NONLINEAR MARKOV CHAINS WITH FINITE STATE SPACE: INVARIANT DISTRIBUTIONS AND LONG-TERM BEHAVIOUR

BERENICE ANNE NEUMANN 🄳,* *University of Trier*

Abstract

Nonlinear Markov chains with finite state space were introduced by Kolokoltsov (*Nonlinear Markov Processes and Kinetic Equations*, 2010). The characteristic property of these processes is that the transition probabilities depend not only on the state, but also on the distribution of the process. Here we provide first results regarding their invariant distributions and long-term behaviour: we show that under a continuity assumption an invariant distribution exists and provide a sufficient criterion for the uniqueness of the invariant distribution. Moreover, we present examples of peculiar limit behaviour that cannot occur for classical linear Markov chains. Finally, we present for the case of small state spaces sufficient (and easy-to-verify) criteria for the ergodicity of the process.

Keywords: Nonlinear Markov chain; invariant distribution; long-term behaviour; ergodicity

2020 Mathematics Subject Classification: Primary 60G65

1. Introduction

Nonlinear Markov processes are a particular class of stochastic processes where the transition probabilities depend not only on the state, but also on the distribution of the process. McKean [13] introduced processes of this type to tackle mechanical transport problems. Thereafter they have been studied by several authors (see the monographs of Kolokoltsov [10] and Sznitman [21]). Recently, the close connection to continuous-time mean-field games has led to significant progress in the analysis of McKean–Vlasov stochastic differential equations, in particular the control of these systems (see for example [5, 18]).

In this paper, we consider a special class of these processes—namely, nonlinear Markov chains in continuous time with a finite state space—and provide first insights regarding their long-term behaviour. Nonlinear Markov chains with finite state space arise naturally, in particular in evolutionary biology, epidemiology, and game theory. Specifically, replicator dynamics, several infection models, and also the dynamics of learning procedures in game theory are nonlinear Markov chains [10]. Moreover, the population dynamics in mean-field games with finite state and action space are also nonlinear Markov chains [15].

Note that the sense in which we use the term 'nonlinear' here is different from the sense in which it occurs in the discussion of general (or nonlinear) birth–death processes, as for example in [3]. Indeed, in the latter context the term 'nonlinear' means that the transition rates depend on the current state not linearly but in a more general way (see [17] and the references

---

therein for an overview). However, in our context the term 'nonlinear' refers to the Markov semigroup $\Phi^t(m_0)$, which is linear for classical Markov chains and nonlinear in our setting, so that the transition probabilities can now depend on the current distribution of the process.

The main focus of this paper lies in the characterization of the long-term behaviour of these processes. We show that an invariant distribution always exists and provide a sufficient criterion for the uniqueness of this invariant distribution. We then turn to the long-term behaviour, for which we first illustrate by two examples that the limit behaviour is much more complex than for classical Markov chains. More precisely, we show that the marginal distributions of a nonlinear Markov chain may be periodic and that irreducibility of the generator does not necessarily imply ergodicity. Then we provide easy-to-verify sufficient criteria for ergodicity for small state spaces (two or three states). All of the conditions that we propose are simple and rely only on the shape of the nonlinear generator, not on the shape of the transition probabilities.

The long-term behaviour of general nonlinear Markov chains in continuous time with a finite state space has not been analysed before. The closest contributions in the literature are results for specific continuous-time Markov chains associated to pressure and resistance games [12], as well as ergodicity criteria for nonlinear Markov processes in discrete time [4, 20]. These latter criteria are a generalization of Dobrushin's ergodicity condition, and the proofs rely crucially on the sequential nature of the problem.

The rest of the paper is structured as follows. In Section 2 we review the relevant definitions and notation. In Section 3 we present the results on existence and uniqueness of the invariant distribution. In Section 4 we provide examples of limit behaviour that cannot arise in the context of classical Markov chains. In Section 5 we present the ergodicity results for small state spaces. Appendix A contains the proofs of two technical results.

2. Continuous-time nonlinear Markov chains with finite state space

This section gives a short overview of the relevant definitions, notation, and preliminary facts regarding nonlinear Markov chains. For more details about these processes we refer the reader to [10, Chapter 1]. Moreover, this section introduces the relevant notions for characterizing the long-term behaviour of these processes.

Let $\mathcal{S} = \{1, \dots, S\}$ be the state space of the nonlinear Markov chain, and denote by $\mathcal{P}(\mathcal{S})$ the probability simplex over $\mathcal{S}$. A nonlinear Markov chain is characterized by a *continuous family of nonlinear transition probabilities* $P(t, m) = (P_{ij}(t, m))_{i,j \in \mathcal{S}}$, which is a family of stochastic matrices that depends continuously on $t \geq 0$ and $m \in \mathcal{P}(\mathcal{S})$, such that the nonlinear Chapman–Kolmogorov equation

$$\sum_{i \in \mathcal{S}} m_i P_{ij}(t + s, m) = \sum_{i,k \in \mathcal{S}} m_i P_{ik}(t, m) P_{kj}\left(s, \sum_{l \in \mathcal{S}} m_l P_l(t, m)\right)$$

is satisfied. As usual, $P_{ij}(t, m_0)$ is interpreted as the probability that the process is in state j at time t given that the initial state was i and the initial distribution of the process was m_0. Such a family yields a nonlinear Markov semigroup $(\Phi^t(\cdot))_{t \geq 0}$ of continuous transformations of $\mathcal{P}(\mathcal{S})$ via

$$\Phi_j^t(m) = \sum_{i \in \mathcal{S}} m_i P_{ij}(t, m) \quad \text{for all } t \geq 0, \; m \in \mathcal{P}(\mathcal{S}), \; j \in \mathcal{S}.$$

Also, $\Phi^t(m_0)$ has the usual interpretation that it represents the marginal distribution of the process at time t when the initial distribution is m_0. A nonlinear Markov chain with initial

distribution $m_0 \in \mathcal{P}(\mathcal{S})$ can then be identified with a time-inhomogeneous Markov chain with initial distribution m_0 and transition probabilities $p(s, i, t, j) = P_{ij}(t - s, \Phi^s(m_0))$.

Before we move on, let us briefly connect nonlinear Markov chains with classical time-homogeneous Markov chains. A classical Markov chain is characterized by a family of transition probabilities $P(t) = (P_{ij}(t))_{i,j \in \mathcal{S}}$ which is a family of stochastic matrices that continuously depends on $t \geq 0$ and which satisfies the Chapman–Kolmogorov equations. Therefore, classical Markov chains are a special case of the nonlinear Markov chains we consider here; namely, they are nonlinear Markov chains where the family of transition probabilities does not depend on the distribution of the process. In this case the associated Markov semigroup reads

$$\Phi_j^t(m_0) = \sum_{i \in \mathcal{S}} m_0^i P_{ij}(t),$$

which is linear in m_0. This gives rise to the name *nonlinear* Markov chain for the objects we investigate in this paper.

As in the theory of classical continuous-time Markov chains, the infinitesimal generator will be the cornerstone of the description and analysis of such processes. Let $\Phi^t(m)$ be differentiable in $t = 0$ and $m \in \mathcal{P}(\mathcal{S})$; then the *(nonlinear) infinitesimal generator* of the semigroup $(\Phi^t(\cdot))_{t \geq 0}$ is given by a transition rate matrix function $Q(\cdot)$ such that for $f(m) := \left. \frac{\partial}{\partial t} \Phi^t(m) \right|_{t=0}$ we have $f_j(m) = \sum_{i \in \mathcal{S}} m_i Q_{ij}(m)$ for all $j \in \mathcal{S}$ and $m \in \mathcal{P}(\mathcal{S})$.

By [10, Section 1.1], any differentiable nonlinear semigroup has a nonlinear infinitesimal generator. However, the converse problem is more important: given a transition rate matrix function (that is, a function $Q : \mathcal{P}(\mathcal{S}) \to \mathbb{R}^{S \times S}$ such that $Q(m)$ is a transition rate matrix for all $m \in \mathcal{P}(\mathcal{S})$), is there a nonlinear Markov semigroup (and thus a nonlinear Markov chain) such that Q is the nonlinear infinitesimal generator of the process? Relying on the semigroup identity $\Phi^{t+s} = \Phi^t \Phi^s$, this problem is equivalent to the following Cauchy problem: is there, for any $m_0 \in \mathcal{P}(\mathcal{S})$, a solution $(\Phi^t(m_0))_{t \geq 0}$ of

$$\frac{\partial}{\partial t} \Phi^t(m_0) = \Phi^t(m_0) Q(\Phi^t(m_0)), \quad \Phi^0(m_0) = m_0,$$

such that $\Phi^t(\cdot)$ is a continuous function ranging from $\mathcal{P}(\mathcal{S})$ to itself, and such that $\Phi^t(m) \in \mathcal{P}(\mathcal{S})$ for all $t \geq 0$ and $m \in \mathcal{P}(\mathcal{S})$?

In the monograph [10] the problem of constructing a semigroup from a given generator is treated in a very general setting. Here, we present a result with easy-to-verify conditions tailored to the specific situation of nonlinear Markov chains with finite state space. The proof of the result, which relies on classical arguments from the theory of ordinary differential equations, is presented in the appendix.

Proposition 1. *Let $Q : \mathcal{P}(\mathcal{S}) \to \mathbb{R}^{S \times S}$ be a transition rate matrix function such that $Q_{ij}(m)$ is Lipschitz continuous for all $i, j \in \mathcal{S}$. Then there is a unique Markov semigroup $(\Phi^t(\cdot))_{t \geq 0}$ such that Q is the infinitesimal generator for $(\Phi^t(\cdot))_{t \geq 0}$.*

This proposition sheds more light on the additional modelling possibilities of nonlinear Markov chains compared to classical Markov chains: indeed, whereas classical Markov chains are characterized through a transition rate matrix $Q \in \mathbb{R}^{S \times S}$, the nonlinear Markov chains that we consider here are described by a function $Q : \mathcal{P}(\mathcal{S}) \to \mathbb{R}^{S \times S}$. This function now allows for the transition rates of the processes to depend on the current distribution of the process. This often occurs in applications, for example in evolutionary game dynamics or infection models (e.g. susceptible–infectious–recovered models) [10]. Moreover, such processes naturally arise

as the limit of weakly interacting particles or agents [11], which is why these processes play a role in mean-field game theory.

Let us illustrate both aspects in a simple toy example. Namely, let the function $Q : \mathcal{P}(\{1, 2\}) \to \mathbb{R}^{2 \times 2}$ be given by

$$m \mapsto Q(m) := \begin{pmatrix} -am_1 & am_1 \\ b & -b \end{pmatrix},$$

where both a and b are positive. This is a simple nonlinear Markov chain, where the transition rate from state 2 to state 1 is as in a classical Markov chain given by the fixed constant b, and the transition rate from state 1 to state 2 increases linearly in the current probability of state 1. This Markov chain naturally arises as a limit of classical Markov chains as follows. Let us assume there are N agents/particles which can be in either state 1 or state 2. Let n_i denote the number of agents/particles in state i, and let $E_N = \{(n_1, n_2) : n_1 + n_2 = N\}$ be the state space. Moreover, let the transition rate matrix of a classical Markov chain describing the motion of these particles be given by

$$Q_{(n_1, n_2), (k_1, k_2)} = \begin{cases} a\frac{n_1}{N} & \text{if } n_1 = k_1 + 1, \\ b & \text{if } n_2 = k_2 + 1, \\ -a\frac{n_1}{N} - b & \text{if } 0 \leq n_1 < N, \; n_1 = k_1, \\ -a & \text{if } n_1 = k_1 = N, \\ 0 & \text{otherwise,} \end{cases}$$

for all $(n_1, n_2), (k_1, k_2) \in E_N$. In this process, the more agents/particles are in state 1, the larger the transition rate is for other agents to go to state 2; i.e., the agents/particles face congestion effects in state 1. Renormalizing the state space to $\frac{1}{N}E_N \subseteq \mathcal{P}(\{1, 2\})$, one can then show that the Markov chain describing the N-particle system converges in distribution and in probability to the nonlinear Markov chain described above (see [11] and the references therein for general results of this type). This means that for a large number of agents/particles the nonlinear Markov chain approximately describes the behaviour of the distribution of the agents/particles.

In this paper we are now mainly interested in the characterization of the long-term behaviour of nonlinear Markov chains. We say that $m \in \mathcal{P}(\mathcal{S})$ is an invariant distribution if $\frac{\partial}{\partial t}\Phi^0(m) = 0$ and thus also $\frac{\partial}{\partial t}\Phi^t(m) = 0$. An equivalent condition with respect to the generator is that a vector $m \in \mathcal{P}(\mathcal{S})$ is an invariant distribution if it solves $0 = m^T Q(m)$.

We say that a nonlinear Markov chain with nonlinear semigroup $(\Phi^t(\cdot))_{t \geq 0}$ is *strongly ergodic* if there exists an $\bar{m} \in \mathcal{P}(\mathcal{S})$ such that for all $m_0 \in \mathcal{P}(\mathcal{S})$ we have

$$\lim_{t \to \infty} \left|\left| \Phi^t(m_0) - \bar{m} \right|\right| = 0.$$

3. Existence and uniqueness of the invariant distribution

The invariant distributions of a nonlinear Markov chain are exactly the fixed points of the set-valued map

$$s : \mathcal{P}(\mathcal{S}) \to 2^{\mathcal{P}(\mathcal{S})}, \quad m \mapsto \{x \in \mathcal{P}(\mathcal{S}) : 0 = x^T Q(m)\}.$$

Using Kakutani's fixed point theorem, we directly obtain the existence of an invariant distribution for any generator, as follows.

Proposition 2. *Let $Q(\cdot)$ be a nonlinear generator such that the map $Q : \mathcal{P}(\mathcal{S}) \to \mathbb{R}^{S \times S}$ is continuous. Then the nonlinear Markov chain with generator $Q(\cdot)$ has an invariant distribution.*

Proof. By [8, Theorem 5.3], the set of all invariant distributions given a fixed generator matrix $Q(m)$ is the convex hull of the invariant distributions given the recurrent communication classes of $Q(m)$. Therefore, the values of the map s are non-empty, convex, and compact. Moreover, the graph of the map s is closed: let $(m^n, x^n)_{n \in \mathbb{N}}$ be a converging sequence such that $x^n \in s(m^n)$. Denote its limit by (m, x). Then $0 = (x^n)^T Q(m^n)$ for all $n \in \mathbb{N}$. By continuity of $Q(\cdot)$ we have $0 = x^T Q(m)$, which implies $x \in s(m)$. Thus, Kakutani's fixed point theorem yields a fixed point of the map s, which is an invariant distribution given $Q(\cdot)$. $\square$

If $Q(m)$ is irreducible for all $m \in \mathcal{P}(\mathcal{S})$, the sets $s(m)$ will be singletons [1, Theorem 4.2]. Let $x(m)$ denote this point. We remark that there are explicit representation formulas for $x(m)$ (e.g. [16, 19]). With these insights we provide the following sufficient criterion for the uniqueness of the invariant distribution.

Theorem 1. *Assume that $Q(m)$ is irreducible for all $m \in \mathcal{P}(\mathcal{S})$. Furthermore, assume that $f(m) := x(m) - m$ is continuously differentiable and that the matrix*

$$
M(m) := \begin{pmatrix} \frac{\partial f_1(m)}{\partial m_1} & \cdots & \frac{\partial f_1(m)}{\partial m_{S-1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{S-1}(m)}{\partial m_1} & \cdots & \frac{\partial f_{S-1}(m)}{\partial m_{S-1}} \end{pmatrix} - \begin{pmatrix} \frac{\partial f_1(m)}{\partial m_S} & \cdots & \frac{\partial f_1(m)}{\partial m_S} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{S-1}(m)}{\partial m_S} & \cdots & \frac{\partial f_{S-1}(m)}{\partial m_S} \end{pmatrix}
$$

is non-singular for all $m \in \mathcal{P}(\mathcal{S})$. Then there is a unique invariant distribution.

Proof. We first note that any invariant distribution of a nonlinear Markov chain with generator $Q(\cdot)$ is an invariant distribution m of a classical Markov chain with generator $Q(m)$. Since any invariant distribution of a classical Markov chain with generator $Q(m)$ has to satisfy that all components are strictly positive [1, Theorem 4.2], no invariant distribution of $Q(\cdot)$ lies on the boundary of $\mathcal{P}(\mathcal{S})$. Therefore, we only need to ensure the existence of a unique invariant distribution in the interior of $\mathcal{P}(\mathcal{S})$.

The set $\mathcal{P}(\mathcal{S})$ is homeomorphic to $\bar{\Omega}$ with

$$
\Omega = \left\{ m \in \mathbb{R}^{S-1} : m_i > 0 \forall i \in \{1, \dots, S-1\} \wedge \sum_{i=1}^{S-1} m_i < 1 \right\},
$$

where the continuous bijections are given as the restrictions of

$$
\phi : \mathbb{R}^{S-1} \to \mathbb{R}^S, \quad (m_1, \dots, m_{S-1}) \mapsto \left(m_1, \dots, m_{S-1}, 1 - \sum_{i=1}^{S-1} m_i \right),
$$

$$
\psi : \mathbb{R}^S \to \mathbb{R}^{S-1}, \quad (m_1, \dots, m_{S-1}, m_S) \mapsto (m_1, \dots, m_{S-1}).
$$

Define $\bar{f} : \bar{\Omega} \to \bar{\Omega}$ by $m \mapsto \psi(f(\phi(m)))$. By the chain rule we obtain

$$\frac{\partial \bar{f}(m)}{\partial m} = \frac{\partial \psi}{\partial m}(f(\phi(m))) \cdot \frac{\partial f}{\partial m}(\phi(m)) \cdot \frac{\partial \phi}{\partial m}(m)$$

$$= \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial f_1(m)}{\partial m_1} & \cdots & \frac{f_1(m)}{\partial m_S} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_S(m)}{\partial m_1} & \cdots & \frac{f_S(m)}{\partial m_S} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \dots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \\ -1 & -1 & \dots & -1 \end{pmatrix}$$

$$= M \left(\left(m_1, \dots, m_{S-1}, 1 - \sum_{i=1}^{S-1} m_i \right)^T \right).$$

The matrix $M(m)$ is, by assumption, non-singular for all $m \in \mathcal{P}(\mathcal{S})$. Thus,

$$\det \left(\frac{\partial \bar{f}(m)}{\partial m} \right) \neq 0 \quad \text{for all } m \in \bar{\Omega}.$$

Since ϕ, ψ, f, and det are continuous functions, we obtain that the function $m \mapsto \det\left(\frac{\partial \bar{f}(m)}{\partial m}\right)$ is also continuous. Thus, the intermediate value theorem yields that $\det\left(\frac{\partial \bar{f}(m)}{\partial m}\right)$ has uniform sign over $\bar{\Omega}$.

Furthermore, we note that by assumption $\mathrm{M}(m)$ is in particular non-singular for all $m \in \phi(\bar{f}^{-1}(\{0\}))$. Thus, 0 is a non-critical value of $\bar{f}$.

The map $\bar{h} : [0, 1] \times \bar{\Omega} \to \mathbb{R}^{S-1}$ given by

$$\bar{h}(t, m) = t \cdot \bar{f}(m) + (1 - t) \cdot \left(\frac{(S-1)}{S}(1, \dots, 1)^T - m \right)$$

$$= t \cdot \psi(x(\phi(m)) + (1 - t) \cdot \frac{S-1}{S}(1, \dots, 1)^T - m$$

is continuous. Furthermore, $0 \notin \bar{h}(t, \partial\Omega)$: indeed, a point $m \in \partial\Omega$ satisfies either $m_i = 0$ for some $i \in \{1, \dots, S - 1\}$ or $\sum_{i=1}^{S-1} m_i = 1$. However, by [1, Theorem 4.2], all components of the invariant distribution for an irreducible generator are strictly positive. Thus, we obtain in the first case that $h_i(t, m) > 0$ and in the second case that the sum of all components is strictly negative, which in both cases implies that $h(t, m) \neq 0$.

With these preparations we can make use of the Brouwer degree (see [6, Sections 1.1 and 1.2]); namely we obtain that

$$\deg\left(\frac{S-1}{S}(1, \dots, 1)^T - m, \Omega, 0 \right) = \deg(\bar{f}, \Omega, 0).$$

Since for continuously differentiable maps g and regular values $y \notin g(\partial \Omega)$ the degree is given by

$$\deg(g, \Omega, y) = \sum_{x \in g^{-1}(\{y\})} \operatorname{sgn} \det\left(\frac{\partial g}{\partial x}(x)\right),$$

we obtain that

$$(-1)^{S-1} = \sum_{m \in \bar{f}^{-1}(\{0\})} \operatorname{sgn} \det\left(\frac{\partial}{\partial m}\bar{f}(m)\right).$$

Because the determinant has uniform sign over $\Omega \supseteq \bar{f}^{-1}(\{0\})$, we obtain that $\bar{f}^{-1}(\{0\})$ consists of exactly one element. Thus, there is a unique stationary point for the nonlinear Markov chain with nonlinear generator $Q(\cdot)$. $\square$

Example 1. We illustrate the use of the result in an example. Consider a nonlinear Markov chain with the following generator:

$$Q(m) = \begin{pmatrix} -(b + em_1 + \epsilon) & b & em_1 + \epsilon \\ 0 & -(em_2 + \epsilon) & em_2 + \epsilon \\ \lambda & \lambda & -2\lambda \end{pmatrix},$$

where all constants are strictly positive. This nonlinear Markov chain arises in a mean-field game model of consumer choice with congestion effects (see [15], which also gives detailed calculations). In this setting the invariant distributions are given as the solution(s) of the nonlinear equation $0 = m^T Q(m)$, for which closed-form solutions are hard or impossible to obtain. However, it is possible to verify that the matrix $M(m)$ is non-singular for all $m \in \mathcal{P}(\mathcal{S})$ yielding a unique invariant distribution. This information can in particular be used to obtain certain characteristic properties of the solutions.

4. Examples of peculiar limit behaviour

The following examples show that the limit behaviour of nonlinear Markov chains (even in the case of small state spaces) is more complex than that of classical continuous-time Markov chains. In particular, it may be that the marginal distributions do not converge, but are periodic; and a nonlinear Markov chain with an irreducible nonlinear generator may not be strongly ergodic, but may exhibit convergence towards several different invariant distributions.

4.1. An example with periodic marginal distributions

Let $B = \mathcal{P}(\{1, 2, 3\}) \cap \{m \in \mathbb{R}^3 : \min\{m_1, m_2, m_3\} \geq \frac{1}{10}\}$, and for all $m \in B$ define the matrix Q as follows:

$$Q_{13}(m) = \frac{1}{m_1}\left(\frac{1}{3} - m_2\right) \mathbb{I}_{\{m_2 \leq \frac{1}{3}\}}, \qquad Q_{23}(m) = \frac{1}{m_2}\left(m_1 - \frac{1}{3}\right) \mathbb{I}_{\{m_1 \geq \frac{1}{3}\}},$$

$$Q_{31}(m) = \frac{1}{m_3}\left(m_2 - \frac{1}{3}\right) \mathbb{I}_{\{m_2 \geq \frac{1}{3}\}}, \qquad Q_{32}(m) = \frac{1}{m_3}\left(\frac{1}{3} - m_1\right) \mathbb{I}_{\{m_1 \leq \frac{1}{3}\}},$$

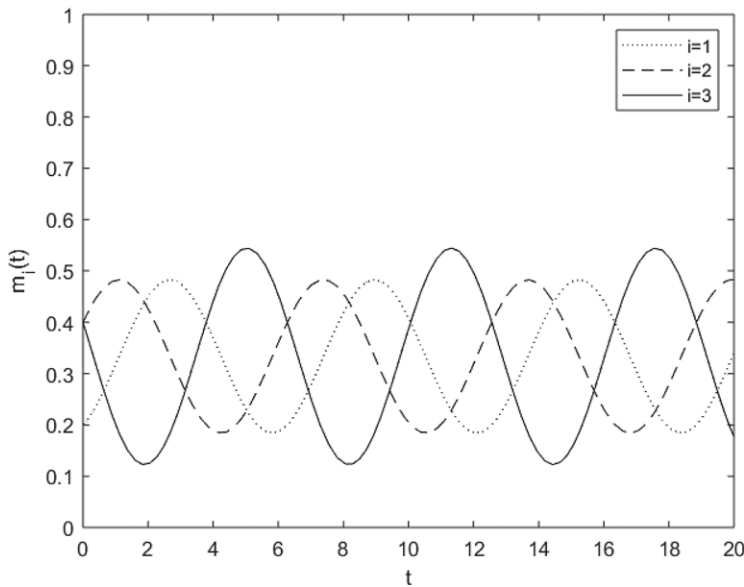$$Q_{12}(m) = Q_{21}(m) = 0, \qquad\qquad Q_{ii}(m) = -\sum_{j \neq i} Q_{ij}(m),$$

FIGURE 1. The marginal distributions of the nonlinear continuous-time Markov chain with initial distribution $m_0 = (0.2, 0.4, 0.4)$.

where $\mathbb{I}_A$ is 1 if A is true and 0 otherwise. Since all transition rates on B are Lipschitz continuous functions, there is an extension of $Q_{ij}(\cdot)$ on $\mathcal{P}(\mathcal{S})$ for all $i, j \in \mathcal{S}$, which is again Lipschitz continuous. Thus, a nonlinear Markov chain with generator Q exists, and whenever $\Phi^t(m_0) \in B$, we have

$$
\frac{\partial}{\partial t}\Phi_1^t(m_0) = \begin{cases} \Phi_1^t(m_0) \cdot \left(-\frac{1}{\Phi_1^t(m_0)}\left(\frac{1}{3} - \Phi_2^t(m_0)\right)\right), & \Phi_2^t(m_0) \le \frac{1}{3}, \\ \Phi_3^t(m_0) \cdot \left(\frac{1}{\Phi_3^t(m_0)}\left(\Phi_2^t(m_0) - \frac{1}{3}\right)\right), & \Phi_2^t(m_0) \ge \frac{1}{3}, \end{cases}
$$

$$
= \Phi_2^t(m_0) - \frac{1}{3},
$$

$$
\frac{\partial}{\partial t}\Phi_2^t(m_0) = \begin{cases} \Phi_2^t(m_0) \cdot \left(-\frac{1}{\Phi_2^t(m_0)}\left(\Phi_1^t(m_0) - \frac{1}{3}\right)\right), & \Phi_1^t(m_0) \ge \frac{1}{3}, \\ \Phi_3^t(m_0) \cdot \left(\frac{1}{\Phi_3^t(m_0)}\left(\frac{1}{3} - \Phi_1^t(m_0)\right)\right), & \Phi_1^t(m_0) \le \frac{1}{3}, \end{cases}
$$

$$
= \frac{1}{3} - \Phi_1^t(m_0),
$$

$$
\frac{\partial}{\partial t}\Phi_3^t(m_0) = \Phi_1^t(m_0) - \Phi_2^t(m_0).
$$

Thus, for any neighbourhood $U \subseteq B$ of $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)^T$ the first two components of the marginal behave like the classical harmonic oscillator. Therefore, there are initial distributions such that the marginals are periodic. An example is the initial distribution $m_0 = (0.2, 0.4, 0.4)$, for which the marginals are plotted in Figure 1.

4.2. An example of a nonlinear Markov chain with irreducible generator that is not strongly ergodic

Let

$$
Q(m) = \begin{pmatrix} -\left(\frac{29}{3}m_1^2 - 16m_1 + \frac{22}{3} \right) & \frac{29}{3}m_1^2 - 16m_1 + \frac{22}{3} \\ m_1^2 + m_1 + 1 & -\left(m_1^2 + m_1 + 1 \right) \end{pmatrix}.
$$

This matrix is irreducible for all $m \in \mathcal{P}(\{1, 2\})$, since $m_1^2 + m_1 + 1 \geq 1$ and $\frac{29}{3}m_1^2 - 16m_1 + \frac{22}{3} \geq \frac{62}{82}$ for all $m_1 \geq 0$.

The ordinary differential equation describing the marginals for the initial condition $m_0 \in \mathcal{P}(\{1, 2\})$ is given by

$$
\frac{\partial}{\partial t} \Phi_1^t(m_0) = -\frac{32}{3} \left(\Phi_1^t(m_0) \right)^3 + 16 \left(\Phi_1^t(m_0) \right)^2 - \frac{22}{3} \Phi_1^t(m_0) + 1 =: f\left(\Phi_1^t(m_0) \right),
$$

$$
\frac{\partial}{\partial t} \Phi_2^t(m_0) = \frac{32}{3} \left(\Phi_1^t(m_0) \right)^3 - 16 \left(\Phi_1^t(m_0) \right)^2 + \frac{22}{3} \Phi_1^t(m_0) - 1 = -f\left(\Phi_1^t(m_0) \right).
$$

We obtain that there are three stationary points, $m^1 = (0.25, 0.75)$, $m^2 = (0.5, 0.5)$, and $m^3 = (0.75, 0.25)$, and the following convergence behaviour:

- Since the function $f(\cdot)$ is strictly positive on $[0, 0.25)$, the trajectories will converge towards $m_1 = 0.25$ for all initial conditions $(m_0)_1 \in [0, 0.25)$.

- Since the function $f(\cdot)$ is strictly negative on $(0.25, 0.5)$, the trajectories will converge towards $m_1 = 0.25$ for all initial conditions $(m_0)_1 \in (0.25, 0.5)$.

- Since the function $f(\cdot)$ is strictly positive on $(0.5, 0.75)$, the trajectories will converge towards $m_1 = 0.75$ for all initial conditions $(m_0)_1 \in (0.5, 0.75)$.

- Since the function $f(\cdot)$ is strictly negative on $(0.75, 1]$, the trajectories will converge towards $m_1 = 0.75$ for all initial conditions $(m_0)_1 \in (0.75, 1]$.

This behaviour is illustrated in Figure 2, where several trajectories for different initial conditions are plotted.

5. Sufficient criteria for ergodicity for small state spaces

Although nonlinear Markov chains have more complex limit behaviour, we still obtain sufficient criteria for ergodicity in the case of a small number of states. Here we present these criteria and discuss applicability as well as the problems that occur for larger state spaces.

Proposition 3. *Let $S = 2$, and assume that $f : [0, 1] \to \mathbb{R}$ defined via*

$$
f(m_1) := m_1 \cdot (Q_{11}(m_1, 1 - m_1)) + (1 - m_1) \cdot Q_{21}(m_1, 1 - m_1)
$$

is continuous. Furthermore, assume that $(\bar{m}, 1 - \bar{m})$ is the unique stationary point given Q. Then the nonlinear Markov chain is strongly ergodic.

Proof. An equilibrium point is characterized by the property that $\frac{\partial}{\partial t} \Phi^t(m) = 0$. By flow-invariance of $\mathcal{P}(\mathcal{S})$ for the ordinary differential equation $\frac{\partial}{\partial t} \Phi^t(m_0) = \Phi^t(m_0) Q(\Phi^t(m_0))$ (see the proof of Theorem 1), which implies that $\frac{\partial}{\partial t} \Phi_1^t(m) + \frac{\partial}{\partial t} \Phi_2^t(m) = 0$, this property is equivalent to the fact that $\frac{\partial}{\partial t} \Phi_1^t(m) = 0$.
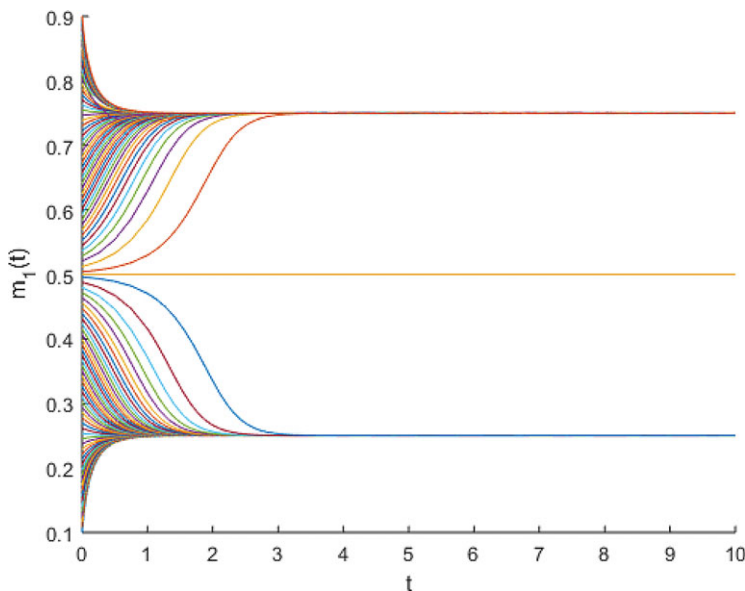
FIGURE 2. The trajectories of the nonlinear Markov chain for several initial conditions.

Since $\frac{\partial}{\partial t}\Phi_1^t(m) = f(m_1)$ and since we have a unique equilibrium point, we obtain that $f(\bar{m}) = 0$ and $f(m_1) \neq 0$ for all $m_1 \neq \bar{m}$. Since $f(\cdot)$ is continuous, we obtain that $f(\cdot)$ is non-vanishing on $[0, \bar{m})$ and $(\bar{m}, 1]$ and has uniform sign on each of these sets. Since $Q(\cdot)$ is a conservative generator we moreover obtain that $f(0) \geq 0$ and $f(1) \leq 0$. Thus, we obtain that $f(m_1) > 0$ for all $m_1 \in [0, \bar{m})$ and $f(m_1) < 0$ for all $m_1 \in (\bar{m}, 1]$. This in turn yields that $[0, 1]$ is flow-invariant for $\dot{m}_1 = f(m_1)$.

Fix $m_0 \in \mathcal{P}(\mathcal{S})$. Then the systems $\frac{\partial}{\partial t}\Phi^t(m_0) = Q(\Phi^t(m_0))^T \Phi^t(m_0)$ and $\frac{\partial}{\partial t}\tilde{\Phi}^t(m_0)_1 = f(\tilde{\Phi}^t(m_0))$ are equivalent in the sense that $\Phi_1^t(m_0) = \tilde{\Phi}^t(m_0)$ for all $t \geq 0$, $m_0 \in \mathcal{P}(\{1, 2\})$: Indeed, let $\Phi^t(m_0) = (\Phi_1^t(m_0), \Phi_2^t(m_0))$ be a solution of the differential equation $\frac{\partial}{\partial t}\Phi^t(m_0) = Q(\Phi^t(m_0))^T \Phi^t(m_0)$ with initial condition $\Phi^0(m_0) = m_0$. By flow-invariance of $\mathcal{P}(\mathcal{S})$ for $\frac{\partial}{\partial t}\Phi^t(m_0) = Q(\Phi^t(m_0))^T \Phi^t(m_0)$ (see Theorem 1), we have $\Phi_2^t(m_0) = 1 - \Phi_1^t(m_0)$ for all $t \geq 0$. Thus, $\frac{\partial}{\partial t}\Phi^t(m_0) = Q(\Phi^t(m_0))^T \Phi^t(m_0)$ is equivalent to

$$\begin{cases} \frac{\partial}{\partial t}\Phi_1^t(m_0) & = \Phi_1^t(m_0) \cdot (Q_{11}(\Phi_1^t(m_0), 1 - \Phi_1^t(m_0))) \\ & \quad + (1 - \Phi_1^t(m_0)) \cdot Q_{21}(\Phi_1^t(m_0), 1 - \Phi_1^t(m_0)), \\ -\frac{\partial}{\partial t}\Phi_1^t(m_0) & = \Phi_1^t(m_0) \cdot (-Q_{12}(\Phi_1^t(m_0), 1 - \Phi_1^t(m_0))) \\ & \quad + (1 - \Phi_1^t(m_0)) \cdot Q_{22}(\Phi_1^t(m_0), 1 - \Phi_1^t(m_0)). \end{cases} \quad (1)$$

Therefore, $\Phi_1^t(m_0)$ is indeed a solution of $\frac{\partial}{\partial t}\Phi_1^t(m_0) = f(\Phi_1^t(m_0))$. For the converse implication we first note that, because $Q(m)$ is conservative for all $m \in \mathcal{P}(\mathcal{S})$, the last equation of (1) is the first equation multiplied by (-1). If $\tilde{\Phi}^t(m_0)$ satisfies $\frac{\partial}{\partial t}\tilde{\Phi}^t(m_0) = f(\tilde{\Phi}^t(m_0))$, $\tilde{\Phi}^0(m_0) = (m_0)_1 \in [0, 1]$, then, by flow-invariance, $\tilde{\Phi}^t(m_0)) \in [0, 1]$ for all $t \geq 0$. Thus, the function $\Phi^t(m_0) = (\tilde{\Phi}^t(m_0), 1 - \tilde{\Phi}^t(m_0))$ satisfies $\frac{\partial}{\partial t}\Phi^t(m_0) = Q(\Phi^t(m_0))^T \Phi^t(m_0)$.

The desired convergence statement follows directly from $f(m_1) > 0$ for all $m_1 \in [0, \bar{m})$ and $f(m_1) < 0$ for all $m_1 \in (\bar{m}, 1]$. □

We also obtain a sufficient criterion for the case of three states. The proof technique is similar to the two-state case. Indeed, we first show that our system is equivalent to a two-dimensional system, for which we can then use standard tools for two-dimensional dynamical systems, exploiting that the dynamical system has a particular shape since $Q(\cdot)$ is a conservative generator.

As mentioned, for systems with three states we obtain that, given $m_0 \in \mathcal{P}(\mathcal{S})$, the function $\Phi^t(m_0) = (\Phi_1^t(m_0), \Phi_2^t(m_0), \Phi_3^t(m_0))$ is a solution of $\frac{\partial}{\partial t}\Phi^t(m_0) = Q(\Phi^t(m_0))^T \Phi^t(m_0)$, $\Phi^0(m_0) = m_0$ if and only of $(\Phi_1^t(m_0), \Phi_2^t(m_0))$ is a solution of

$$
\begin{pmatrix} \frac{\partial}{\partial t}\Phi_1^t(m_0) \\ \frac{\partial}{\partial t}\Phi_2^t(m_0) \end{pmatrix} = f \begin{pmatrix} \Phi_1^t(m_0) \\ \Phi_2^t(m_0) \end{pmatrix}, \quad \begin{pmatrix} \Phi_1^0(m_0) \\ \Phi_2^0(m_0) \end{pmatrix} = \begin{pmatrix} (m_0)_1 \\ (m_0)_2 \end{pmatrix},
$$

where

$$
f \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} Q_{31}(\hat{m}) + (Q_{11}(\hat{m}) - Q_{31}(\hat{m}))m_1 + (Q_{21}(\hat{m}) - Q_{31}(\hat{m}))m_2 \\ Q_{32}(\hat{m}) + (Q_{12}(\hat{m}) - Q_{32}(\hat{m}))m_1 + (Q_{22}(\hat{m}) - Q_{32}(\hat{m}))m_2 \end{pmatrix} \tag{2}
$$

and $\hat{m} = (m_1, m_2, 1 - m_1 - m_2)$. Indeed, the proof is analogous to the proof for the two-state case; the central adjustment is to prove the flow-invariance of $\{(m_1, m_2) \in [0, \infty) : m_1 + m_2 \le 1\}$ for

$$
\left(\frac{\partial}{\partial t}\Phi_1^t(m_0), \frac{\partial}{\partial t}\Phi_2^t(m_0) \right)^T = f(\Phi_1^t(m_0), \Phi_2^t(m_0))
$$

instead of the flow-invariance of $[0, 1]$ for $\Phi_1^t(m_0) = f(\Phi_1^t(m_0))$. This statement is proven in the appendix (Lemma 1).

To show the desired convergence statement, we now rely on the Poincaré–Bendixson theorem [22, Chapter 7], which characterizes the ω-limit sets $\omega_+(m_0)$ of a trajectory with initial condition $\Phi^0(m_0) = m_0$.

Theorem 2. *Let $O \supseteq \{(m_1, m_2) \in [0, \infty)^2 : m_1 + m_2 \le 1\}$ be a simply connected and bounded region such that there is a continuously differentiable function $f : O \to \mathbb{R}^2$ satisfying (2) on $\mathcal{P}(\mathcal{S})$. Let $\bar{m}$ be the unique stationary point given $Q(\cdot)$. Furthermore, assume that*

(a) $\frac{\partial f_1}{\partial m_1}(m) + \frac{\partial f_2}{\partial m_2}(m)$ *is non-vanishing for all $m \in O$ and has uniform sign on O, and*

(b) *it holds that*

$$
\frac{\partial f_1}{\partial m_1}(\bar{m}) \cdot \frac{\partial f_2}{\partial m_2}(\bar{m}) - \frac{\partial f_1}{\partial m_2}(\bar{m}) \cdot \frac{\partial f_2}{\partial m_1}(\bar{m}) > 0
$$

or that

$$
\left(\frac{\partial f_1}{\partial m_1}(\bar{m}) + \frac{\partial f_2}{\partial m_2}(\bar{m}) \right)^2 - 4\left(\frac{\partial f_1}{\partial m_1}(\bar{m}) \cdot \frac{\partial f_2}{\partial m_2}(\bar{m}) - \frac{\partial f_1}{\partial m_2}(\bar{m}) \cdot \frac{\partial f_2}{\partial m_1}(\bar{m}) \right) < 0.
$$

Then the nonlinear Markov chain is strongly ergodic.

Proof. Since the set $F := \{(m_1, m_2)^T \in \mathbb{R}^2 : m_1, m_2 \ge 0 \wedge m_1 + m_2 \le 1\}$ is flow-invariant for

$$
\left(\frac{\partial}{\partial t}\Phi_1^t(m_0), \frac{\partial}{\partial t}\Phi_2^t(m_0) \right)^T = f(\Phi_1^t(m_0), \Phi_2^t(m_0)),
$$

any trajectory will stay in this set. Since the set F is compact, we obtain by [22, Lemma 6.6] that $\omega_+(m_0)$ lies in F. Since there is, by assumption, only one stationary point, we can apply

the Poincaré–Bendixson theorem [22, Theorem 7.16], which yields that one of the following three cases holds:

(i) $\omega_+(m_0) = \{\bar{m}\}$,

(ii) $\omega_+(m_0)$ is a regular periodic orbit, or

(iii) $\omega_+(m_0)$ consists of (finitely many) fixed points $x_1, \ldots, x_k$ and non-closed orbits $\gamma(z)$ such that $\omega_\pm(z) \in \{x_1, \ldots, x_k\}$.

By the condition (a) and Bendixson's criterion [9, Theorem 3.5], the case (ii) is not possible. Since, by the condition (b), the point $\bar{m}$ is not a saddle point, there is no homoclinic path joining $\bar{m}$ to itself. Therefore, since $\bar{m}$ is the only stationary point, the case (iii) is also not possible. Thus, $\omega_+(m_0) = \{\bar{m}\}$. Since the trajectory considered lies in the compact set F, we moreover obtain by [22, Lemma 6.7] that

$$0 = \lim_{t\to\infty} d\big(\Phi^t(m_0), \omega_+(m_0)\big) = \lim_{t\to\infty} d\big(\Phi^t(m_0), \bar{m}\big).$$

$\square$

Remark 1. The equivalence of the systems considered above and $S - 1$ systems on some subset of $\mathbb{R}^{S-1}$, as well as the construction performed in Section 4.1, hint at the general problem for a larger number of states ($S \geq 4$). It might happen that the dynamics of the nonlinear Markov chain describe a classical 'chaotic' nonlinear system like the Lorentz system. In other words, the difficulties that arise in the classical theory of dynamical systems might also arise here, for which reason criteria for a larger number of states are more complex.

Example 2. Theorem 2 now yields strong ergodicity of the nonlinear Markov chain introduced at the end of Section 3. In this setting the function f is given by

$$f\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} \lambda - em_1^2 - (b + \epsilon + \lambda)m_1 - \lambda m_2 \\ \lambda + (b - \lambda)m_1 - em_2^2 - (\epsilon + \lambda)m_2 \end{pmatrix};$$

moreover, we have $\frac{\partial f_1}{\partial m_1}(m) + \frac{\partial f_2}{\partial m_2}(m) < 0$ for all $m \in N_\epsilon([0, 1]^2)$, as well as

$$\frac{\partial f_1}{\partial m_1}(m) \frac{\partial f_2}{\partial m_2}(m) - \frac{\partial f_1}{\partial m_2}(m) \frac{\partial f_2}{\partial m_1}(m) > 0$$

for all $m \in [0, 1]^2$ and thus in particular for the unique invariant distribution. Therefore, by Theorem 2 we obtain strong ergodicity.

Appendix A. Proofs

Proof of Theorem 1. We first note that

$$f(m) := \left(\sum_{i\in\mathcal{S}} m_i Q_{ij}(m) \right)_{j\in\mathcal{S}}$$

is Lipschitz continuous on $\mathcal{P}(\mathcal{S})$. Indeed, let L be a Lipschitz constant for all functions $Q_{ij}(\cdot)$ $(i, j \in \mathcal{S})$ simultaneously. Moreover, since $\mathcal{P}(\mathcal{S})$ is compact there is a finite constant

$$M := \sup_{m\in\mathcal{P}(\mathcal{S}), i,j\in\mathcal{S}} Q_{ij}(m).$$

Thus, we have

$$|f(m^1) - f(m^2)|_1 \leq (M + L)S \cdot \left|m^1 - m^2\right|_1 .$$

By McShane's extension theorem [14], there is a Lipschitz continuous extension $\tilde{f} : \mathbb{R}^S \to \mathbb{R}^S$ of f. Let us fix an arbitrary $m_0 \in \mathcal{P}(\mathcal{S})$. By the classical existence and uniqueness theorem for ordinary differential equations, we obtain that there is a unique solution of $\Phi^{\cdot}(m_0) : [0, \infty) \to \mathbb{R}^S$ of $\frac{\partial}{\partial t} \Phi^t(m_0) = \tilde{f}(\Phi^t(m_0))$, $\Phi^0(m_0) = m_0$.

As a next step we show that the vectors $f(m) = \tilde{f}(m)$ lie, for all $m \in \mathcal{P}(\mathcal{S})$, in the Bouligand tangent cone

$$T_{\mathcal{P}(\mathcal{S})}(m) = \left\{ y \in \mathbb{R}^S : \liminf_{h \downarrow 0} \frac{d(m + hy, \mathcal{P}(\mathcal{S}))}{h} = 0 \right\}$$

$$= \left\{ y \in \mathbb{R}^S : y_i \geq 0 \forall i \in \mathcal{S} \text{ s.t. } m_i = 0 \wedge \sum_{i \in \mathcal{S}} y_i = 0 \right\},$$

where the second line follows from [2, Proposition 5.1.7]. Indeed, since for all interior points of $\mathcal{P}(\mathcal{S})$ the condition is trivially satisfied, it suffices to consider the boundary points $m \in \partial \mathcal{P}(\mathcal{S})$. These points satisfy that there is at least one $j \in \mathcal{P}(\mathcal{S})$ such that $m_j = 0$. Since the only non-positive column entry of $Q_{\cdot j}$ (which is Q_{jj}) gets weight m_j, the vector $f(m) = (\sum_{i \in \mathcal{S}} m_i Q_{ija}(m))_{j \in \mathcal{S}}$ will have non-negative entries at each $j \in \mathcal{S}$ such that $m_j = 0$. Since Q is conservative, we moreover obtain that

$$\sum_{j \in \mathcal{S}} \sum_{i \in \mathcal{S}} m_i Q_{ija}(m) = \sum_{i \in \mathcal{S}} \underbrace{\sum_{j \in \mathcal{S}} Q_{ija}(m)}_{=0} m_i = 0.$$

Thus, $f(m) = \tilde{f}(m) \in T_{\mathcal{P}(\mathcal{S})}(m)$ for all $m \in \mathcal{P}(\mathcal{S})$. Therefore, by the classical flow-invariance statement for ordinary differential equations ([23, Theorem 10.XVI]), we find that the solution satisfies $m(t) \in \mathcal{P}(\mathcal{S})$ for all $t \geq 0$. Thus, $\Phi^{\cdot}(m_0) : [0, \infty) \to \mathbb{R}^S$ is also the unique solution of $\frac{\partial}{\partial t} \Phi^t(m_0) = f(\Phi^t(m_0))$, $\Phi^0(m_0) = m_0$. The continuity of $\Phi^t(\cdot)$ follows from a classical general dependence theorem [23, Theorem 12.VII]. $\qquad\square$

Lemma 1. *The set $N = \{(m_1, m_2) \in [0, \infty) : m_1 + m_2 \leq 1\}$ is flow-invariant for*

$$\left(\frac{\partial}{\partial t} \Phi^t_1(m_0), \frac{\partial}{\partial t} \Phi^t_2(m_0) \right)^T = f(\Phi^t_1(m_0), \Phi^t_2(m_0)).$$

Proof. The statement follows from [7, Lemma 1]. This lemma states that for an open set $O \subseteq \mathbb{R}^S$ and a family of continuously differentiable functions $g_i : O \to \mathbb{R}$ ($i \in \{1, \ldots, k\}$), the set

$$M = \{x \in O : g_i(x) \leq 0 \text{ for all } i \in \{1, \ldots, k\}\}$$

is flow-invariant for $\dot{x} = f(x)$ whenever for any $x \in \partial M$ there is an $i \in \{1, \ldots, k\}$ such that $g_i(x) = 0$ and

$$\langle f(x), \nabla g_i(x) \rangle < 0.$$

Indeed, in our case we have

$$M = \{x \in \mathbb{R}^S : -m_1 \leq 0 \wedge -m_2 \leq 0 \wedge m_1 + m_2 \leq 1\},$$

and the boundary points of this set satisfy either $m_i = 0$ for at least one $i \in \{1, 2\}$, or $m_1 + m_2 = 1$. Since $Q(\cdot)$ is conservative and irreducible, we obtain

$$\left\langle f((m_1, m_2)^T), \nabla(-m_i) \right\rangle < 0$$

in the first case and

$$\left\langle f((m_1, m_2)^T), \nabla(m_1 + m_2 - 1) \right\rangle < 0$$

in the second case. Therefore, the claim follows. $\square$

Funding information

Competing interests

References

[1] ASMUSSEN, S. (2003). *Applied Probability and Queues*. Springer, New York.

[2] AUBIN, J.-P. AND CELLINA, A. (1984). *Differential Inclusions: Set-Valued Maps and Viability Theory*. Springer, Berlin, Heidelberg.

[3] BATHER, J. A. (1963). Two non-linear birth and death processes. *J. Austral. Math. Soc.* **3**, 104–116.

[4] BUTKOVSKY, O. A. (2014). On ergodic properties of nonlinear Markov chains and stochastic McKean–Vlasov equations. *Theory Prob. Appl.* **58**, 661–674.

[5] CARMONA, R. AND DELARUE, F. (2015). Forward–backward stochastic differential equations and controlled McKean–Vlasov dynamics. *Ann. Prob.* **43**, 2647–2700.

[6] DEIMLING, K. (1985). *Nonlinear Functional Analysis*. Springer, Berlin.

[7] FERNANDES, M. L. AND ZANOLIN, F. (1987). Remarks on strongly flow-invariant sets. *J. Math. Anal. Appl.* **128**, 176–188.

[8] IOSIFESCU, M. (1980). *Finite Markov Processes and Their Applications*. John Wiley, Chichester.

[9] JORDAN, D. W. AND SMITH, P. (2005). *Nonlinear Ordinary Differential Equations: An Introduction to Dynamical Systems*. Oxford University Press.

[10] KOLOKOLTSOV, V. N. (2010). *Nonlinear Markov Processes and Kinetic Equations*. Cambridge University Press.

[11] KOLOKOLTSOV, V. N. (2012). Nonlinear Markov games on a finite state space (mean-field and binary interactions). *Internat. J. Statist. Prob.* **1**, 77–91.

[12] KOLOKOLTSOV, V. N. AND MALAFEYEV, O. A. (2019). *Many Agent Games in Socio-economic Systems: Corruption, Inspection, Coalition Building, Network Growth, Security*. Springer, Cham.

[13] MCKEAN, H. P., Jr. (1966). A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. USA* **56**, 1907–1911.

[14] MCSHANE, E. J. (1934). Extension of range of functions. *Bull. Amer. Math. Soc.* **40**, 837–842.

[15] NEUMANN, B. A. (2019). Stationary equilibria of mean field games with finite state and action space : existence, computation, stability, and a myopic adjustment process. Doctoral Thesis, Universität Hamburg.

[16] NEUMANN, B. A. (2020). Stationary equilibria of mean field games with finite state and action space. *Dynam. Games Appl.* **10**, 845–871.

[17] NOVOZHILOV, A. S., KAREV, G. P. AND KOONIN, E. V. (2006). Biological applications of the theory of birth-and-death processes. *Briefings Bioinformat.* **7**, 70–85.

[18] PHAM, H. AND WEI, X. (2017). Dynamic programming for optimal control of stochastic McKean–Vlasov dynamics. *SIAM J. Control Optimization* **55**, 1069–1101.

[19] RESNICK, S. I. (1992). *Adventures in Stochastic Processes*. Birkhäuser, Boston.

[20] SABUROV, M. (2016). Ergodicity of nonlinear Markov operators on the finite dimensional space. *Nonlinear Anal. Theory Methods Appl.* **143**, 105–119.

[21] SZNITMAN, A.-S. (1991). Topics in propagation of chaos. In *Ecole d'Eté de Probabilités de Saint-Flour XIX—1989*, ed. P.-L. Hennequin, Springer, Berlin, Heidelberg, pp. 165–251.

[22] TESCHL, G. (2012). *Ordinary Differential Equations and Dynamical Systems*. American Mathematical Society, Providence, RI.

[23] WALTER, W. (1998). *Ordinary Differential Equations*. Springer, New York.