamma(s)\big|_{s=t^*} \geq 0$. It therefore follows from Eq. (13) that $\min \Gamma(t) \geq \min \Gamma(0)$. □

Proof of Proposition 8 We first prove L1. Consider any $f \in \mathcal{L}(\mathcal{X})$. It then follows from Lemma 22 that $\underline{T}_s f$ is continuously differentiable on $[0, t]$. Therefore, and because of Eq. (8), we infer from Lemma 23 that $\min \underline{T}_t f \geq \min \underline{T}_0 f$. Since $\underline{T}_0 f = f$, this implies that $\min \underline{T}_t f \geq \min f$, which in turn implies that $\underline{T}_t f \geq \min f$.

Let us now prove L2. Consider any $f, g \in \mathcal{L}(\mathcal{X})$. It follows from Lemma 22 that $\underline{T}_s f, \underline{T}_s g$, and $\underline{T}_s(f + g)$ are continuously differentiable on $[0, t]$, which implies that $\Gamma(s) := \underline{T}_s(f + g) - \underline{T}_s f - \underline{T}_s g$ is continuously differentiable on $[0, t]$. Furthermore, for all $s \in [0, t]$, it follows from Eq. (8) and R2 that

$$\begin{aligned}
\frac{d}{ds}\Gamma(s) &= \frac{d}{ds}\underline{T}_s(f + g) - \frac{d}{ds}\underline{T}_s f - \frac{d}{ds}\underline{T}_s g \\
&= \underline{Q}\,\underline{T}_s(f + g) - \underline{Q}\,\underline{T}_s f - \underline{Q}\,\underline{T}_s g \\
&= \underline{Q}\big(\Gamma(s) + \underline{T}_s f + \underline{T}_s g\big) - \underline{Q}\,\underline{T}_s f - \underline{Q}\,\underline{T}_s g \geq \underline{Q}\Gamma(s).
\end{aligned}$$

Therefore, we infer from Lemma 23 that $\min \Gamma(t) \geq \min \Gamma(0)$. Since $\Gamma(0) = \underline{T}_0(f + g) - \underline{T}_0 f - \underline{T}_0 g = 0$, this implies that $\min \Gamma(t) \geq 0$, which in turn implies that $\Gamma(t) \geq 0$ or, equivalently, that $\underline{T}_t(f + g) \geq \underline{T}_t f + \underline{T}_t g$.

We end by proving L3. Consider any $f \in \mathcal{L}(\mathcal{X})$ and $\lambda \geq 0$. It then follows from Eq. (8) and R3 that

$$\frac{d}{ds}(\lambda \underline{T}_s f) = \lambda \frac{d}{ds}\underline{T}_s f = \lambda \underline{Q}\,\underline{T}_s f = \underline{Q}(\lambda \underline{T}_s f) \quad \text{for all } s \geq 0.$$

Since we also have that $\lambda \underline{T}_0 f = \lambda f = \underline{T}_0(\lambda f)$, it follows that $\lambda \underline{T}_s f$ satisfies the same differential equation and boundary condition as $\underline{T}_s(\lambda f)$. Since we know that this differential equation and boundary condition lead to a unique solution on $[0, \infty)$, it follows that $\underline{T}_t(\lambda f) = \lambda \underline{T}_t(f)$. □

Lemma 24 *Let $\underline{Q}$ be a lower transition rate operator. Then,*

$$\lim_{\Delta \to 0} \underline{T}_\Delta = I \text{ and } \lim_{\Delta \to 0} {}^1\!/\!\Delta(\underline{T}_\Delta - I) = \underline{Q}.$$

Proof For any $f \in \mathcal{L}(\mathcal{X})$, it follows from Eq. (8) that $\underline{T}_t f$ is continuous on $[0, \infty)$, which implies that $\lim_{\Delta \to 0} \underline{T}_\Delta f = \underline{T}_0 f = f$. Therefore, we infer from Proposition 2 that $\lim_{\Delta \to 0} \underline{T}_\Delta = I$, which proves the first part of this lemma. We end by proving the second part. For any $f \in \mathcal{L}(\mathcal{X})$, it follows from Eq. (8) that

$$\lim_{\Delta \to 0} {}^1\!/_\Delta (\underline{T}_\Delta - I)(f) = \lim_{\Delta \to 0} {}^1\!/_\Delta (\underline{T}_\Delta f - f)$$
$$= \lim_{\Delta \to 0} {}^1\!/_\Delta (\underline{T}_\Delta f - \underline{T}_0 f) = \underline{Q}\underline{T}_0 f = \underline{Q}f.$$

Therefore, and since, for all $\Delta > 0$, ${}^1\!/_\Delta (\underline{T}_\Delta - I)$ is a lower transition rate operator because of Proposition 6, it follows from Proposition 7 that $\lim_{\Delta \to 0} {}^1\!/_\Delta (\underline{T}_\Delta - I) = \underline{Q}$.
□

Proof of Proposition 9 Since $\underline{T}_0 f := f$ for all $f \in \mathcal{L}(\mathcal{X})$, it follows trivially that $\underline{T}_0 = I$. Consider now any $t \geq 0$. In order to prove that $\frac{d}{dt}\underline{T}_t = \underline{Q}\underline{T}_t$, it suffices to show that for all $\epsilon > 0$, there is some $\delta > 0$ such that

$$\left\| \frac{\underline{T}_s - \underline{T}_t}{s - t} - \underline{Q}\underline{T}_t \right\| < \epsilon \quad \text{for all } s \geq 0 \text{ such that } 0 < |t - s| < \delta. \quad (15)$$

So consider any $\epsilon > 0$. If $\underline{Q} = 0$, Eq. (15) is trivially true because, since I clearly satisfies Eq. (8), it follows from the unicity of the solution of Eq. (8) that $\underline{T}_t = \underline{T}_s = I$. Therefore, in the remainder of this proof, we may assume that $\underline{Q} \neq 0$, which implies that $\|\underline{Q}\| \neq 0$. It then follows from Lemma 24 that there are $\delta_1 > 0$ and $\delta_2 > 0$ such that $\|\underline{T}_q - I\| < \epsilon/4\|\underline{Q}\|$ for all $0 < q < \delta_1$ and $\|{}^1\!/_\Delta(\underline{T}_\Delta - I) - \underline{Q}\| < \epsilon/2$ for all $0 < \Delta < \delta_2$. Now define $\delta := \min\{\delta_1, \delta_2\}$ and consider any $s \geq 0$ such that $0 < |t - s| < \delta$. Let $u := \min\{s, t\}$, $\Delta := |t - s|$ and $q := t - u$, which implies that $0 \leq q \leq \Delta < \delta \leq \delta_1$ and $0 < \Delta < \delta \leq \delta_2$. If $q = 0$, then $\underline{T}_q = \underline{T}_0 = I$ and therefore $\|\underline{Q}\underline{T}_q - \underline{Q}\| = \|\underline{Q} - \underline{Q}\| = 0$. If $q > 0$, it follows from R12 and Proposition 8 that $\|\underline{Q}\underline{T}_q - \underline{Q}\| \leq 2\|\underline{Q}\|\|\underline{T}_q - I\| < 2\|\underline{Q}\|\epsilon/4\|\underline{Q}\| = \epsilon/2$. Hence, in all cases, we find that $\|\underline{Q}\underline{T}_q - \underline{Q}\| < \epsilon/2$. The result now holds because

$$\left\| \frac{\underline{T}_s - \underline{T}_t}{s - t} - \underline{Q}\underline{T}_t \right\| = \left\| \frac{\underline{T}_{\Delta+u} - \underline{T}_u}{\Delta} - \underline{Q}\underline{T}_{q+u} \right\| = \left\| \frac{\underline{T}_\Delta \underline{T}_u - \underline{T}_u}{\Delta} - \underline{Q}\underline{T}_q \underline{T}_u \right\|$$
$$\leq \left\| \frac{\underline{T}_\Delta - I}{\Delta} - \underline{Q}\underline{T}_q \right\| \|\underline{T}_u\| \leq \left\| \frac{\underline{T}_\Delta - I}{\Delta} - \underline{Q}\underline{T}_q \right\|$$
$$\leq \left\| \frac{\underline{T}_\Delta - I}{\Delta} - \underline{Q} \right\| + \left\| \underline{Q}\underline{T}_q - \underline{Q} \right\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

where the second equality follows from Eq. (10), the first inequality follows from N2 and the second inequality follows from Proposition 8 and L9. □

Proof of Proposition 10 The result is trivial if $t = 0$. In the remainder of this proof, we assume that $t > 0$. The result for $\underline{Q} = 0$ is also trivial because, since I then clearly satisfies Eq. (8), it follows from the unicity of the solution of Eq. (8) that $\underline{T}_t = I$. Therefore, in the remainder of this proof, we assume that $\underline{Q} \neq 0$, which implies that $\|\underline{Q}\| \neq 0$. We will now prove that for every $\epsilon > 0$, there is some $n \in \mathbb{N}$ such that

$$\left\| \underline{T}_t - \left(I + \frac{t}{k}\underline{Q} \right)^k \right\| < \epsilon \quad \text{for all } k \geq n.$$

So consider any $\epsilon > 0$. It then follows from Lemma 24 that there is some $\delta > 0$ such that $\|1/\Delta(\underline{T}_\Delta - I) - \underline{Q}\| < \epsilon/t$ for all $0 < \Delta < \delta$. Now choose $n \in \mathbb{N}$ such that $n > \max\{t/\delta, t\,\|\underline{Q}\|\}$ and consider any $k \geq n$. Let $\Delta := t/k \leq t/n$, which implies that $0 < \Delta < \delta$ and $\Delta\,\|\underline{Q}\| < 1$. Then

$$
\begin{aligned}
&\left\|(\underline{T}_\Delta)^k - (I + \Delta\underline{Q})^k\right\| \\
&= \left\|(\underline{T}_\Delta)^k - (\underline{T}_\Delta)^{k-1}(I + \Delta\underline{Q}) + (\underline{T}_\Delta)^{k-1}(I + \Delta\underline{Q}) - (I + \Delta\underline{Q})^k\right\| \\
&\leq \left\|(\underline{T}_\Delta)^k - (\underline{T}_\Delta)^{k-1}(I + \Delta\underline{Q})\right\| + \left\|(\underline{T}_\Delta)^{k-1}(I + \Delta\underline{Q}) - (I + \Delta\underline{Q})^k\right\| \\
&\leq \left\|\underline{T}_\Delta - (I + \Delta\underline{Q})\right\| + \left\|(\underline{T}_\Delta)^{k-1} - (I + \Delta\underline{Q})^{k-1}\right\|\,\|I + \Delta\underline{Q}\| \\
&\leq \left\|\underline{T}_\Delta - (I + \Delta\underline{Q})\right\| + \left\|(\underline{T}_\Delta)^{k-1} - (I + \Delta\underline{Q})^{k-1}\right\|,
\end{aligned}
$$

where the second inequality follows from Proposition 8 and L11 [by applying them repeatedly] and N2, and the third inequality follows from L9 and Proposition 5. By continuing in this way, we find that

$$
\left\|(\underline{T}_\Delta)^k - (I + \Delta\underline{Q})^k\right\| \leq k\left\|\underline{T}_\Delta - (I + \Delta\underline{Q})\right\|.
$$

Therefore, since

$$
\left\|\underline{T}_\Delta - (I + \Delta\underline{Q})\right\| = \Delta\left\|\frac{\underline{T}_\Delta - I}{\Delta} - \underline{Q}\right\| < \Delta\frac{\epsilon}{t} = \frac{t}{k}\frac{\epsilon}{t} = \frac{\epsilon}{k},
$$

and because it follows from Eq. (10) that $\underline{T}_t = (\underline{T}_{t/k})^k = (\underline{T}_\Delta)^k$, we find that

$$
\left\|\underline{T}_t - \left(I + \frac{t}{k}\underline{Q}\right)^k\right\| = \left\|(\underline{T}_\Delta)^k - (I + \Delta\underline{Q})^k\right\| \leq k\left\|\underline{T}_\Delta - (I + \Delta\underline{Q})\right\| < k\frac{\epsilon}{k} = \epsilon.
$$

$\square$

Proofs of Results in Sect. 7

Lemma 25 *Let $\underline{Q}$ be a lower transition rate operator. Then, $\lim_{\Delta\to 0+}\underline{T}_\Delta = I$.*

Proof Fix any $\epsilon > 0$ and consider any $\delta_1, \alpha > 0$ such that $\delta_1\left(\alpha + \|\underline{Q}\|\right) < \epsilon$. Since Proposition 9 implies that $\frac{d}{dt}\underline{T}_t\big|_{t=0} = \underline{Q}\underline{T}_0 = \underline{Q}I = \underline{Q}$, there is some $\delta_2 > 0$ such that $\|1/\Delta(\underline{T}_\Delta - I) - \underline{Q}\| \leq \alpha$ for all $0 < \Delta < \delta_2$. Hence, if we let $\delta_\epsilon := \min\{\delta_1, \delta_2\}$, then

$$
\begin{aligned}
\|\underline{T}_\Delta - I\| &= \Delta\left\|\frac{\underline{T}_\Delta - I}{\Delta}\right\| \leq \Delta\left\|\frac{\underline{T}_\Delta - I}{\Delta} - \underline{Q}\right\| \\
&+ \Delta\,\|\underline{Q}\| \leq \Delta(\alpha + \|\underline{Q}\|) \leq \delta_1(\alpha + \|\underline{Q}\|) < \epsilon
\end{aligned}
$$

for all $0 \leq \Delta < \delta_\epsilon$. Since $\epsilon > 0$ is arbitrary, this implies that $\lim_{\Delta\to 0+}\underline{T}_\Delta = I$. $\square$

Proof of Proposition 11 Let p be the natural number whose existence is guaranteed by Proposition 3, and let $t := \Delta/p$. Fix any $f \in \mathcal{L}(\mathcal{X})$. Since $\underline{T}_t$ is a lower transition operator, it follows from Proposition 3 that $\underline{T}_t^{pn} f$ converges to a limit in $\mathcal{L}(\mathcal{X})$ as n goes to infinity. Hence, since we know from Eq. (10) that $\underline{T}_t^{pn} = \underline{T}_{tpn} = \underline{T}_{\Delta n}$, it follows that $\underline{T}_{\Delta n} f$ converges to a limit in $\mathcal{L}(\mathcal{X})$ as n goes to infinity. We will denote this limit by $\underline{T} f$. By doing this for every $f \in \mathcal{L}(\mathcal{X})$, we obtain an operator $\underline{T}$. Since we know from Proposition 8 that the operators $\underline{T}_{\Delta n}, n \in \mathbb{N}$, are lower transition operators, it now follows from Proposition 1 that $\underline{T}$ is a lower transition operator. $\qquad\square$

Proof of Theorem 12 Consider any $\Delta > 0$, and let $\underline{T}_\infty := \lim_{n \to +\infty} \underline{T}_{\Delta n}$ be the lower transition operator whose existence is guaranteed by Proposition 11. We will now prove that $\underline{T}_\infty = \lim_{t \to +\infty} \underline{T}_t$, or equivalently, that

$$(\forall \epsilon > 0)\, (\exists t_\epsilon \geq 0)\, (\forall t \geq t_\epsilon)\, \left\| \underline{T}_t - \underline{T}_\infty \right\| < \epsilon.$$

To this end, fix any $\epsilon > 0$. It then follows from Lemma 25 that there is some $\delta_1 > 0$ such that $\left\| \underline{T}_\delta - I \right\| < \epsilon/2$ for all $0 < \delta < \delta_1$. Let k be any natural number such that $\Delta < \delta_1 k$, and let $\Delta^* := \Delta/k$. It then follows from Proposition 11 that the sequence $\{\underline{T}_{\Delta^* i}\}_{i \in \mathbb{N}}$ converges to a lower transition operator that, since $\{\underline{T}_{\Delta n}\}_{n \in \mathbb{N}}$ is clearly a subsequence of $\{\underline{T}_{\Delta^* i}\}_{i \in \mathbb{N}}$, is equal to $\underline{T}_\infty$. Hence, there is some $i_\epsilon \in \mathbb{N}$ such that $\left\| \underline{T}_{\Delta^* i} - \underline{T}_\infty \right\| < \epsilon/2$ for all $i \geq i_\epsilon$. Now, let $t_\epsilon := i_\epsilon \Delta^*$ and consider any $t \geq t_\epsilon$. We will prove that $\left\| \underline{T}_t - \underline{T}_\infty \right\| < \epsilon$. In order to do that, we let i be the unique natural number such that $\Delta^* i < t \leq \Delta^*(i+1)$ and let $\delta := t - \Delta^* i$. Since we then clearly have that $i \geq i_\epsilon$ and $0 < \delta \leq \Delta^* < \delta_1$, we find that indeed, as required,

$$\begin{aligned}
\left\| \underline{T}_t - \underline{T}_\infty \right\| &\leq \left\| \underline{T}_t - \underline{T}_{\Delta^* i} \right\| + \left\| \underline{T}_{\Delta^* i} - \underline{T}_\infty \right\| \\
&= \left\| \underline{T}_{\Delta^* i} \underline{T}_\delta - \underline{T}_{\Delta^* i} I \right\| + \left\| \underline{T}_{\Delta^* i} - \underline{T}_\infty \right\| \\
&\leq \left\| \underline{T}_\delta - I \right\| + \left\| \underline{T}_{\Delta^* i} - \underline{T}_\infty \right\| < \epsilon/2 + \epsilon/2 = \epsilon,
\end{aligned}$$

where the first equality follows from Eq. (10) and the fact that $\underline{T}_0 = I$—see Proposition 9—and where the second inequality follows from L11. $\qquad\square$

Proofs of Results in Sect. 8

Proof of Proposition 13 First assume that $\underline{Q}$ is ergodic. For all $f \in \mathcal{L}(\mathcal{X})$, it then follows from that $\lim_{s \to +\infty} \underline{T}_s f$ exists and is a constant function. Therefore, for all $f \in \mathcal{L}(\mathcal{X})$, it follows from Eq. (10) that

$$\lim_{n \to +\infty} \underline{T}_t^n f = \lim_{n \to +\infty} \underline{T}_{nt} f = \lim_{s \to +\infty} \underline{T}_s f$$

exists and is a constant function, which implies that $\underline{T}_t$ is ergodic.

Next, assume that $\underline{T}_t$ is ergodic. This means that, for all $f \in \mathcal{L}(\mathcal{X})$, there is some $c_f \in \mathbb{R}$ such that

$$(\forall \epsilon > 0)(\exists n \in \mathbb{N})(\forall k \geq n) \left\| \underline{T}_t^k f - c_f \right\| < \epsilon. \tag{16}$$

Consider now any $f \in \mathcal{L}(\mathcal{X})$ and any $\epsilon > 0$. It then follows from Eq. (16) that there is some $n_\epsilon \in \mathbb{N}$ such that $\left\| \underline{T}_t^{n_\epsilon} f - c_f \right\| < \epsilon$, which, because of Proposition 8 and L5, implies that $\left\| \underline{T}_t^{n_\epsilon} (f - c_f) \right\| < \epsilon$. Now, let $s_\epsilon := n_\epsilon t$. Then, for all $s \geq s_\epsilon$, we have that

$$\left\| \underline{T}_s f - c_f \right\| = \left\| \underline{T}_s (f - c_f) \right\| = \left\| \underline{T}_{s-s_\epsilon} \underline{T}_t^{n_\epsilon} (f - c_f) \right\|$$
$$\leq \left\| \underline{T}_{s-s_\epsilon} \right\| \left\| \underline{T}_t^{n_\epsilon} (f - c_f) \right\| < \epsilon,$$

where the first equality follows from Proposition 8 and L5, the second equality follows from Eq. (10), the first inequality follows from N1, and the last inequality follows from Proposition 8, L9 and the fact that $\left\| \underline{T}_t^{n_\epsilon} f - c_f \right\| < \epsilon$. Hence, we have found that for all $\epsilon > 0$, there is some $s_\epsilon > 0$ such that $\left\| \underline{T}_s f - c_f \right\| < \epsilon$ for all $s \geq s_\epsilon$. In other words: $\lim_{s \to +\infty} \underline{T}_s f = c_f$. Since this is true for all $f \in \mathcal{L}(\mathcal{X})$, it follows from Definition 6 that $\underline{Q}$ is ergodic. □

Proof of Corollary 14 Immediate consequence of Propositions 4, 8 and 13. □

Lemma 26 *Let $\underline{Q}$ be a lower transition rate operator. Consider any $f \in \mathcal{L}(\mathcal{X})$ and $x \in \mathcal{X}$ such that $f(x) > \min f$. Then, for all $t \geq 0$: $\underline{T}_t f(x) > \min f$.*

Proof Since we know from Lemma 22 that $\underline{T}_t f$ is continuously differentiable on $[0, \infty)$, we know that $r_t := \underline{T}_t f - \min f$ and therefore also $r_t(x)$ is continuously differentiable on $[0, \infty)$. Furthermore, for all $t \geq 0$, it follows from Proposition 8 and L1 that $r_t \geq 0$, which in turn implies that

$$\frac{\mathrm{d}}{\mathrm{d}t} r_t(x) = \frac{\mathrm{d}}{\mathrm{d}t} \underline{T}_t f(x) = \underline{Q}(\underline{T}_t f)(x) = \underline{Q}(r_t)(x)$$
$$\geq \sum_{y \in \mathcal{X}} r_t(y) \underline{Q}(\mathbb{I}_y)(x) \geq r_t(x) \underline{Q}(\mathbb{I}_x)(x),$$

where the second equality follows from Eq. (8), the third equality follows from R6, the first inequality follows from R2 and R3, and the last inequality follows from R4. Hence, for all $t \geq 0$, we find that $r_t(x) \geq r_0(x) e^{[\underline{Q}(\mathbb{I}_x)(x)]t}$. Since we also know that $r_0(x) = \underline{T}_0 f(x) - \min f = f(x) - \min f > 0$, this implies that for all $t \geq 0$:

$$\underline{T}_t f(x) - \min f = r_t(x) \geq r_0(x) e^{[\underline{Q}(\mathbb{I}_x)(x)]t} > 0.$$

□

Lemma 27 *Let $\underline{Q}$ be a lower transition rate operator. Consider any $f \in \mathcal{L}(\mathcal{X})$, $x \in \mathcal{X}$ and $s \geq 0$ such that $\underline{T}_s f(x) > \min f$. Then, for all $t \geq s$: $\underline{T}_t f(x) > \min f$.*

Proof Because of Eq. (10), it suffices to prove that $\underline{T}_{t-s} \underline{T}_s f(x) > \min f$. We consider two cases: $\min \underline{T}_s f > \min f$ and $\min \underline{T}_s f = \min f$; $\min \underline{T}_s f < \min f$ is not possible because of Proposition 8 and L1. If $\min \underline{T}_s f > \min f$, it follows from Proposition 8 and L1 that $\underline{T}_{t-s} \underline{T}_s f(x) \geq \min \underline{T}_s f > \min f$. If $\min \underline{T}_s f = \min f$, then $\underline{T}_s f(x) > \min \underline{T}_s f$ and therefore, because of Lemma 26, $\underline{T}_{t-s} \underline{T}_s f(x) > \min \underline{T}_s f = \min f$. □

Lemma 28 *Let $\underline{Q}$ be a lower transition rate operator. Consider any $f \in \mathcal{L}(\mathcal{X})$, $x \in \mathcal{X}$ and $t, s > 0$. Then,*

$$\underline{T}_t f(x) > \min f \Leftrightarrow \underline{T}_s f(x) > \min f.$$

Proof For any $\tau \geq 0$, let

$$\mathcal{X}_\tau := \{ y \in \mathcal{X} : \underline{T}_\tau f(y) > \min f \}. \tag{17}$$

It then follows from Lemma 27 that $\mathcal{X}_\tau$ is an increasing function of τ:

$$\tau \leq \tau' \Rightarrow \mathcal{X}_\tau \subseteq \mathcal{X}_{\tau'}. \tag{18}$$

Assume *ex absurdo* that

$$(\forall \tau' > 0)\ (\forall \mathcal{X}' \subseteq \mathcal{X})\ (\exists \tau \in (0, \tau'])\ \mathcal{X}_\tau \neq \mathcal{X}'. \tag{19}$$

Choose any $\tau_1 > 0$. Then clearly, $\mathcal{X}_{\tau_1} \subseteq \mathcal{X}$. Therefore, due to Eq. (19), we know that there is some $0 < \tau_2 < \tau_1$ such that $\mathcal{X}_{\tau_2} \neq \mathcal{X}_{\tau_1}$, which, because of Eq. (18), implies that $\mathcal{X}_{\tau_2} \subset \mathcal{X}_{\tau_1}$. Similarly, we infer that there is some $0 < \tau_3 < \tau_2$ such that $\mathcal{X}_{\tau_3} \subset \mathcal{X}_{\tau_2}$. By continuing in this way, we obtain an infinite sequence of time points $\tau_1 > \tau_2 > \tau_3 > \cdots > \tau_i > \cdots > 0$ such that $\mathcal{X} \supseteq \mathcal{X}_{\tau_1} \supset \mathcal{X}_{\tau_2} \supset \mathcal{X}_{\tau_3} \supset \cdots \supset \mathcal{X}_{\tau_i} \supset \cdots$. Since $\mathcal{X}$ is a finite set, this is a contradiction, leading us to conclude that Eq. (19) is false. This implies that there is some $\tau^* > 0$ and $\mathcal{X}^* \subseteq \mathcal{X}$ such that

$$(\forall \tau \in (0, \tau^*])\ \mathcal{X}_\tau = \mathcal{X}^*. \tag{20}$$

Fix any $\tau > \tau^*$ and choose $n \in \mathbb{N}$ high enough such that $2\tau/n \leq \tau^*$. It then follows from Eq. (20) that $\mathcal{X}_{\tau/n} = \mathcal{X}_{2\tau/n} = \mathcal{X}^*$. Furthermore, because of Proposition 8, L1 and Eq. (17), we know that $\underline{T}_{\tau/n} f(y) = \underline{T}_{2\tau/n} f(y) = \min f$ for all $y \in \mathcal{X} \backslash \mathcal{X}^*$. Therefore, we infer from Eq. (17) that there is some $\lambda > 0$ such that

$$\underline{T}_{2\tau/n} f - \min f \leq \lambda (\underline{T}_{\tau/n} f - \min f),$$

which, because of Eq. (10), Proposition 8 and L5 implies that

$$\underline{T}_{\tau/n}^2 (f - \min f) = \underline{T}_{2\tau/n}(f - \min f) = \underline{T}_{2\tau/n} f - \min f \leq \lambda (\underline{T}_{\tau/n} f - \min f).$$

Hence, it follows from Proposition 8, L5, Eq. (10), L6 and L3 that

$$\underline{T}_\tau f - \min f = \underline{T}_\tau (f - \min f) = \underline{T}_{\tau/n}^n (f - \min f) \leq \lambda^{n-1} (\underline{T}_{\tau/n} f - \min f). \tag{21}$$

Consider now any $y \in \mathcal{X} \backslash \mathcal{X}^*$. Since $\underline{T}_{\tau/n} f(y) = \min f$, it follows from Eq. (21) that $\underline{T}_\tau f(y) \leq \min f$, which in turn implies that $y \notin \mathcal{X}_\tau$. Since this holds for all $y \in \mathcal{X} \backslash \mathcal{X}^*$, we find that $\mathcal{X}_\tau \subseteq \mathcal{X}^* = \mathcal{X}_{\tau/n}$. Furthermore, since $\tau/n \leq \tau$, it follows

from Eq. (18) that $\mathcal{X}_{\tau/n} \subseteq \mathcal{X}_\tau$. Hence, we find that $\mathcal{X}_\tau = \mathcal{X}^*$. Since this is true for all $\tau > \tau^*$, it follows from Eq. (20) that

$$\mathcal{X}_\tau = \mathcal{X}^* \quad \text{for all } \tau > 0.$$

Therefore, due to Eq. (17), we find that

$$\underline{T}_t f(x) > \min f \Leftrightarrow x \in \mathcal{X}_t \Leftrightarrow x \in \mathcal{X}_s \Leftrightarrow \underline{T}_s f(x) > \min f.$$

□

Lemma 29 *Let $\underline{Q}$ be a lower transition rate operator. Consider any $f \in \mathcal{L}(\mathcal{X})$, $x \in \mathcal{X}$ and $s \geq 0$ such that $\overline{T}_s f(x) > \min f$. Then, for all $t \geq s$: $\overline{T}_t f(x) > \min f$.*

Proof Because of Eqs. (5) and (10), it suffices to prove that $\overline{T}_{t-s}\overline{T}_s f(x) > \min f$. We consider two cases: $\min \overline{T}_s f > \min f$ and $\min \overline{T}_s f = \min f$; $\min \overline{T}_s f < \min f$ is not possible because of Proposition 8 and L4. If $\min \overline{T}_s f > \min f$, it follows from Proposition 8 and L4 that $\overline{T}_{t-s}\overline{T}_s f(x) \geq \min \overline{T}_s f > \min f$. If $\min \overline{T}_s f = \min f$, then $\overline{T}_s f(x) > \min \overline{T}_s f$ and therefore, it follows from Proposition 8, L4 and Lemma 26 that $\overline{T}_{t-s}\overline{T}_s f(x) \geq \underline{T}_{t-s}\overline{T}_s f(x) > \min \overline{T}_s f = \min f$. □

Lemma 30 *Let $\underline{Q}$ be a lower transition rate operator. Consider any $f \in \mathcal{L}(\mathcal{X})$, $x \in \mathcal{X}$ and $t, s > 0$. Then,*

$$\overline{T}_t f(x) > \min f \Leftrightarrow \overline{T}_s f(x) > \min f.$$

Proof For any $\tau \geq 0$, let

$$\mathcal{X}_\tau := \{y \in \mathcal{X} : \overline{T}_\tau f(y) > \min f\}. \tag{22}$$

It then follows from Lemma 29 that $\mathcal{X}_\tau$ is an increasing function of τ:

$$\tau \leq \tau' \Rightarrow \mathcal{X}_\tau \subseteq \mathcal{X}_{\tau'}. \tag{23}$$

Using an argument that is identical to that in Lemma 28, we find that this implies that there is some $\tau^* > 0$ and $\mathcal{X}^* \subseteq \mathcal{X}$ such that

$$(\forall \tau \in (0, \tau^*]) \, \mathcal{X}_\tau = \mathcal{X}^*. \tag{24}$$

Fix any $\tau > \tau^*$ and choose $n \in \mathbb{N}$ high enough such that $2\tau/n \leq \tau^*$. It then follows from Eq. (24) that $\mathcal{X}_{\tau/n} = \mathcal{X}_{2\tau/n} = \mathcal{X}^*$. Furthermore, because of Proposition 8, L4 and Eq. (22), we know that $\overline{T}_{\tau/n} f(y) = \overline{T}_{2\tau/n} f(y) = \min f$ for all $y \in \mathcal{X}\backslash\mathcal{X}^*$. Therefore, we infer from Eq. (22) that there is some $\lambda > 0$ such that

$$\overline{T}_{2\tau/n} f - \min f \leq \lambda(\overline{T}_{\tau/n} f - \min f),$$

which, because of Eqs. (5) and (10), Proposition 8 and L5 implies that

$$\overline{T}_{\tau/n}^2(f - \min f) = \overline{T}_{2\tau/n}(f - \min f) = \overline{T}_{2\tau/n} f - \min f \leq \lambda(\overline{T}_{\tau/n} f - \min f).$$

Hence, it follows from Proposition 8, L5, Eqs. (5) and (10), L6 and L3 that

$$\overline{T}_\tau f - \min f = \overline{T}_\tau(f - \min f) = \overline{T}_{\tau/n}^n(f - \min f) \le \lambda^{n-1}(\overline{T}_{\tau/n} f - \min f). \quad (25)$$

Consider now any $y \in \mathcal{X} \setminus \mathcal{X}^*$. Since $\overline{T}_{\tau/n} f(y) = \min f$, it follows from Eq. (25) that $\overline{T}_\tau f(y) \le \min f$, which in turn implies that $y \notin \mathcal{X}_\tau$. Since this holds for all $y \in \mathcal{X} \setminus \mathcal{X}^*$, we find that $\mathcal{X}_\tau \subseteq \mathcal{X}^* = \mathcal{X}_{\tau/n}$. Furthermore, since $\tau/n \le \tau$, it follows from Eq. (23) that $\mathcal{X}_{\tau/n} \subseteq \mathcal{X}_\tau$. Hence, we find that $\mathcal{X}_\tau = \mathcal{X}^*$. Since this is true for all $\tau > \tau^*$, it follows from Eq. (24) that

$$\mathcal{X}_\tau = \mathcal{X}^* \quad \text{for all } \tau > 0.$$

Therefore, due to Eq. (22), we find that

$$\overline{T}_t f(x) > \min f \Leftrightarrow x \in \mathcal{X}_t \Leftrightarrow x \in \mathcal{X}_s \Leftrightarrow \overline{T}_s f(x) > \min f.$$

$\square$

Proposition 31 *Let $\underline{Q}$ be a lower transition rate operator. Then, for all $f \in \mathcal{L}(\mathcal{X})$, $x \in \mathcal{X}$ and $t, s > 0$:*

$$f(x) > \min f \;\Rightarrow\; \underline{T}_t f(x) > \min f \;\Leftrightarrow\; \underline{T}_s f(x) > \min f;$$
$$f(x) < \max f \;\Rightarrow\; \underline{T}_t f(x) < \max f \;\Leftrightarrow\; \underline{T}_s f(x) < \max f;$$
$$f(x) > \min f \;\Rightarrow\; \overline{T}_t f(x) > \min f \;\Leftrightarrow\; \overline{T}_s f(x) > \min f;$$
$$f(x) < \max f \;\Rightarrow\; \overline{T}_t f(x) < \max f \;\Leftrightarrow\; \overline{T}_s f(x) < \max f;$$

Proof The first implication $[f(x) > \min f \Rightarrow \underline{T}_t f(x) > \min f]$ follows from Lemma 26, and the first equivalence $[\underline{T}_t f(x) > \min f \Leftrightarrow \underline{T}_s f(x) > \min f]$ follows from Lemma 28. Since $\overline{T}_0 f = f$, the third implication $[f(x) > \min f \Rightarrow \overline{T}_t f(x)]$ follows from Lemma 29. The third equivalence $[\overline{T}_t f(x) > \min f \Leftrightarrow \overline{T}_s f(x) > \min f]$ follows from Lemma 30. The rest of the result now follows directly because we know from Eq. (5) that $\overline{T}_t f(x) = -\underline{T}_t(-f)(x)$, $\overline{T}_s f(x) = -\underline{T}_s(-f)(x)$ and $\max f = -\min(-f)$. $\square$

Corollary 32 *Let $\underline{Q}$ be a lower transition rate operator. Then, for all $A \subseteq \mathcal{X}$, all $x \in \mathcal{X}$ and all $t, s > 0$:*

$$x \in A \;\Rightarrow\; \underline{T}_t \mathbb{I}_A(x) > 0 \;\Leftrightarrow\; \underline{T}_s \mathbb{I}_A(x) > 0;$$
$$x \notin A \;\Rightarrow\; \underline{T}_t \mathbb{I}_A(x) < 1 \;\Leftrightarrow\; \underline{T}_s \mathbb{I}_A(x) < 1;$$
$$x \in A \;\Rightarrow\; \overline{T}_t \mathbb{I}_A(x) > 0 \;\Leftrightarrow\; \overline{T}_s \mathbb{I}_A(x) > 0;$$
$$x \notin A \;\Rightarrow\; \overline{T}_t \mathbb{I}_A(x) < 1 \;\Leftrightarrow\; \overline{T}_s \mathbb{I}_A(x) < 1.$$

Proof If $A = \emptyset$ or $A = \mathcal{X}$, the result follows trivially from Proposition 8 and L4. In all other cases, the result follows directly from Proposition 31, with $f = \mathbb{I}_A$. $\square$

Proof of Proposition 15 If $t = 0$, we know from Proposition 9 that $\underline{T}_t = I$, which implies that $\underline{T}_t^n = I = \underline{T}_t$ and $\overline{T}_t^n = I = \overline{T}_t$ for all $n \in \mathbb{N}$. In that case, Definitions 3 and 4 are trivially equal. If $t > 0$, then since we know from Eqs. (10) and (5) that $\underline{T}_t^n = \underline{T}_{nt}$ and $\overline{T}_t^n = \overline{T}_{nt}$ for all $n \in \mathbb{N}$, the equivalence of Definitions 3 and 4 follows directly from Corollary 32. $\qquad\square$

Proof of Corollary 16 This result is a trivial consequence of Corollary 14 and Proposition 15. $\qquad\square$

Proof of Proposition 17 First assume that $\overline{T}_t \mathbb{I}_x(y) > 0$. It then follows from Proposition 10 and Eqs. (5) and (7) that there is some $n \in \mathbb{N}$ such that $n \geq t \left\| \underline{Q} \right\|$ and

$$\left(\left(I + \frac{t}{n} \overline{Q} \right)^n \mathbb{I}_x \right)(y) > 0. \tag{26}$$

Let $\Delta := t/n \geq 0$ and define $\underline{T}_* := I + \Delta \underline{Q}$. Since $n \geq t \left\| \underline{Q} \right\|$ implies that $\Delta \left\| \underline{Q} \right\| \leq 1$, it then follows from Proposition 5 that $\underline{T}_*$ is a lower transition operator. Therefore, for all $z \in \mathcal{X}$ and $w \in \mathcal{X}$, it follows from L4 that $c(w, z) := \left(\overline{T}_* \mathbb{I}_z \right)(w) \geq 0$. For all $z \in \mathcal{X}$, we now have that

$$\overline{T}_* \mathbb{I}_z = \sum_{w \in \mathcal{X}} \mathbb{I}_w \cdot \left(\overline{T}_* \mathbb{I}_z \right)(w) = \sum_{w \in \mathcal{X}} c(w, z) \mathbb{I}_w.$$

Hence, for all $x_n \in \mathcal{X}$, it follows from Eq. (5), L2 and L6 that

$$\overline{T}_*^n \mathbb{I}_{x_n} = \overline{T}_*^{n-1} \overline{T}_* \mathbb{I}_{x_n} = \overline{T}_*^{n-1} \sum_{x_{n-1} \in \mathcal{X}} c(x_{n-1}, x_n) \mathbb{I}_{x_{n-1}}$$
$$\leq \sum_{x_{n-1} \in \mathcal{X}} c(x_{n-1}, x_n) \overline{T}_*^{n-1} \mathbb{I}_{x_{n-1}}$$

and, by continuing in this way, that

$$\overline{T}_*^n \mathbb{I}_{x_n} \leq \sum_{x_{n-1} \in \mathcal{X}} c(x_{n-1}, x_n) \sum_{x_{n-2} \in \mathcal{X}} c(x_{n-2}, x_{n-1}) \cdots \sum_{x_1 \in \mathcal{X}} c(x_1, x_2) \overline{T}_* \mathbb{I}_{x_1}.$$

Therefore, for all $x_n \in \mathcal{X}$ and $x_0 \in \mathcal{X}$, we find that

$$\left(\overline{T}_*^n \mathbb{I}_{x_n} \right)(x_0) \leq \sum_{x_{n-1} \in \mathcal{X}} c(x_{n-1}, x_n) \sum_{x_{n-2} \in \mathcal{X}} c(x_{n-2}, x_{n-1}) \cdots \sum_{x_1 \in \mathcal{X}} c(x_1, x_2) c(x_0, x_1).$$

Hence, if we let $x_0 := y$ and $x_n := x$, it follows from Eq. (26) that

$$\sum_{x_{n-1} \in \mathcal{X}} c(x_{n-1}, x_n) \sum_{x_{n-2} \in \mathcal{X}} c(x_{n-2}, x_{n-1}) \cdots \sum_{x_1 \in \mathcal{X}} c(x_1, x_2) c(x_0, x_1) > 0.$$

This implies that there is some sequence $y = x_0, x_1, \ldots, x_n = x$ such that

$$c(x_{n-1}, x_n)c(x_{n-2}, x_{n-1}) \cdots c(x_1, x_2)c(x_0, x_1) > 0.$$

Since each of the factors in this product is non-negative, it follows that $c(x_{k-1}, x_k) > 0$ for all $k \in \{1, \ldots, n\}$. Therefore, for any $k \in \{1, \ldots, n\}$ such that $x_k \neq x_{k-1}$, it follows that

$$\begin{aligned}
\overline{Q}(\mathbb{I}_{x_k})(x_{k-1}) &= \frac{1}{\Delta}\left(\mathbb{I}_{x_k}(x_{k-1}) + \Delta\overline{Q}(\mathbb{I}_{x_k})(x_{k-1})\right) \\
&= \frac{1}{\Delta}\left((\mathbb{I}_{x_k} + \Delta\overline{Q}(\mathbb{I}_{x_k}))(x_{k-1})\right) \\
&= \frac{1}{\Delta}\left(((I + \Delta\overline{Q})\mathbb{I}_{x_k})(x_{k-1})\right) \\
&= \frac{1}{\Delta}\left((\overline{T}_*\mathbb{I}_{x_k})(x_{k-1})\right) = \frac{1}{\Delta}c(x_{k-1}, x_k) > 0.
\end{aligned}$$

If $x_k \neq x_{k-1}$ for all $k \in \{1, \ldots, n\}$, this implies that x is upper reachable from y. Otherwise, let $x'_0, \ldots, x'_m$ be a new sequence, obtained by removing from $x_0, \ldots, x_n$ those elements x_k for which $x_k = x_{k-1}$; $n - m$ is the number of elements that is removed. Then, $x'_0 = y$, $x'_m = x$ and, for all $k \in \{1, \ldots, m\}$, we have that $x'_k \neq x'_{k-1}$ and $\overline{Q}(\mathbb{I}_{x'_k})(x'_{k-1}) > 0$. Therefore, x is upper reachable from y.

Conversely, assume that x is upper reachable from y, meaning that there is some sequence $y = x_0, x_1, \ldots, x_n = x$ such that, for all $k \in \{1, \ldots, n\}$, $x_k \neq x_{k-1}$ and $\overline{Q}\mathbb{I}_{x_k}(x_{k-1}) > 0$. If $n = 0$, then $x = y$, and therefore, it follows from Corollary 32 that $\overline{T}_t\mathbb{I}_x(y) > 0$. Hence, for the remainder of this proof, we may assume that $n \geq 1$. Fix any $k \in \{1, \ldots, n\}$. We then have that $\overline{T}_0\mathbb{I}_{x_k}(x_{k-1}) = \mathbb{I}_{x_k}(x_{k-1}) = 0$ and

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\overline{T}_s\mathbb{I}_{x_k}(x_{k-1})\Big|_{s=0} &= -\frac{\mathrm{d}}{\mathrm{d}s}\underline{T}_s(-\mathbb{I}_{x_k})(x_{k-1})\Big|_{s=0} \\
&= -\underline{Q}(\underline{T}_0(-\mathbb{I}_{x_k}))(x_{k-1}) \\
&= -\underline{Q}(-\mathbb{I}_{x_k})(x_{k-1}) = \overline{Q}(\mathbb{I}_{x_k})(x_{k-1}) > 0,
\end{aligned}$$

where the first equality follows from Eq. (5), the second equality follows from Eq. (8), and the last equality follows from Eq. (7). Therefore, there is some $\epsilon_k > 0$ such that $\overline{T}_{\epsilon_k}\mathbb{I}_{x_k}(x_{k-1}) > 0$. Consequently, if we let $c_k := \overline{T}_{\epsilon_k}\mathbb{I}_{x_k}(x_{k-1}) > 0$, then because it follows from Proposition 8 and L4 that $\overline{T}_{\epsilon_k}\mathbb{I}_{x_k} \geq 0$, we have that $\overline{T}_{\epsilon_k}\mathbb{I}_{x_k} \geq c_k\mathbb{I}_{x_{k-1}}$. Let $\epsilon := \sum_{k=1}^n \epsilon_k > 0$. Then,

$$\overline{T}_\epsilon\mathbb{I}_{x_n} = \overline{T}_{\epsilon_1}\cdots\overline{T}_{\epsilon_{n-1}}\overline{T}_{\epsilon_n}\mathbb{I}_{x_n} \geq c_n\overline{T}_{\epsilon_1}\cdots\overline{T}_{\epsilon_{n-1}}\mathbb{I}_{x_{n-1}} \geq \cdots \geq \left(\prod_{k=1}^n c_k\right)\mathbb{I}_{x_0},$$

where the equality follows from Eqs. (5) and (10) and where the inequalities follow from Proposition 8, L6, L3, and Eq. (5). Therefore, we find that

$$\overline{T}_\epsilon \mathbb{I}_x(y) = \overline{T}_\epsilon \mathbb{I}_{x_n}(x_0) \geq \prod_{k=1}^n c_k \mathbb{I}_{x_0}(x_0) = \prod_{k=1}^n c_k > 0,$$

which implies that $\overline{T}_t \mathbb{I}_x(y) > 0$ because of Corollary 32. □

Proof of Proposition 18 Let $\{A_k\}_{k \in \mathbb{N}_0}$ and n be defined as in Definition 8. We need to prove that $x \in A_n$ if and only if $\underline{T}_t \mathbb{I}_A(x) > 0$.

First assume that $\underline{T}_t \mathbb{I}_A(x) > 0$. It then follows from Proposition 10 that there is some $m \in \mathbb{N}$ such that $m \geq n$, $m \geq t \, \|\underline{Q}\|$ and

$$\left(\left(I + \frac{t}{m} \underline{Q} \right)^m \mathbb{I}_A \right)(x) > 0. \tag{27}$$

Let $\Delta := t/m \geq 0$ and define $\underline{T}_* := I + \Delta \underline{Q}$. Since $m \geq t \, \|\underline{Q}\|$ implies that $\Delta \, \|\underline{Q}\| \leq 1$, it then follows from Proposition 5 that $\underline{T}_*$ is a lower transition operator. Consider any $k \in \mathbb{N}_0$ such that $k \leq m$ and any $y \in \mathcal{X} \setminus A_{k+1}$. Since $A_k \subseteq A_{k+1}$, this implies that $y \notin A_k$. Assume *ex absurdo* that $\underline{Q}(\mathbb{I}_{A_k})(y) > 0$. It then follows from Eq. (12) that $y \in A_{k+1}$, a contradiction. Hence, we find that $\underline{Q}(\mathbb{I}_{A_k})(y) \leq 0$. Since $y \notin A_k$, it follows from R2 and R4 that $\underline{Q}(\mathbb{I}_{A_k})(y) \geq \sum_{z \in A_k} \underline{Q}(\mathbb{I}_z)(y) \geq 0$. Hence, we infer that $\underline{Q}(\mathbb{I}_{A_k})(y) = 0$. Furthermore, since $y \notin A_k$, we also have that $\mathbb{I}_{A_k}(y) = 0$. Hence, we find that $(\underline{T}_* \mathbb{I}_{A_k})(y) = ((I + \Delta \underline{Q}) \mathbb{I}_{A_k})(y) = 0$. Since this holds for all $y \in \mathcal{X} \setminus A_{k+1}$, there is some $c_k > 0$ such that $\underline{T}_* \mathbb{I}_{A_k} \leq c_k \mathbb{I}_{A_{k+1}}$. Due to L3 and L6, this implies that $\underline{T}_*^{m-k} \mathbb{I}_{A_k} \leq c_k \underline{T}_*^{m-k-1} \mathbb{I}_{A_{k+1}}$. Since this holds for all $k \in \mathbb{N}_0$ such that $k \leq m$, we find that

$$\underline{T}_*^m \mathbb{I}_{A_0} \leq c_0 \underline{T}_*^{m-1} \mathbb{I}_{A_1} \leq c_0 c_1 \underline{T}_*^{m-2} \mathbb{I}_{A_2} \leq \cdots \leq c_0 c_1 \cdots c_{m-1} \mathbb{I}_{A_m}.$$

Therefore, since $A_0 = A$, it follows from Eq. (27) that $x \in A_m$. Since $A_n = A_{n+1}$, it follows from Eq. (12) that $A_r = A_n$ for all $r \geq n$ and therefore, in particular, that $A_m = A_n$. Since $x \in A_m$, this implies that $x \in A_n$.

Conversely, assume that $x \in A_n$. If $n = 0$, then $A_n = A_0 = A$ and therefore $x \in A$, which, due to Corollary 32, implies that $\underline{T}_t \mathbb{I}_A(x) > 0$. Therefore, for the remainder of this proof, we may assume that $n \geq 1$. Fix any $k \in \{0, \ldots, n-1\}$. Consider any $y \in A_{k+1} \setminus A_k$. Then, $\underline{T}_0 \mathbb{I}_{A_k}(y) = \mathbb{I}_{A_k}(y) = 0$ and

$$\left. \frac{d}{ds} \underline{T}_s \mathbb{I}_{A_k}(y) \right|_{s=0} = \underline{Q}(\underline{T}_0(\mathbb{I}_{A_k}))(y) = \underline{Q}(\mathbb{I}_{A_k})(y) > 0,$$

where the first equality follows from Eq. (8) and the inequality follows from Eq. (12). Therefore, there is some $\epsilon_{k,y} > 0$ such that $\underline{T}_s \mathbb{I}_{A_k}(y) > 0$ for all $s \in (0, \epsilon_{k,y}]$. Hence, if we let $\epsilon_k := \min_{y \in A_{k+1} \setminus A_k} \epsilon_{k,y}$, then $\underline{T}_{\epsilon_k} \mathbb{I}_{A_k}(y) > 0$ for all $y \in A_{k+1} \setminus A_k$. For all $y \in A_k$, it follows from Corollary 32 that $\underline{T}_{\epsilon_k} \mathbb{I}_{A_k}(y) > 0$. Hence, in summary, we have that $\underline{T}_{\epsilon_k} \mathbb{I}_{A_k}(y) > 0$ for all $y \in A_{k+1}$. Since we know from Proposition 8 and L1 that $\underline{T}_{\epsilon_k} \mathbb{I}_{A_k} \geq 0$, this implies that there is some $c_k > 0$ such that $\underline{T}_{\epsilon_k} \mathbb{I}_{A_k} \geq c_k \mathbb{I}_{A_{k+1}}$. Let $\epsilon := \sum_{k=0}^{n-1} \epsilon_k$. Then,

$$\underline{T}_\epsilon \mathbb{I}_{A_0} = \underline{T}_{\epsilon_{n-1}} \cdots \underline{T}_{\epsilon_1} \underline{T}_{\epsilon_0} \mathbb{I}_{A_0} \geq c_0 \underline{T}_{\epsilon_{n-1}} \cdots \underline{T}_{\epsilon_1} \mathbb{I}_{A_1} \geq \cdots \geq \left(\prod_{k=0}^{n-1} c_k \right) \mathbb{I}_{A_n},$$

where the equality follows from Eq. (10) and the inequalities follow from Proposition 8, L3 and L6. Therefore, we find that

$$\underline{T}_\epsilon \mathbb{I}_A(x) = \underline{T}_\epsilon \mathbb{I}_{A_0}(x) \geq \left(\prod_{k=0}^{n-1} c_k \right) \mathbb{I}_{A_n}(x) = \prod_{k=0}^{n-1} c_k > 0,$$

which implies that $\underline{T}_t \mathbb{I}_A(x) > 0$ because of Corollary 32. □

Proof of Theorem 19 Fix any $t > 0$. It then follows from Corollary 16 that $\underline{Q}$ is ergodic if and only if $\underline{T}_t$ is 1-step absorbing or, equivalently, because of Definition 4, if

$$\mathcal{X}_{1A} := \{x \in \mathcal{X} \colon \min \overline{T}_t \mathbb{I}_x > 0\} \neq \emptyset \text{ and } (\forall x \in \mathcal{X} \backslash \mathcal{X}_{1A}) \ \underline{T}_t \mathbb{I}_{\mathcal{X}_{1A}}(x) > 0.$$

The result now follows immediately because we know from Proposition 17 that, for all $x \in \mathcal{X}$,

$$\min \overline{T}_t \mathbb{I}_x > 0 \ \Leftrightarrow \ (\forall y \in \mathcal{X}) \ \overline{T}_t \mathbb{I}_x(y) > 0 \ \Leftrightarrow \ (\forall y \in \mathcal{X}) \ y \dashrightarrow x,$$

and because we know from Proposition 18 that, for all $x \in \mathcal{X} \backslash \mathcal{X}_{1A}$, $\underline{T}_t \mathbb{I}_{\mathcal{X}_{1A}}(x) > 0$ if and only if $x \dashrightarrow \mathcal{X}_{1A}$. □

References

Anderson, W.J.: Continuous-Time Markov Chains, An Applications-Oriented Approach, Springer Series in Statistics. Springer, New York (1991)

Chung, K.L.: Markov Chains with Stationary Transition Probabilities. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer, Berlin (1967)

Galdino, S.: Interval continuous-time Markov chains simulation. In: Proceedings of the 1013 International Conference on Fuzzy Theory and Its Applications, pp. 273–278 (2013)

Guo, X., Hernández-Lerma, O.: Continuous-time controlled Markov chains. Ann. Appl. Probab. **13**(1), 363–388 (2003)

Guo, X., Hernández-Lerma, O.: Continuous-Time Markov Decision Processes: Theory and Applications, Stochastic Modelling and Applied Probability, vol. 62. Springer, Berlin (2009)

Hermans, F., de Cooman, G.: Characterisation of ergodic upper transition operators. Int. J. Approx. Reason. **53**(4), 573–583 (2012)

Krak, T., De Bock, J.: Imprecise continuous-time Markov chains (submitted)

Roy, B.: Transitivité et connexité. C. R. Acad. Sci. Paris **249**, 216–218 (1959)

Royden, H.L., Fitzpatrick, P.M.: Real Analysis, 4th edn. Prentice Hall, Englewood Cliffs (2010)

Seneta, E.: Non-negative Matrices and Markov Chains. Springer, New York (2006)

Sine, R.: A nonlinear Perron–Frobenius theorem. Proc. Am. Math. Soc. **109**(2), 331–336 (1990)

Škulj, D.: Interval matrix differential equations (2012). arXiv:1204.0467

Škulj, D.: Efficient computation of the bounds of continuous time imprecise Markov chains. Appl. Math. Comput. **250**(C), 165–180 (2015)

Škulj, D., Hable, R.: Coefficients of ergodicity for Markov chains with uncertain parameters. Metrika **76**(1), 107–133 (2011)

Tornambè, A.: Discrete-Event System Theory: An Introduction. World Scientific, Singapore (1995)

Troffaes, M.C.M., de Cooman, G.: Lower Previsions. Wiley, London (2014)

Troffaes, M.C.M., Gledhill, J., Škulj, D., Blake, S.: Using imprecise continuous time Markov chains for assessing the reliability of power networks with common cause failure and non-immediate repair. In: ISIPTA '15: Proceedings of the Ninth International Symposium on Imprecise Probability: Theories and Applications, pp. 287–294 (2015)

Walley, P.: Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London (1991)

Warshall, S.: A theorem on boolean matrices. J. ACM **9**(1), 11–12 (1962)