Check for
updates

# Extinction scenarios in evolutionary processes: a multinomial Wright–Fisher approach

**Alexander Roitershtein[1] · Reza Rastegar[2] · Robert S. Chapkin[3] · Ivan Ivanov[4]**

## Abstract

We study a discrete-time multi-type Wright–Fisher population process. The mean-field dynamics of the stochastic process is induced by a general replicator difference equation. We prove several results regarding the asymptotic behavior of the model, focusing on the impact of the mean-field dynamics on it. One of the results is a limit theorem that describes sufficient conditions for an almost certain path to extinction, first eliminating the type which is the least fit at the mean-field equilibrium. The effect is explained by the metastability of the stochastic system, which under the conditions of the theorem spends almost all time before the extinction event in a neighborhood of the equilibrium. In addition to the limit theorems, we propose a maximization principle for a general deterministic replicator dynamics and study its implications for the stochastic model.

✉ Ivan Ivanov
iivanov@cvm.tamu.edu

Alexander Roitershtein
alexander@stat.tamu.edu

Reza Rastegar
reza.j.rastegar@gmail.com

Robert S. Chapkin
r-chapkin@tamu.edu

1 Department of Statistics, Texas A&M University, College Station, TX 77843, USA

2 Occidental Petroleum Corporation, Houston, TX 77046, USA

3 Department of Nutrition - Program in Integrative Nutrition & Complex Diseases, Texas A&M University, College Station, TX 77843, USA

4 Department of Veterinary Physiology and Pharmacology, Texas A&M University, College Station, TX 77843, USA

**Mathematics Subject Classification** Primary 92D15 · 92D25; Secondary 37N25 · 91A22 · 60J10

## 1 Introduction

In this paper, we study a multi-type Wright–Fisher process, modeling the evolution of a population of cells or microorganisms in discrete time. Each cell belongs to one of the several given types, and its type evolves stochastically while the total number of cells does not change with time. The process is a time-homogeneous Markov chain that exhibits a rich and complex behavior.

Due to the apparent universality of its dynamics, availability of effective approximation schemes, and its relative amenability to numerical simulations, the Wright–Fisher process constitutes a popular modelling tool in applied biological research. In particular, there is a growing body of literature using the Wright–Fisher model of evolutionary game theory (Imhof and Nowak 2006) and its variations along with other tools of game theory and evolutionary biology to study compartmental models of complex microbial communities and general metabolic networks (Cavaliere et al. 2017; Hummert et al. 2014; Tarnita 2017; West et al. 2007; Zeng et al. 2017; Zomorrodi and Segrè 2017). For instance, recent studies (Zomorrodi and Segrè 2017) incorporate evolutionary game theory and concepts from behavioral economics into their hybrid models of metabolic networks including microbial communities with so-called Black Queen functions. See also Cai et al. (2019), Hummert et al. (2014), Ruppin et al. (2010) for related research. The Wright–Fisher model has been employed to study the evolution of host-associated microbial communities (Zeng et al. 2017). In a number of studies, the Wright–Fisher model was applied to examine adaptive dynamics in living organisms, for instance viral and immune populations perpetually adapting (Nourmohammad et al. 2016) and response to T-cell-mediated immune pressure (Barton et al. 2016; Chen and Kardar 2019) in HIV-infected patients. Another interesting application of the Wright–Fisher model to the study of evolution of viruses has been previously reported (Zinger et al. 2019).

Other biological applications of the Wright–Fisher model supported by experimental data include, for instance, substitutions in protein-coding genes (dos Reis 2015), analysis of the single-nucleotide polymorphism differentiation between populations (Dewar et al. 2011), correlations between genetic diversity in the soil and above-ground population (Koopmann et al. 2017), and persistence of decease-associated gene alleles in wild populations (Vallier et al. 2017). In cancer-related research, the Wright–Fisher model has been used to model tumor progression (Beerenwinkel et al. 2007; Datta et al. 2013), clonal interference (Garcia et al. 2018; Park and Krug 2007), cancer risk across species (Peto's paradox) (Caulin et al. 2015), and intra-tumoral heterogeneity (Iwasa and Michor 2011). Not surprisingly, the Wright–Fisher model is also extensively used to model phenomena in theoretical evolutionary biology (Cvijović et al. 2018; Goudenège and Zitt 2015; Hobolth and Sireén 2016; Hofrichter et al. 2017; Houchmandzadeh 2015; John and Seetharaman 2016; Proulx 2011; Simonsen Speed et al. 2019; Waxman and Loewe 2010).

We now describe the Wright–Fisher model and state our results in a semi-formal setting. For a population with $N$ cells, we define $M$-dimensional vectors

$$\mathbf{Z}_k^{(N)} := \big(Z_k^{(N)}(1), Z_k^{(N)}(2), \ldots, Z_k^{(N)}(M)\big), \qquad k = 0, 1, 2, \ldots,$$

where $Z_k^{(N)}(i)$ represents the number of type $i$ cells in the generation $k$. We denote by

$$\mathbf{X}_k^{(N)} := \big(X_k^{(N)}(1), X_k^{(N)}(2), \ldots, X_k^{(N)}(M)\big)$$

the frequency vectors of the population. That is,

$$X_k^{(N)}(i) = \frac{Z_k^{(N)}(i)}{N} = \frac{1}{N} \cdot \#\{\text{cells of type } i \text{ in the } k\text{-th generation}\}, \qquad (1)$$

for all $1 \leq i \leq M$. We refer to the frequency vector $\mathbf{X}_k^{(N)}$ as the *profile of the population at time $k$*.

With biological applications in mind, the terms "particles" and "cells" are used interchangeably throughout this paper. We assume that the sequence $\mathbf{Z}_k^{(N)}$ forms a time-homogeneous Markov chain with the Wright–Fisher frequency-dependent transition mechanism. That is, conditioned on $\mathbf{Z}_k^{(N)}$, the next generation vector $\mathbf{Z}_{k+1}^{(N)}$ has a multinomial distribution with the profile-dependent parameter $\mathbf{\Gamma}\big(\mathbf{X}_k^{(N)}\big)$, where $\mathbf{\Gamma}$ is a vector field that shapes the fitness landscape of the population (see Sect. 2.2 for details). In Example 1 we list several examples of the fitness taken from either biological or biologically inspired mathematical literature and briefly indicate a rationale behind their introduction.

It is commonly assumed in evolutionary population biology that particles reproduce or adopt their behavior according to their fitness, which depends on the population composition through a parameter representing utility of a random interaction within the population. In Sect. 3, in order to relate the stochastic dynamics of the Wright–Fisher model to *first biological principles*, we interpret the mean-field model in terms of the reproductive fitness of the stochastic process and study some basic properties of the fitness. Our main contribution here is the following maximization principle for a general deterministic replicator dynamics. An informal statement of this contribution is as follows. We refer to a vector $\mathbf{x}$ in $\mathbb{R}^m$ as a *population profile vector* if all its components are non-negative and their sum is one (i. e., $\mathbf{x}$ is a probability vector).

**Theorem 1** *Assume that $\mathbf{\Gamma}$ is a continuous vector-field. Then there exists a "nice" (non-decreasing, bounded, complete Lyapunov, see Definition 2) function $h$ mapping population profiles into the set of reals $\mathbb{R}$, such that:*

(i) *$h(\mathbf{x})$ is the average $\sum_i x(i)\varphi_i(\mathbf{x})$ of an ordinal Darwinian fitness given by*

$$\varphi_i(\mathbf{x}) = \frac{\Gamma_i(\mathbf{x})}{x(i)} h(\mathbf{x}), \qquad i = 1, \ldots, M. \qquad (2)$$

(ii) $h\big(\mathbf{\Gamma}(\mathbf{x})\big) \geq h(\mathbf{x})$ *for any population profile* $\mathbf{x}$.

(iii) $h\big(\mathbf{\Gamma}(\mathbf{x})\big) > h(\mathbf{x})$ *for any population profile* $\mathbf{x}$ *which is, loosely speaking, "not recurrent" (see Definition* 1 *for the definition of the "recurrence").*

We refer to Theorem 3.2 for a formal statement of the result. The theorem is a simple consequence of the "fundamental theorem of dynamical systems" due to Conley (1978), see Sect. 3.2 for details.

Let $\mathbf{x}$ be an arbitrary population profile, set $\boldsymbol{\psi}_0 = \mathbf{x}$, and define recursively,

$$\boldsymbol{\psi}_{k+1} = \mathbf{\Gamma}(\boldsymbol{\psi}_k), \qquad k \geq 0. \tag{3}$$

We will occasionally use the notation $\boldsymbol{\psi}_k(\mathbf{x})$ to indicate the initial state $\boldsymbol{\psi}_0 = \mathbf{x}$. It turns out that dynamical system $\boldsymbol{\psi}_k$ serves as mean-field approximation for the stochastic model (see Sect. 2.2 and Theorem 4.1 for details). It follows from Theorem 1 that the average fitness $h$ is non-decreasing along the orbits of $\boldsymbol{\psi}_k$ :

$$h(\boldsymbol{\psi}_{k+1}) \geq h(\boldsymbol{\psi}_k),$$

and, moreover, (again, informally) "typically" we have a strong inequality $h(\boldsymbol{\psi}_{k+1}) > h(\boldsymbol{\psi}_k)$. Furthermore, since $h$ is a bounded function,

$$\lim_{k \to \infty} h(\boldsymbol{\psi}_k) = \sup_k h(\boldsymbol{\psi}_k)$$

exists for all initial conditions $\boldsymbol{\psi}_0$.

In fact, the average/population Darwinian fitness (as defined in (14) below) generically remains constant during the evolution of any Wright–Fisher model driven by a replicator mean-field dynamics (see Sect. 3.1 for details). Theorem 1 offers an alternative notion of fitness, a distinguished ordinal utility, that is free of this drawback. In contrast to the most of the known results of this type (Ao 2005; Birch 2016; Bratus et al. 2018; Doebeli et al. 2017; Edwards 2016; Ewens 2011; Ewens and Lessard 2015; Gavrilets 2010; Obolski et al. 2018; Orr 2009; Queller 2017), our maximization principle 1) is genuinely universal, such that it can be applied to any discrete-time and continuous in state space replicator dynamics; and 2) refers to maximization of the entire reproduction fitness, rather than of its part.

In Sect. 4.3 we study fixation probabilities of a certain class of Wright–Fisher models, induced by a classical linear-fractional replicator mean-field dynamics with a symmetric payoff matrix which corresponds to a partnership normal-form game with a unique evolutionary stable equilibrium. Understanding the route to fixation of multivariate population models is of a great importance to many branches of biology (Der et al. 2011; Gavrilets 2010; Hofrichter et al. 2017; Nowak and Sigmund 2004; Obolski et al. 2018; Waxman and Loewe 2010). For instance, extinction risk of biologically important components, regulated by metabolism in living cells, has been the focus of intensive research and is a key factor in numerous biotechnologically and biomedically relevant applications (Assaf and Meerson 2008; Samoilov and Arkin 2006). Our contribution can be summarized as follows. Firstly, Theorem 4.6 offers an interpretation of the notion "the least fit" which is in the direct accordance to the law of competitive

exclusion and Fisher's fundamental theorem of natural selection. Secondly, on a more practical note, the theorem suggests that for the underlying class of population models, the path to extinction is almost deterministic when the population size is large. We remark that for a different class of models, a similar fixation mechanism has been described in Assaf and Meerson (2010), Assaf and Mobilia (2011), Park and Traulsen (2017). However, our proof is based on a different idea, and is mathematically rigorous in contrast to the previous work (where, in order to establish the result, analytical methods are combined with some heuristic arguments).

Informally, the result in Theorem 4.6 can be stated as follows:

**Theorem 2** *Suppose that* $\boldsymbol{\Gamma}$ *satisfies the following set of conditions:*

(i) *There is* $\boldsymbol{\chi_{eq}} \in \mathbb{R}^M$ *such that* $\chi_{eq}(i) > 0$ *for all* $i = 1, \ldots, M$ *and* $\lim_{k \to \infty} \boldsymbol{\psi}_k = \boldsymbol{\chi_{eq}}$ *whenever* $\psi_0(i) > 0$ *for all* $i = 1, \ldots, M$.

(ii) *There are no mutations, that is* $\Gamma_i(x) = 0$ *if* $x(i) = 0$.

(iii) *The Markov chain* $\mathbf{X}_k^{(N)}$ *is metastable, that is it spends a long time confined in a small neighborhood of the deterministic equilibrium* $\boldsymbol{\chi_{eq}}$ *before a large (catastrophic) fluctuation occurs and one of the types gets instantly extinct.*

*Then the probability that the least fit at the equilibrium* $\boldsymbol{\chi_{eq}}$ *type will extinct first converges to one when N goes to infinity.*

*Metastability* or *stable coexistence* is described in Broekman et al. (2019) as "the long-term persistence of multiple competing species, without any species being competitively excluded by the others, and with each species able to recover from perturbation to low density." Various mathematical conditions equivalent to a variant of the above concept can be found in Auletta et al. (2018), Bovier and den Hollander (2015), Landim (2019), see also Assaf and Meerson (2010). The interplay between metastability, quasi-stationary distribution, strong persistence, and extinction probabilities in multitype population models has been discussed by many authors, see, for instance, Ao (2005), Ashcroft et al. (2015), Benaïm and Schreiber (2019), Block and Allen (2000), Gyllenberg and Silvestrov (2008), Hofbauer and Sigmund (1998), Hutson and Schmitt (1992), Kang and Chesson (2010), Méléard and Villemonais (2012), Steinsaltz and Evans (2004), Xu et al. (2014), Zhou and Qian (2011) and references therein.

Heuristically, the result in our theorem suggests that if a stochastic population model has a metastable state, then under some technical conditions that grantee a strong enough attraction to the quasi-equilibrium, the extinction by a series of intermediate steps is "costly" to the system (that is where the existence of a Lyapunov function is relevant to the result), and the overwhelmingly most likely extinction scenario is an instantaneous "collapse" of the "weakest link", namely an instantaneous extinction of the type which is the least abundant at the quasi-equilibrium state. We also emphasize the importance of the strong attraction to the quasi-equilibrium factor, without which the mechanism of extinction can be completely different as it has been illustrated in the literature.

The benefit of the kind of results stated in Theorem 2 for applied biological research is that they elucidate conditions under which the ultimate outcome of a complex stochastic evolution is practically deterministic, and thus can be robustly predicted in a general accordance with a simple variation of the *competitive exclusion principle*.

In Sect. 4.3, in addition to our main result stated in Theorem 4.6, we prove an auxiliary Proposition 4.7 which complements (Losert and Akin 1983) and explains the mechanism of convergence to the equilibrium under the conditions of Theorem 4.6. The proposition establishes a strong contraction near the equilibrium property for the systems considered in Theorem 4.6. The claim is a "reverse engineering" in the sense that it would immediately imply Theorem 2-(i), whereas the latter has been obtained in Losert and Akin (1983) in a different way and is used here to prove the claim in Proposition 4.7.

The Wright–Fisher Markov chains are known to exhibit complex multi-scale dynamics, partially because of the non-smooth boundary of their natural phase space. The analytic intractability of Wright–Fisher models has motivated an extensive literature on the approximation of these stochastic processes (Chalub and Souza 2014; Ewens 2012; Hartle and Clark 2007; Hofrichter et al. 2017; Nagylaki 1990). These approximations "illuminate the evolutionary process by focusing attention on the appropriate scales for the parameters, random variables, and time and by exhibiting the parameter combinations on which the limiting process depends" (Nagylaki 1990). Our results in Sect. 4.3 concerning the model's path to extinction rely partially on the Gaussian approximation constructed in Theorem 4.1 and error estimates proved in Sect. 4.2. The construction is fairly standard and has been used to study other stochastic population models. Our results in Sects. 4.1 and 4.2 can be informally summarized as follows:

**Theorem 3** *Assuming that the vector field $\boldsymbol{\Gamma}$ is smooth enough, the following is true:*

(i) $\mathbf{X}_k^{(N)} = \boldsymbol{\psi}_k + \frac{\mathbf{u}_k^{(N)}}{\sqrt{N}}$, *where $\boldsymbol{\psi}_k$ is the deterministic mean-field sequence defined in* (3) *(in particular, independent of $N$) and $\mathbf{u}_k^{(N)}$ is a random sequence that, as $N$ goes to infinity, converges in distribution to a solution of certain random linear recursion.*

(ii) *For a fixed $N$ and any given $\varepsilon > 0$, with an overwhelmingly large probability, the norm of the noise perturbation vector $\frac{\mathbf{u}_k^{(N)}}{\sqrt{N}}$ will not exceed $\varepsilon$ for an order of $e^{\alpha \varepsilon^2 N}$ first steps of the Markov chain, where $\alpha > 0$ is a constant independent of $N$ and $\varepsilon$.*

The theorem implies that when the population size $N$ is large, the stochastic Wright–Fisher model can be interpreted as a perturbation of the mean-field dynamical system $\boldsymbol{\Gamma}$ with a small random noise.

Due to the particular geometry of the phase space boundary, the stochastic dynamics of the Wright–Fisher model often exhibits structural patterns (specifically, invariant and quasi-invariant sets) inherited from the mean-field model. "Appendix A" of this paper is devoted to the exploration of the peculiar connection between invariant sets of the mean-field dynamical system and the form of the equilibria of the stochastic model when mutation is allowed. The results in the appendix aim to contribute to the general theory of Wright–Fisher processes with mutations and are not used directly anywhere in the paper.

The rest of the paper is organized as follows. The mathematical model is formally introduced in Sect. 2. Section 3 is concerned with the geometry of the fitness

landscape and fitness maximization principle for the deterministic mean-field model. Our main results for the stochastic dynamics are stated and discussed in Sect. 4, with proofs deferred to Sect. 5. Finally, conclusions are outlined in Sect. 6. "Appendix A" ties the geometric properties of the fitness to the form of stochastic equilibria of the Wright–Fisher process. In order to facilitate the analysis of stochastic persistence in the Wright–Fisher model considered in Sect. 4, we survey in "Appendix B" a mathematical theory of coexistence for a class of deterministic interacting species models, known as permanent (or uniform persistent) dynamical systems.

## 2 The model

The aim of this section is to formally present general form of the Wright–Fisher model and introduce some of the key notions pertinent to the general framework. The section is divided into two subsections. Section 2.1 introduces a broader geometric context and thus sets the stage for a formal definition of the model given in Sect. 2.2.

### 2.1 Geometric setup, notation

Fix $M \in \mathbb{N}$ and define $S_M = \{1, 2, \ldots, M\}$. Let $\mathbb{Z}_+$ and $\mathbb{R}_+$ denote, respectively, the set of non-negative integers and the set of non-negative reals. Given a vector $\mathbf{x} \in \mathbb{R}^M$, we usually use the notation $x(i)$ to denote its $i$-th coordinate. The only exception is in the context of vector fields, whose $i$-th component will typically be indicated with the subscript $i$.

Let $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ denote the unit vector in the direction of the $i$-th axis in $\mathbb{R}^M$. We define the $(M - 1)$-dimensional simplex

$$\Delta_M = \left\{ \mathbf{x} = \big(x(1), \ldots, x(M)\big) \in \mathbb{R}_+^M : \sum_{i \in S_M} x(i) = 1 \right\}$$

and denote by $V_M$ the set of its vertices:

$$V_M = \{\mathbf{e}_j : j \in S_M\} \subset \mathbb{R}^M. \tag{4}$$

$\Delta_M$ is a natural state space for the entire collection of stochastic processes $(\mathbf{X}^{(N)})_{N \in \mathbb{N}}$ of (1). We equip $\Delta_M$ with the usual Euclidean topology inherited from $\mathbb{R}^M$. For an arbitrary closed subset $\Delta$ of $\Delta_M$, we denote by $\mathrm{bd}(\Delta)$ its topological boundary and by $\mathrm{Int}(\Delta)$ its topological interior $\Delta \backslash \mathrm{bd}(\Delta)$. For any non-empty subset $J$ of $S_M$ we denote by $\Delta_{[J]}$ the simplex built on the vertices corresponding to $J$. That is,

$$\Delta_{[J]} = \left\{ \mathbf{x} \in \mathbb{R}_+^M : \sum_{i \in J} x(i) = 1 \right\} \subset \mathrm{bd}(\Delta_M).$$

We denote by $\partial(\Delta_{[J]})$ the manifold boundary of $\Delta_{[J]}$ and by $\Delta_{[J]}^\circ$ the corresponding interior $\Delta_{[J]} \backslash \partial(\Delta_{[J]})$. That is,

$$\partial(\Delta_{[J]}) = \big\{\mathbf{x} \in \Delta_{[J]} : x(j) = 0 \text{ for some } j \in J\big\}$$

and

$$\Delta^{\circ}_{[J]} = \big\{\mathbf{x} \in \Delta_{[J]} : x(j) > 0 \text{ for all } j \in J\big\}.$$

We write $\partial(\Delta_M)$ and $\Delta^{\circ}_M$ for, respectively, $\partial(\Delta_{[S_M]})$ and $\Delta^{\circ}_{[S_M]}$. Note that $\mathrm{bd}(\Delta_{[J]}) = \Delta_{[J]}$, and hence $\partial(\Delta_{[J]}) \neq \mathrm{bd}(\Delta_{[J]})$ for any *proper* subset $J$ of $S_M$.

Furthermore, let

$$\Delta_{M,N} = \left\{0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}, 1\right\}^M.$$

Note that $\mathbf{X}^{(N)}_k \in \Delta_{M,N}$ for all $k \in \mathbb{Z}_+$ and $N \in \mathbb{N}$. For any set of indices $J \subset S_M$ we write

$$\Delta_{[J],N} = \Delta_{[J]} \bigcap \Delta_{M,N}, \quad \partial(\Delta_{[J],N}) = \partial(\Delta_{[J]}) \bigcap \Delta_{M,N},$$

$$\Delta^{\circ}_{[J],N} = \Delta^{\circ}_{[J]} \bigcap \Delta_{M,N}. \tag{5}$$

We simplify the notations $\partial(\Delta_{[S_M],N})$ and $\Delta^{\circ}_{[S_M],N}$ to $\partial(\Delta_{M,N})$ and $\Delta^{\circ}_{M,N}$, respectively.

We use the maximum ($L^{\infty}$) norms for vectors in $\Delta_M$ and functions in $C(\Delta_M)$, the space of real-valued continuous functions on $\Delta_M$:

$$\|\mathbf{x}\| := \max_{i \in S_M} |x(i)| \text{ for } \mathbf{x} \in \Delta_M \quad \text{and} \quad \|f\| := \max_{\mathbf{x} \in \Delta_M} |f(\mathbf{x})| \text{ for } f \in C(\Delta_M).$$

The Hadamard (element-wise) product of two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^M$ is denoted by $\mathbf{u} \circ \mathbf{v}$. That is, $\mathbf{u} \circ \mathbf{v} \in \mathbb{R}^M$ and

$$(u \circ v)(i) = u(i)v(i) \qquad \forall i \in S_M. \tag{6}$$

The usual dot-notation is used to denote the regular scalar product of two vectors in $\mathbb{R}^M$.

## 2.2 Formulation of the model

We now proceed with the formal definition of the Markov chain $\mathbf{X}^{(N)}_k$. Let $P$ denote the underlying probability measure defined in some large probability space that includes all random variables considered in this paper. We use $E$ and $COV$ to denote the corresponding expectation and covariance operators. For a random vector $\mathbf{X} = (X(1), X(2), \dots, X(M)) \in \mathbb{R}^M$, we set $E(\mathbf{X}) := \big(E\big(X(1)\big), E\big(X(2)\big), \dots, E\big(X(M)\big)\big)$.

Let $\mathbf{\Gamma} : \Delta_M \to \Delta_M$ be a mapping of the probability simplex $\Delta_M$ into itself. Suppose that the population profile at time $k \in \mathbb{Z}_+$, is given by a vector $\mathbf{x} =$

$(x(1), \dots, x(M)) \in \Delta_{M,N}$, i.e. $\mathbf{X}_k^{(N)} = \mathbf{x}$. Then the next generation's profile, $\mathbf{X}_{k+1}^{(N)}$, is determined by a multinomial trial where the probability of producing an $i$-th type particle is equal to $\Gamma_i(\mathbf{x})$. In other words, $\mathbf{X}^{(N)}$ is a Markov chain whose transition probability matrix is defined as

$$P\left(\mathbf{X}_{k+1}^{(N)} = \left(\frac{j_1}{N}, \dots, \frac{j_M}{N}\right) \mid \mathbf{X}_k^{(N)} = \mathbf{x}\right) = \frac{N!}{\prod_{i=1}^{M} j_i!} \prod_{i=1}^{M} \Gamma_i(\mathbf{x})^{j_i} \quad (7)$$

for all $k \in \mathbb{Z}_+$ and $\left(\frac{j_1}{N}, \dots, \frac{j_M}{N}\right) \in \Delta_{M,N}$. Note that for any state $\mathbf{x} \in \Delta_{M,N}$ and time $k \in \mathbb{Z}_+$ we have

$$E\left(\mathbf{X}_{k+1}^{(N)} \mid \mathbf{X}_k^{(N)} = \mathbf{x}\right) = \Gamma(\mathbf{x}) \quad (8)$$

and

$$COV\left(X_{k+1}^{(N)}(i) X_{k+1}^{(N)}(j) \mid X_k^{(N)} = x\right) = \begin{cases} -\frac{1}{N}\Gamma_i(\mathbf{x})\Gamma_j(\mathbf{x}) & \text{if } i \neq j, \\ \frac{1}{N}\Gamma_i(\mathbf{x})\left(1 - \Gamma_i(\mathbf{x})\right) & \text{if } i = j. \end{cases} \quad (9)$$

We refer to any Markov chain $\mathbf{X}^{(N)}$ on $\Delta_{M,N}$ that satisfies (7) as a *Wright–Fisher model associated with the update rule* $\Gamma$.

We denote transition kernel of $\mathbf{X}^{(N)}$ on $\Delta_{M,N}$ by $P_N$. That is, $P_N(\mathbf{x}, \mathbf{y})$ is the conditional probability in (7) with $y(i) = j_i/N$. For $k \in \mathbb{N}$, we denote by $P_N^k$ the $k$-th iteration of the kernel. Thus,

$$P_N^k(\mathbf{x}, \mathbf{y}) = P\left(\mathbf{X}_{m+k}^{(N)} = \mathbf{y} \mid \mathbf{X}_m^{(N)} = \mathbf{x}\right)$$

for any $m \in \mathbb{Z}_+$.

The Wright–Fisher model is completely characterized by the choice of the parameters $M$, $N$ and the update rule $\Gamma(\mathbf{x})$. Theorem 4.1 below states that when $N$ approaches infinity while $M$ and $\Gamma$ remain fixed, the update rule emerges as the model's mean-field map, in that the path of the Markov chain $\mathbf{X}^{(N)}$ with $\mathbf{X}_0^{(N)} = \mathbf{x}$ converges to the orbit of the (deterministic) discrete-time dynamical system $\boldsymbol{\psi}_k$ introduced in (3). Thus, for the large population size $N$, the Wright–Fisher process can be perceived as a small-noise perturbation of the deterministic dynamical system $\boldsymbol{\psi}_k$.

Throughout the paper we assume that the mean-field dynamics is induced by a replicator equation

$$\Gamma_i(\mathbf{x}) = \frac{x(i)\varphi_i(\mathbf{x})}{\sum_{j \in S_M} x(j)\varphi_j(\mathbf{x})} \qquad \forall i \in S_M, \quad (10)$$

where the vector-field $\boldsymbol{\varphi} : \Delta_M \to \mathbb{R}_+^M$ serves as a fitness landscape of the model (see Sect. 3 for details). For biological motivations of this definition, see (Cressman and Tao 2014; Hofbauer and Sigmund 2003; Sigmund 1986). The replicator dynamics can be viewed as a natural discretization scheme for its continuous-time counterpart

described in the form of a differential equation (Garay and Hofbauer 2003; Hofbauer and Sigmund 1998). The normalization factor $\sum_{j \in S_m} x(j)\varphi_j(\mathbf{x})$ in (10) is the total population fitness and plays a role similar to that of the partition function in statistical mechanics (Barton and Coe 2009).

**Example 1** (i) One standard choice of the fitness $\boldsymbol{\varphi}$ in the replicator equation (10) is Hofbauer (2011), Hofbauer and Sigmund (1998), Hofrichter et al. (2017), Imhof and Nowak (2006):

$$\boldsymbol{\varphi}(\mathbf{x}) = (1 - \omega)\mathbf{b} + \omega \mathbf{A}\mathbf{x}, \tag{11}$$

where $\mathbf{A} = (A_{ij})_{i,j \in S_M}$ is an $M \times M$ payoff matrix, $\mathbf{b}$ is a constant $M$-vector, and $\omega \in (0, 1)$ is a selection parameter. The evolutionary game theory interprets $(Ax)(i) = \sum_{j \in S_M} A_{ij}x(j)$ as the expected payoff of a type $i$ particle in a single interaction (game) with a particle randomly chosen from the population profile $\mathbf{x}$ when the utility of the interaction with a type $j$ particle is $A_{ij}$. The constant vector $\mathbf{b}$ represents a baseline fitness and the selection parameter $w$ is a proxy for modeling the strength of the effect of interactions on the evolution of the population compared to inheritance (Imhof and Nowak 2006).

 (ii) In replicator models associated with multiplayer games (Broom and Rychtář 2013; Gokhale and Traulsen 2014), the pairwise interaction of case (i) (two-person game) is replaced by a more complex interaction within a randomly selected group (multiplayer game). In this case, the fitness $\boldsymbol{\varphi}$ is typically a (multivariate) polynomial of $\mathbf{x}$ of a degree higher than one.

(iii) An evolutionary game dynamics with non-uniform interaction rates has been considered in Taylor and Nowak (2006). Mathematically, their assumptions amount to replacing (11) with

$$\varphi_i(\mathbf{x}) = (1 - \omega)b(i) + \omega \frac{Bx(i)}{Rx(i)},$$

where $\mathbf{R}$ is the matrix of rates reflecting the frequency of pair-wise interaction between different types, and $\mathbf{B} = \mathbf{A} \circ \mathbf{R}$, the Hadamard product of the payoff matrix $\mathbf{A}$ and $\mathbf{R}$.

(iv) Negative frequency dependence of the fitness can be captured by, for instance, allowing matrix $\mathbf{A}$ in (11) to have negative entries. In fact, it suffices that the matrix off-diagonal elements are larger than the respective diagonal elements (Broekman et al. 2019).

 (v) A common choice of a non-linear fitness is an exponential function (Hofbauer and Sigmund 1998; Sandholm 2010):

$$\varphi_i(\mathbf{x}) = e^{\mathbf{e}taAx(i)}, \tag{12}$$

where, as before, $\mathbf{A}$ is a payoff matrix, and $\mathbf{e}ta > 0$ is a "cooling" parameter [analogue of the inverse temperature in statistical mechanics (Barton and Coe 2009)].

## 3 Mean-field model: fitness landscape

In Sect. 3.1 we relate the update rule $\Gamma$ to a family of fitness landscapes of the mean-field model in the absence of mutations. All the fitness functions in the family yield the same *relative fitness* and differ from each other only by a scalar normalization function. Section 3.2 contains a version of Fisher's maximization principle for the deterministic mean-field model. In Sect. 3.3, we incorporate mutations into the framework presented in Sects. 3.1 and 3.2.

The main result of this section is Theorem 3.2 asserting the existence of a non-constant reproductive fitness which is maximized during the evolutionary process. The result is an immediate consequence of a version of Conley's "fundamental theorem of dynamical systems" (1978, Theorem B in Section II.6.4), a celebrated mathematical result that appears to be underutilized in biological applications.

### 3.1 Reproductive fitness and replicator dynamics

The main goal of this section is to introduce a notion of fitness suitable for supporting a formalization of Fisher's fitness maximization principle stated below in Sect. 3.2 as Theorem 3.2. Proposition 3.1 provides a useful characterization of the fitness, while Theorem 1 [which is a citation from Garay and Hofbauer (2003)] addresses some of its more technical properties, all related to the characterization of the support of the fitness function.

Let $\mathcal{C} : \Delta_M \to 2^{S_M}$ be the *composition set-function* defined as

$$\mathcal{C}(\mathbf{x}) = \{j \in S_M : x(j) > 0\} = \text{"support of vector } \mathbf{x}\text{"}. \tag{13}$$

According to (8), for a population profile $\mathbf{x} \in \Delta_{M,N}$, the $M$-dimensional vector $\mathbf{f}(\mathbf{x})$ with $i$-th component equal to

$$f_i(\mathbf{x}) = \frac{\Gamma_i(\mathbf{x})}{x(i)} \quad \text{if} \quad i \in \mathcal{C}(\mathbf{x}) \tag{14}$$

represents the *Darwinian fitness* of the profile $\mathbf{x}$ (Hartle and Clark 2007; Orr 2009), see also Doebeli et al. (2017), Svensson and Connallon (2019). We will adopt the terminology of Chalub and Souza (2017), Traulsen et al. (2005) and call a vector field $\boldsymbol{\varphi} : \Delta_M \to \mathbb{R}_+^M$ a *reproductive fitness* of the population if for all $\mathbf{x} \in \Delta_M$ we have:

$$\sum_{j \in \mathcal{C}(\mathbf{x})} x(j)\varphi_j(\mathbf{x}) > 0 \quad \text{and} \quad \Gamma_i(\mathbf{x}) = \frac{x(i)\varphi_i(\mathbf{x})}{\sum_{j \in S_M} x(j)\varphi_j(\mathbf{x})} \quad \forall i \in \mathcal{C}(\mathbf{x}). \tag{15}$$

Notice that even though we assume that both $\boldsymbol{\varphi}(\mathbf{x})$ and $\mathbf{f}(\mathbf{x})$ are vectors in $\mathbb{R}_+^M$, the definitions (14) and (15) only consider $i \in \mathcal{C}(\mathbf{x})$, imposing no restriction on the rest of the vectors components. Notice that, since $\boldsymbol{\Gamma}$ is a probability vector, the Darwinian fitness defined in (14) can serve as an example of the reproductive fitness. Furthermore, a comparison of (15) and (14) shows that $\varphi(\mathbf{x})$ is a reproductive fitness if and only if

$\varphi(\mathbf{x}) = h(\mathbf{x})\mathbf{f}(\mathbf{x})$ for all $\mathbf{x} \in \Delta_M$, where $h : \Delta_M \to \mathbb{R}_+$ is an arbitrary strictly positive scalar-valued function.

Both concepts of fitness are pertinent to the amount of reproductive success of particles of a certain type per capita of the population. We remark that in McAvoy et al. (2018), $\varphi$ is referred to as the *reproductive capacity* rather than the reproductive fitness. A similar definition of reproductive fitness, but with the particular normalization $\mathbf{x} \cdot \varphi = 1$ and the restriction of the domain of $\mathbf{\Gamma}$ to the interior $\Delta_M^\circ$, is given in Sandholm (2010, p. 160).

In the following straightforward extension of Lemma 1 in Chalub and Souza (2017), we show that two notions of fitness essentially agree under the following no-mutations condition:

$$\sum_{j \in \mathcal{C}(\mathbf{x})} \Gamma_j(\mathbf{x}) = 1 \quad \text{or, equivalently,} \quad \mathbf{\Gamma}(\mathbf{x}) \in \Delta_{[\mathcal{C}(\mathbf{x})]} \quad \forall\, \mathbf{x} \in \Delta_M. \quad (16)$$

**Proposition 3.1** *The following is true for any given* $\mathbf{\Gamma} : \Delta_M \to \Delta_M$ :

(i) *A reproductive fitness exists if and only if condition* (16) *is satisfied.*

(ii) *Suppose that condition* (16) *is satisfied and let* $h : \Delta_M \to (0, \infty)$ *be an arbitrary strictly positive function defined on the simplex. Then any* $\varphi : \Delta_M \to \mathbb{R}_+^M$ *such that*

$$\varphi_i(\mathbf{x}) = \frac{\Gamma_i(\mathbf{x})}{x(i)} h(\mathbf{x}) \quad \forall\, i \in \mathcal{C}(\mathbf{x}), \quad (17)$$

*is a reproductive fitness.*

(iii) *Conversely, if* $\varphi : \Delta_M \to \mathbb{R}_+^M$ *is a reproductive fitness, then* (17) *holds for* $h(\mathbf{x}) = \mathbf{x} \cdot \varphi(\mathbf{x})$.

(iv) *If* $\varphi : \Delta_M \to \mathbb{R}_+^M$ *is a reproductive fitness,* $\mathbf{x} \in \Delta_M$ *and* $i \in \mathcal{C}(\mathbf{x})$, *then* $\varphi_i(\mathbf{x}) > 0$ *if and only if* $\Gamma_i(\mathbf{x}) > 0$.

Note that (16) is equivalent to the condition that $\mathbf{\Gamma}\big(\Delta_{[I]}^\circ\big) \subset \Delta_{[I]}$ for all $I \subset S_M$. Therefore, if $\mathbf{\Gamma}$ is a continuous map, then (16) is equivalent to the condition

$$\mathbf{\Gamma}\big(\Delta_{[I]}\big) \subset \Delta_{[I]} \quad \forall\, I \subset S_M.$$

For continuous fitness landscapes, the following refined version of Proposition 3.1 has been obtained in Garay and Hofbauer (2003). Recall that $\mathbf{\Gamma}$ is called a homeomorphism of $\Delta_M$ if it is a continuous bijection of $\Delta_M$ and its inverse function is also continuous. Intuitively, a bijection $\mathbf{\Gamma} : \Delta_M \to \Delta_M$ is a homeomorphism when the distance $\|\mathbf{\Gamma}(x) - \mathbf{\Gamma}(y)\|$ is small if and only if $\|\mathbf{x} - \mathbf{y}\|$ is small. $\mathbf{\Gamma}$ is a diffeomorphism if it is a homeomorphism and both $\mathbf{\Gamma}$ and $\mathbf{\Gamma}^{-1}$ are continuously differentiable.

**Theorem 1** (Garay and Hofbauer 2003) *Suppose that* $\mathbf{\Gamma} = \mathbf{x} \circ \varphi$ *for a continuous* $\varphi : \Delta_M \to \mathbb{R}_+^M$ *and* $\mathbf{\Gamma}$ *is a homeomorphism of* $\Delta_M$. *Then:*

(i) *We have:*

$$\Gamma\big(\Delta_{[I]}^\circ\big) = \Delta_{[I]}^\circ \quad \text{and} \quad \Gamma\big(\partial(\Delta_{[I]})\big) = \partial\big(\Delta_{[I]}\big), \quad \forall\, I \subset S_M.$$

(ii) *If, in addition, $\mathbf{\Gamma}$ is a diffeomorphism of $\Delta_M$ and each $\varphi_i$, $i \in S_M$, can be extended to an open neighborhood $U$ of $\Delta_M$ in $\mathbb{R}^M$ in such a way that the extension is continuously differentiable on $U$, then $\varphi_i(\mathbf{x}) > 0$ for all $\mathbf{x} \in \Delta_M$ and $i \in S_M$.*

The first part of the above theorem is the Surjectivity Theorem on p. 1022 of Garay and Hofbauer (2003) and the second part is Corollary 6.1 there. The assumption that $\mathbf{\Gamma}$ is a bijection, i. e. one-to-one and onto on $\Delta_M$, seems to be quite restrictive in biological applications (Bratus et al. 2018; de Visser and Krug 2014; Gavrilets 2010; Obolski et al. 2018). Nevertheless, the theorem is of a considerable interest for us because it turns out that its conditions are satisfied for $\mathbf{\Gamma}$ in (11) and (12).

## 3.2 Maximization principle for the reproductive fitness

The aim of this section is to introduce a maximization principle suitable for the deterministic replicator dynamics and reproductive fitness. The result is stated in Theorem 3.2. An example of application to the stochastic Wright–Fisher model is given in Proposition 3.3.

For the Darwinian fitness we have $\bar{\mathbf{f}}(\mathbf{x}) := \sum_{i \in S_M} x_i f_i(\mathbf{x}) = \sum_{i \in S_M} \Gamma_i(\mathbf{x}) = 1$. Thus the average Darwinian fitness is independent of the population profile, and in particular does not change with time. In contrary, the average reproductive fitness $\bar{\boldsymbol{\varphi}}(\mathbf{x}) := \sum_{i \in S_M} x_i \varphi_i(\mathbf{x})$ is equal to the (arbitrary) normalization factor $h(\mathbf{x})$ in (17), and hence in principle can carry more structure. If the mean-field dynamics is within the domain of applicability of Fisher's fundamental theorem of natural selection, we should have (Birch 2016; Edwards 2016; Ewens and Lessard 2015; Queller 2017; Schneider 2010):

$$h\big(\mathbf{\Gamma}(\mathbf{x})\big) \geq h(\mathbf{x}) \qquad \forall \mathbf{x} \in \Delta_M. \tag{18}$$

In order to elaborate on this point, we need to introduce the notion of chain recurrence.

**Definition 1** (Conley 1978) An $\varepsilon$-chain from $\mathbf{x} \in \Delta_M$ to $\mathbf{y} \in \Delta_M$ for $\mathbf{\Gamma}$ is a sequence of points in $\Delta_M$, $\mathbf{x} = \mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n = \mathbf{y}$, with $n \in \mathbb{N}$, such that $\left\| \mathbf{\Gamma}(\mathbf{x}_i) - \mathbf{x}_{i+1} \right\| < \varepsilon$ for $0 \leq i \leq n-1$. A point $\mathbf{x} \in \Delta_M$ is called chain recurrent if for every $\varepsilon > 0$ there is an $\varepsilon$-chain from $x$ to itself. The set $\mathcal{R}(\mathbf{\Gamma})$ of chain recurrent points is called the chain recurrent set of $\mathbf{\Gamma}$.

Write $\mathbf{x} \looparrowright \mathbf{y}$ if for every $\varepsilon > 0$ there is an $\varepsilon$-chain from $\mathbf{x}$ to $\mathbf{y}$ and $\mathbf{x} \sim \mathbf{y}$ if both $\mathbf{x} \looparrowright \mathbf{y}$ and $\mathbf{y} \looparrowright \mathbf{x}$ hold true. It is easy to see that $\sim$ is an equivalence relation on $\mathcal{R}(\mathbf{\Gamma})$. For basic properties of the chain recurrent set we refer the reader to Conley (1978) and Block and Franke (1985), Fathi and Pageault (2015), Franke and Selgrade (1976). In particular, Conley showed [see Proposition 2.1 in Franke and Selgrade (1976)] that equivalence classes of $\mathcal{R}(\mathbf{\Gamma})$ are precisely its connected components. Note that if $\mathbf{\Gamma}(\mathbf{x}) = \mathbf{x}$, then $\mathbf{x}_0 = \mathbf{x}$ and $\mathbf{x}_1 = \mathbf{x}$ form an $\varepsilon$-chain for any $\varepsilon > 0$. Thus, any fixed point of $\mathbf{\Gamma}$ is chain recurrent

Somewhat surprisingly, the following general result is true:

**Theorem 3.2** *Assume that $\mathbf{\Gamma} : \Delta_M \to \Delta_M$ is continuous. Then there exists a reproductive fitness function $\boldsymbol{\varphi}$ such that*

(i) *$\boldsymbol{\varphi}$ is continuous in $\Delta_{[I]}^{\circ}$ for all $I \subset \Delta_M$.*
(ii) *The average $h(\mathbf{x}) = \mathbf{x} \cdot \boldsymbol{\varphi}(\mathbf{x})$ satisfies (18) and, moreover, is a complete Lyapunov function for $\mathbf{\Gamma}$ (see Definition 2).*

*In particular, the difference $h(\mathbf{\Gamma}(\mathbf{x})) - h(\mathbf{x})$ is strictly positive on $\Delta_M \backslash \mathcal{R}(\mathbf{\Gamma})$.*

Theorem 3.2 is an immediate consequence of a well-known mathematical result which is sometimes called a "fundamental theorem of dynamical systems" (Franks; Norton 1995; Papadimitriou and Piliouras 2016). The following definition and theorem are adapted from Hurley (1998). To simplify notation, in contrast to the conventional definition, we define Lyapunov function as a function non-decreasing (rather than non-increasing) along the orbits of $\boldsymbol{\psi}_k$.

**Definition 2** A complete Lyapunov function for $\mathbf{\Gamma} : \Delta_M \to \Delta_M$ is a continuous function $h : \Delta_M \to [0, 1]$ with the following properties:

1. If $\mathbf{x} \in \Delta_M \backslash \mathcal{R}(\mathbf{\Gamma})$, then $h(\mathbf{\Gamma}(\mathbf{x})) > h(\mathbf{x})$.
2. If $\mathbf{x}, \mathbf{y} \in \mathcal{R}(\mathbf{\Gamma})$ and $\mathbf{x} \hookrightarrow \mathbf{y}$, then $h(\mathbf{x}) \leq h(\mathbf{y})$. Moreover, $h(\mathbf{x}) = h(\mathbf{y})$ if and only if $\mathbf{x} \sim \mathbf{y}$.
3. The image $h(\mathcal{R}(\mathbf{\Gamma}))$ is a compact nowhere dense subset of $[0, 1]$.

We remark that while $h$ must be bounded as a continuous functions defined on a compact domain $\Delta_M$, the normalization $h(\mathbf{x}) \in [0, 1]$ is rather arbitrary and chosen for the convenience only. Indeed, any multiplier $ch$ of a complete Lyapunov function $h$ with a positive scalar $c > 0$, is a complete Lyapunov function itself.

We have:

**Theorem 2** (Fundamental theorem of dynamical systems) *If $\mathbf{\Gamma} : \Delta_M \to \Delta_M$ is continuous, then there is a complete Lyapunov function for $\mathbf{\Gamma}$.*

To derive Theorem 3.2 from this fundamental result, set $\varphi_i(\mathbf{x}) = \frac{\Gamma_i(\mathbf{x})}{x_i} h(\mathbf{x})$ for $i \in \mathcal{C}(\mathbf{x})$ and, for instance, $\varphi_i(\mathbf{x}) = 0$ for $i \notin \mathcal{C}(\mathbf{x})$.

Theorem 2 was established by Conley for continuous-time dynamical systems on a compact space (Conley 1978). The above discrete-time version of the theorem is taken from Hurley (1998), where it is proven for an arbitrary separable metric state space. For recent progress and refinements of the theorem see Bernardi and Florio (2019), Fathi and Pageault (2019), Pageault (2009) and references therein.

The Lyapunov functions constructed in Conley (1978) and Hurley (1998) map $\mathcal{R}(\mathbf{\Gamma})$ into a subset of the Cantor middle-third set. Thus, typically, $\Delta_M \backslash \mathcal{R}(\mathbf{\Gamma})$ is a large set [cf. Section 6.2 in Conley (1978), Theorem A in Block and Franke (1985), and an extension of Theorem 2 in Fathi and Pageault (2015)]. For example, for the class of Wright–Fisher models discussed in Sect. 4.3, $\mathcal{R}(\mathbf{\Gamma})$ consists of the union of a single interior point (asymptotically stable global attractor in $\Delta_M^{\circ}$) and the boundary of $\Delta_M$ (repeller). For a related more general case, see "Appendix B" and Example 7 below.

Recall that a point $\mathbf{x} \in \Delta_M$ is called recurrent if the orbit $(\boldsymbol{\psi}_k)_{k \in \mathbb{Z}_+}$ with $\boldsymbol{\psi}_0 = \mathbf{x}$ visits any open neighborhood of $\mathbf{x}$ infinitely often. Any recurrent point is chain recurrent, but the converse is not necessarily true. It is easy to verify that the equality

$h(\Gamma(\mathbf{x})) = h(\mathbf{x})$ holds for any recurrent point $\mathbf{x}$. The following example builds upon this observation. It shows that for a general chaotic map the complete Lyapunov function needs to be constant.

***Example 2*** The core of this example is the content of Theorem 4 in Basener (2013). Assume that (18) holds for a continuous function $h : \Delta_M \to \mathbb{R}_+$ in (17). Suppose in addition that the mean-field dynamics is chaotic. More precisely, assume that $\Gamma$ is topologically transitive, that is $\Gamma$ is continuous and for any two non-empty open subsets $U$ and $V$ of $\Delta_M$, there exists $k \in \mathbb{N}$ such that $\boldsymbol{\psi}_k(U) \bigcap V \neq \emptyset$ [this is not a standard definition of a chaotic dynamical system, we omitted from a standard Devaney's set of conditions a part which is not essential for our purpose, cf. Banks et al. (1992), Basener (2013), Silverman (1992)]. Equivalently (Silverman 1992), there is $\mathbf{x}_0 \in \Delta_M$ such that the orbit $(\boldsymbol{\psi}_k(\mathbf{x}_0))_{k \in \mathbb{Z}_+}$ is dense in $\Delta_M$.

Let $\mathbf{x} = \boldsymbol{\psi}_k(\mathbf{x}_0)$ be an arbitrary point on this orbit. We will now show that the assumption $\varepsilon := [h(\Gamma(\mathbf{x})) - h(\mathbf{x})] > 0$ leads to a contradiction. To this end, let $\delta > 0$ be so small that $|h(\mathbf{x}) - h(\mathbf{y})| < \varepsilon/2$ whenever $\|\mathbf{x} - \mathbf{y}\| < \delta$. By our assumption, there exists $m > k$ such that $\|\boldsymbol{\psi}_m(\mathbf{x}_0) - \mathbf{x}\| < \delta$, and hence $h(\boldsymbol{\psi}_m(\mathbf{x}_0)) < h(\Gamma(\mathbf{x})) = h(\boldsymbol{\psi}_{k+1}(\mathbf{x}_0))$, which is impossible in view of the monotonicity condition (18). Thus $h$ is constant along the forward orbit of $\mathbf{x}_0$. Since the normalization function $h(\mathbf{x})$ is constant on a dense subset of $\Delta_M$ and $\Delta_M$ is compact, $h(\mathbf{x})$ is a constant function. In other words, for a chaotic mean-field dynamics a constant function is the only choice of a continuous normalization $h(\mathbf{x})$ that satisfies (18).

A common example of a chaotic dynamical system is the iterations of a one-dimensional logistic map. In our framework this can be translated into the two-dimensional case $M = 2$ and $\Gamma(x, y) = (4x(1-x), 1-4x(1-x)) = (4xy, 1-4xy)$. In principle, the example can be extended to higher-dimensions using, for instance, techniques of Akhmet and Fen (2016).

A generic instance of the mean-field update function $\Gamma$ with a non-constant complete Lyapunov function is discussed in Example 7 deferred to "Appendix B".

The existence of a non-decreasing reproductive fitness does not translate straightforwardly into a counterpart for the stochastic dynamics. Indeed, in the absence of mutations $\mathbf{X}_k^{(N)}$ converges almost surely to an absorbing state, i.e. to a random vertex of the simplex $\Delta_M$. Therefore, by the bounded convergence theorem,

$$\lim_{k \to \infty} E(h(\mathbf{X}_k^{(N)}) \mid \mathbf{X}_0^{(N)} = \mathbf{x}) = \sum_{j \in S_M} p_j(\mathbf{x}) h(e_j),$$

where $p_j(\mathbf{x}) = \lim_{k \to \infty} P(\mathbf{X}_k^{(N)} = \mathbf{e}_j \mid \mathbf{X}_0^{(N)} = \mathbf{x})$. It then follows from part 2 of Definition 2 that if the boundary $\partial(\Delta_{M,N})$ is a repeller for $\Gamma$ (cf. "Appendix B"), then generically the limit is not equal to the $\sup_{k \in \mathbb{N}} E(h(\mathbf{X}_k^{(N)}) \mid \mathbf{X}_0^{(N)} = \mathbf{x})$.

The following example shows that under some additional convexity assumptions, the mean of a certain continuous reproductive fitness is a strictly increasing function of time. More precisely, for the class of Wright–Fisher models in the example, $h(\mathbf{X}_k^{(N)})$ is a submartingale for an explicit average reproductive fitness $h$.

**Example 3** Let $h : \Delta_M \to \mathbb{R}_+$ be an arbitrary convex (and hence, continuous) function that satisfies (18), not necessarily the one given by Theorem 3.2. Then, by the multivariate Jensen's inequality (Ferguson 1967, p. 76), with probability one, for all $k \in \mathbb{Z}_+$,

$$E\big(h\big(\mathbf{X}_{k+1}^{(N)}\big) \mid \mathbf{X}_k^{(N)}\big) \geq h\big(E\big(\mathbf{X}_{k+1}^{(N)} \mid \mathbf{X}_k^{(N)}\big)\big) = h\big(\mathbf{\Gamma}\big(\mathbf{X}_k^{(N)}\big)\big) \geq h\big(\mathbf{X}_k^{(N)}\big). \quad (19)$$

Thus, $M_k = h\big(\mathbf{X}_k^{(N)}\big)$ is a submartingale in the filtration $\mathcal{F}_k = \sigma\big(\mathbf{X}_0^{(N)}, \dots, \mathbf{X}_k^{(N)}\big)$. In particular, the sequence $\mathfrak{f}_k = E(M_k)$ is non-decreasing and converges to its supremum as $k \to \infty$. In fact, since the first inequality in (19) is strict whenever $\mathbf{X}_k^{(N)} \notin V_M$, the mean of the reproductive fitness is strictly increasing under the no-mutations assumption (16):

$$0 \leq E\big(h\big(\mathbf{X}_k^{(N)}\big)\big) < E\big(h\big(\mathbf{X}_{k+1}^{(N)}\big)\big) \leq \max_{\mathbf{x} \in \Delta_M} h(\mathbf{x}) < \infty$$

as long as $P\big(\mathbf{X}_0^{(N)} \in \Delta_{M,N}^\circ\big) = 1$. Note that the maximum exists and is finite because $h$ is continuous and $\Delta_M$ is a compact set.

The effect of the fitness convexity assumption on the evolution of biological populations has been discussed by several authors, see, for instance, Christie and Beekman (2017), Hintze et al. (2015), Spichtig and Kawecki (2004) and references therein. Generally speaking, the inequality $h\big(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}\big) \leq \alpha h(\mathbf{x}) + (1-\alpha)h(\mathbf{y})$, $\alpha \in [0, 1]$, $\mathbf{x}, \mathbf{y} \in \Delta_M$, that characterizes convexity, manifests an evolutionary disadvantage of the intermediate population $\alpha\mathbf{x} + (1-\alpha)\mathbf{y}$ in comparison to the extremes $\mathbf{x}, \mathbf{y}$. A typical biological example of such a situation is competitive foraging (Hintze et al. 2015).

It can be shown [see Section 6 in Hofbauer (2011) or Losert and Akin (1983)] that if $\mathbf{A}$ is a non-negative symmetric $M \times M$ matrix, then (18) holds with $h(\mathbf{x}) = \mathbf{x}^T \mathbf{A}\mathbf{x}$ for (11) and (12). Here and henceforth, $\mathbf{A}^T$ for a matrix $\mathbf{A}$ (of arbitrary dimensions) denotes the transpose of $\mathbf{A}$. Moreover, the inequality is strict unless $\mathbf{x}$ is a game equilibrium. The evolutionary games with $\mathbf{A} = \mathbf{A}^T$ are referred to in Hofbauer (2011) as partnership games, such games are a particular case of so-called potential games [see Hofbauer and Sigmund (2003), Sandholm (2010) and references therein]. In the case of the replicator mean-field dynamics (10) and potential games, the potential function $h(\mathbf{x})$ coincides with the average reproductive fitness $\mathbf{x} \cdot \boldsymbol{\varphi}(\mathbf{x})$. Let

$$W_M = \left\{ \mathbf{x} \in \mathbb{R}^M : \sum_{i \in S_M} x(i) = 0 \right\}. \quad (20)$$

It is easy to check that if $\mathbf{A}$ is positive definite on $W_M$, that is

$$\mathbf{w}^T \mathbf{A}\mathbf{w} > 0 \qquad \forall\, \mathbf{w} \in W_M, \mathbf{w} \neq \mathbf{0}, \quad (21)$$

then $\mathbf{f}(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ is a convex function on $\Delta_M$, that is

$$\frac{1}{2}\mathbf{x}^T \mathbf{A} \mathbf{x} + \frac{1}{2}\mathbf{y}^T \mathbf{A} \mathbf{y} > \frac{(\mathbf{x}+\mathbf{y})^T}{2}\mathbf{A}\frac{\mathbf{x}+\mathbf{y}}{2} \qquad \forall \mathbf{x}, \mathbf{y} \in \Delta_M, \mathbf{x} \neq \mathbf{y}.$$

We thus have:

**Proposition 3.3** (Hofbauer 2011; Losert and Akin 1983) *Suppose that* $\mathbf{A}$ *is a symmetric invertible* $M \times M$ *matrix with positive entries such that* (21) *holds true (i.e.* $\mathbf{A}$ *is positive-definite on* $W_M$). *Then* (19) *holds with* $h(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$.

The above proposition shows that a normalized (in order to scale its range into the interval [0, 1]) version of $h$, namely $\tilde{h}(\mathbf{x}) = \frac{1}{\|A\|}\mathbf{x}^T \mathbf{A} \mathbf{x}$ can serve as a complete Lyapunov function, strictly increasing within the interior of $\Delta_M$. It is shown in Losert and Akin (1983) that under the conditions of the proposition, $\boldsymbol{\psi}_k$ converges, as $k$ tends to infinity, to an equilibrium on the boundary of the simplex $\Delta_M$ (cf. Theorem 4 in "Appendix B"). The proposition thus indicates that when $\mathbf{A}$ is positive-definite on $W_M$, the stochastic and deterministic mean-field model might have very similar asymptotic behavior when $N$ is large. This observation partially motivated our Theorem 4.6 stated below in Sect. 4.3.

### 3.3 Incorporating mutations

In order to incorporate mutations into the interpretation of the update rule $\boldsymbol{\Gamma}$ in terms of the population's fitness landscape, we adapt the Wright–Fisher model with neutral selection and mutation of Hobolth and Sireén (2016). Namely, for the purpose of this discussion we will make the following assumption:

**Assumption 3.4** The update rule $\boldsymbol{\Gamma}$ can be represented in the form $\boldsymbol{\Gamma}(\mathbf{x}) = \boldsymbol{\Upsilon}(\mathbf{x}^T \boldsymbol{\Theta})$, where $\mathbf{x}^T$ denotes the transpose of $\mathbf{x}$, and

1. $\boldsymbol{\Upsilon} : \Delta_M \to \Delta_M$ is a vector field that satisfies the "no-mutations" condition (cf. (16))

$$x(j) = 0 \implies \Upsilon_j(\mathbf{x}) = 0 \tag{22}$$

    for all $\mathbf{x} \in \Delta_M$ and $j \in S_M$.

2. $\boldsymbol{\Theta} = (\Theta_{ij})_{i,j \in S_M}$ is $M \times M$ stochastic matrix, that is $\Theta_{ij} \geq 0$ for all $i, j \in S_m$ and $\sum_{j \in S_M} \Theta_{ij} = 1$ for all $i \in S_M$.

The interpretation is that while the entry $\Theta_{ij}$ of the *mutation matrix* $\boldsymbol{\Theta}$ specifies the probability of mutation (mutation rate) of a type $i$ particle into a particle of type $j$ for $j \neq i$, the update rule $\boldsymbol{\Upsilon}$ encompasses the effect of evolutional forces other than mutation, for instance genetic drift and selection. The total probability of mutation of a type $i$ particle is $\sum_{j \neq i} \Theta_{ij} = 1 - \Theta_{ii}$, and, accordingly, $\boldsymbol{\Gamma}$ satisfies the no-mutation condition if $\boldsymbol{\Theta}$ is the unit matrix. Note that, since $\boldsymbol{\Theta}$ is assumed to be stochastic, the linear transformation $\mathbf{x} \mapsto \mathbf{x}^T \boldsymbol{\Theta}$ leaves the simplex $\Delta_M$ invariant, that is $\mathbf{x}^T \boldsymbol{\Theta} \in \Delta_M$ for all $\mathbf{x} \in \Delta_M$. Both the major implicit assumptions forming a foundation for writing

$\boldsymbol{\Gamma}$ as the composition of a non-mutative map $\boldsymbol{\Upsilon}$ and the action of a mutation matrix $\boldsymbol{\Theta}$, namely that mutation can be explicitly separated from other evolutionary forces and that mutations happen at the beginning of each cycle before other evolutionary forces take effect, are standard (Antal et al. 2009; Hobolth and Sireén 2016; Hofrichter et al. 2017) even though at least the latter is a clear idealization (Stephens 2014). Certain instances of Fisher's fundamental theorem of natural selection have been extended to include mutations in Basener and Sanford (2018), Hofbauer (1985).

We now can extend the definition (15) as follows:

**Definition 3** Let Assumption 3.4 hold. A vector field $\boldsymbol{\varphi} : \Delta_M \to \mathbb{R}_+^M$ is called the reproductive fitness landscape of the model if for all $\mathbf{x} \in \Delta_M$ we have:

$$\mathbf{x}^T \boldsymbol{\Theta} \boldsymbol{\varphi}(\mathbf{x}) > 0 \quad \text{and} \quad \Gamma_i(\mathbf{x}) = \frac{(x^T \Theta)(i) \cdot \varphi_i(\mathbf{x})}{\mathbf{x}^T \boldsymbol{\Theta} \boldsymbol{\varphi}(\mathbf{x})} \quad \forall i \in S_M, \qquad (23)$$

where $\mathbf{x}^T$ is the transpose of $\mathbf{x}$.

An analogue of Proposition 3.1 for a model with mutation reads:

**Proposition 3.5** *Let Assumption 3.4 hold. Then $\boldsymbol{\varphi} : \Delta_M \to \Delta_M$ is a reproductive fitness landscape if and only if*

$$\varphi_i(\mathbf{x}) = \begin{cases} \frac{\Gamma_i(\mathbf{x})}{(x^T \Theta)(i)} h(\mathbf{x}) & \text{if } i \in \mathcal{C}(\mathbf{x}^T \boldsymbol{\Theta}), \\ 0 & \text{if } i \notin \mathcal{C}(\mathbf{x}^T \boldsymbol{\Theta}) \end{cases} \qquad (24)$$

*for some function $h : \Delta_M \to (0, \infty)$.*

**Proof** For the "if part" of the proposition, observe that if $\boldsymbol{\varphi}$ is given by (24), then

$$\mathbf{x}^T \boldsymbol{\Theta} \boldsymbol{\varphi}(\mathbf{x}) = h(\mathbf{x}) \cdot \sum_{i \in \mathcal{C}(\mathbf{x}^T \boldsymbol{\Theta})} \Gamma_i(\mathbf{x}) = h(\mathbf{x}) \cdot \sum_{i \in \mathcal{C}(\boldsymbol{\Gamma}(\mathbf{x}))} \Gamma_i(\mathbf{x}) = h(\mathbf{x}),$$

where in the second step we used the fact that $\mathcal{C}(\boldsymbol{\Gamma}(\mathbf{x})) \subset \mathcal{C}(\mathbf{x}^T \boldsymbol{\Theta})$ by virtue of (22). Thus $\boldsymbol{\varphi}$ is a solution to (23).

For the "only if part" of the proposition, note that if $\boldsymbol{\varphi}$ is a reproductive fitness landscape, then (23) implies that (24) holds with $h(\mathbf{x}) = \mathbf{x}^T \boldsymbol{\Theta} \boldsymbol{\varphi}(\mathbf{x})$. $\qquad \square$

If the weighted average $\mathbf{x}^T \boldsymbol{\Theta} \boldsymbol{\varphi}(\mathbf{x})$ is adopted as the definition of the mean fitness, Theorem 3.2 can be carried over verbatim to the setup with mutations.

## 4 Stochastic dynamics: main results

In this section we present our main results for the stochastic model. The proofs are deferred to Sect. 5.

First, we prove two different approximation results. The Markov chain $\mathbf{X}^{(N)}$, even for the classical two-allele haploid model of genetics, is fairly complex. In fact, most of

the known results about the asymptotic behavior of Wright–Fisher Markov chains were obtained through a comparison to various limit processes, including the mean-field deterministic system, branching process, diffusions and Gaussian processes (Chalub and Souza 2014; Ewens 2012; Hartle and Clark 2007; Hofrichter et al. 2017; Nagylaki 1990).

In Sect. 4.1 we are concerned with the so-called Gaussian approximation, which can be considered as an intermediate between the deterministic and the diffusion approximations. This approximation process can be thought of as the mean-field iteration scheme perturbed by an additive Gaussian noise. Theorem 4.1 constructs such an approximation for the Wright–Fisher model. Results of this type are well-known in literature, and our proof of Theorem 4.1 is a routine adaptation of classical proof methods.

In Sect. 4.2, we study the longitudinal error of the deterministic approximation by the mean-field model. Specifically, Theorem 4.2 provides an exponential lower bound on the decoupling time, namely the first time when an error of approximation exceeds a given threshold. A variation of the result is used later in the proof of Theorem 4.6, the main result of this section.

Conceptually, Theorem 4.1 is a limit theorem suggesting that for large values of $N$, the stochastic component is the major factor in determination of the asymptotic behavior of the stochastic model. Theorem 4.2 then complements this limit theorem by quantifying this intuition for the stochastic model of a given (though large) population size $N$ (versus the infinite-population-limit result of Theorem 4.1).

The impact of the mean-field dynamics on the underlying stochastic process is further studied in Sect. 4.3. The main result is Theorem 4.6 which specifies the route to extinction of the stochastic system in a situation when the mean-field model is "strongly permanent" (heuristically, strongly repelling from the boundary, cf. "Appendix B" and the discussion on p. 22) and has a unique interior equilibrium. In general, the problem of determining the route to extinction for a multi-type stochastic population model is of an obvious importance to biological sciences, but is notoriously difficult in practice.

The proof of Theorem 4.6 highlights a mechanism leading to an almost deterministic route to extinction for the model. Intuitively, under the conditions of Theorem 4.6 the system is trapped in a quasi-equilibrium stochastic state, fluctuating near the deterministic equilibrium for a very long time, eventually "untying the Gordian knot", and instantly jumping to the boundary. A similar mechanism of extinction was previously described for different stochastic population models in Assaf and Meerson (2010), Assaf and Mobilia (2011), Park and Traulsen (2017). We note that analysis in these papers involves a mixture of rigorous mathematical and heuristic arguments, and is different in nature from our approach.

### 4.1 Gaussian approximation of the Wright–Fisher model

Let $\xrightarrow{P}$ denote convergence in probability as the population size $N$ goes to infinity. The following theorem is an adaptation of Theorems 1 and 3 in Buckley and Pollett (2010) for our multivariate setup. We also refer to Klebaner and Nerman (1994), Nagylaki

(1986), Nagylaki (1990), Parra-Rojas et al. (2014) for earlier related results and to
(Rao 2001, p. 527) for a rigorous construction of a degenerate Gaussian distribution.

Recall $\boldsymbol{\psi}_k$ from (3). We have:

**Theorem 4.1** *Suppose that*

$$\mathbf{X}_0^{(N)} \xrightarrow{P} \boldsymbol{\psi}_0 \tag{25}$$

*for some $\boldsymbol{\psi}_0 \in \Delta_M$. Then the following holds true:*

(a) $\mathbf{X}_k^{(N)} \xrightarrow{P} \boldsymbol{\psi}_k$ *for all $k \in \mathbb{Z}_+$.*
(b) *Let*

$$\mathbf{u}_k^{(N)} = \sqrt{N}\big(\mathbf{X}_k^{(N)} - \boldsymbol{\psi}_k\big), \qquad N \in \mathbb{N}, \ k \in \mathbb{Z}_+,$$

*and $\boldsymbol{\Sigma}(\mathbf{x})$ be the matrix $M \times M$ with entries (cf. (9) in Sect. 2.2)*

$$\Sigma_{i,j}(\mathbf{x}) := \begin{cases} -\Gamma_i(\mathbf{x})\Gamma_j(\mathbf{x}) & \text{if } i \neq j, \\ \Gamma_i(\mathbf{x})\big(1 - \Gamma_i(\mathbf{x})\big) & \text{if } i = j \end{cases}, \qquad \mathbf{x} \in \Delta_M.$$

*Suppose that in addition to (25), $\boldsymbol{\Gamma}$ is twice continuously differentiable and $\mathbf{u}_0^{(N)}$ converges weakly, as $N$ goes to infinity, to some (possibly random) $\mathbf{u}_0$.*

*Then the sequence $\mathbf{u}^{(N)} := \big(\mathbf{u}_k^{(N)}\big)_{k \in \mathbb{Z}_+}$ converges in distribution, as $N$ goes to infinity, to a time-inhomogeneous Gaussian $AR(1)$ sequence $(\mathbf{U}_k)_{k \in \mathbb{Z}_+}$ defined by*

$$\mathbf{U}_{k+1} = \mathbf{D}_x(\boldsymbol{\psi}_k)\mathbf{U}_k + \mathbf{g}_k, \tag{26}$$

*where $\mathbf{D}_x(\boldsymbol{\psi}_k)$ denotes the Jacobian matrix of $\boldsymbol{\Gamma}$ evaluated at $\boldsymbol{\psi}_k$, and $\mathbf{g}_k$, $k \in \mathbb{Z}_+$, are independent degenerate Gaussian vectors, each $\mathbf{g}_k$ distributed as $N\big(0, \boldsymbol{\Sigma}(\boldsymbol{\psi}_k)\big)$.*

The proof of the theorem is included in Sect. 5.1. It is not hard to prove [cf. Remark
(v) on p. 61 of Buckley and Pollett (2010), see also Klebaner and Nerman (1994)] that
if $\chi_{eq}$ is a unique global stable point of $\boldsymbol{\Gamma}$, $\boldsymbol{\psi}_0 \xrightarrow{P} \chi_{eq}$ in the statement of Theorem 4.1
and, in addition, $\mathbf{u}_0^{(N)}$ converges weakly, as $N$ goes to infinity, to some $\mathbf{u}_0$, then the
linear recursion (26) can be replaced with

$$\mathbf{U}_{k+1} = \mathbf{D}_x(\chi_{eq})\mathbf{U}_k + \widetilde{\mathbf{g}}_k, \tag{27}$$

where $\widetilde{\mathbf{g}}_k$, $k \in \mathbb{Z}_+$, are i. i. d. degenerate Gaussian vectors in $\mathbb{R}^M$, each $\widetilde{\mathbf{g}}_k$ distributed
as $N\big(0, \boldsymbol{\Sigma}(\chi_{eq})\big)$. One then can show that if the spectral radius of $\mathbf{D}_x(\chi_{eq})$ is strictly
less than one, Markov chain $\mathbf{U}_k$ has a stationary distribution, see Buckley and Pollett
(2010, p. 61) for more details. In the case when $\boldsymbol{\Gamma}$ has a unique global stable point, (27)
was obtained in Parra-Rojas et al. (2014) by different from ours, analytic methods.

### 4.2 Difference equation approximation

Theorem 4.1 indicates that when $N$ is large, the trajectory of the deterministic dynamical system $\boldsymbol{\psi}_k$ can serve as a good approximation to the path of the Markov chain $\mathbf{X}_k^{(N)}$. The following theorem offers some insight into the duration of the time for which the deterministic and the stochastic paths stay fairly close to each other, before separating significantly for the first time. The theorem is a suitable modification of some results of Aydogmus (2016) in a one-dimensional setup.

For $\varepsilon > 0$, let

$$\tau_N(\varepsilon) = \inf\left\{k \in \mathbb{Z}_+ : \left\|\mathbf{X}_k^{(N)} - \boldsymbol{\psi}_k\right\| > \varepsilon\right\}, \tag{28}$$

where $\boldsymbol{\psi}_0 = \mathbf{X}_0^{(N)}$ and $\boldsymbol{\psi}_k$ is the sequence introduced in (3). Thus $\tau_N(\varepsilon)$ is the first time when the path of the Markov chain $\mathbf{X}_k^{(N)}$ deviates from the trajectory of the deterministic dynamical system (3) by more than a given threshold $\varepsilon > 0$. We have:

**Theorem 4.2** *Suppose that the map* $\boldsymbol{\Gamma} : \Delta_M \to \Delta_M$ *is Lipschitz continuous, that is there exists a constant* $\rho > 0$ *such that*

$$\|\boldsymbol{\Gamma}(\mathbf{x}) - \boldsymbol{\Gamma}(\mathbf{y})\| \le \rho\|\mathbf{x} - \mathbf{y}\| \tag{29}$$

*for all* $\mathbf{x}, \mathbf{y} \in \Delta_M$. *Then the following holds for any* $\varepsilon > 0$, $K \in \mathbb{N}$, *and* $N \in \mathbb{N}$ :

$$P\big(\tau_N(\varepsilon) \le K\big) \le 2KM \exp\left(-\frac{\varepsilon^2 c_K^2}{2}N\right), \tag{30}$$

*where*

$$c_K = \begin{cases} \frac{1-\rho}{1-\rho^K} & \text{if } \rho \ne 1, \\ \frac{1}{K} & \text{if } \rho = 1. \end{cases} \tag{31}$$

*In particular, if* $\rho < 1$ *we have:*

$$P\big(\tau_N(\varepsilon) \le K\big) \le 2KM \exp\left(-\frac{\varepsilon^2(1-\rho)^2}{2}N\right).$$

The proof of the theorem is given in Sect. 5.2. Note that the upper bound in (30) is meaningful for large values of $N$ even when $\rho \ge 1$. When $\boldsymbol{\Gamma}$ is a contraction, i.e. $\rho < 1$, the Banach fixed point theorem implies that there is a unique point $\boldsymbol{\chi}_{eq} \in \Delta_M$ such that $\boldsymbol{\Gamma}(\boldsymbol{\chi}_{eq}) = \boldsymbol{\chi}_{eq}$. Furthermore, for any $\mathbf{x} \in \Delta_M$ and $k \in \mathbb{N}$ we have

$$\left\|\boldsymbol{\psi}_k - \boldsymbol{\chi}_{eq}\right\| \le \rho^{k-1}\|\mathbf{X}_0^{(N)} - \boldsymbol{\chi}_{eq}\|,$$

that is, as $k$ goes to infinity, $\boldsymbol{\psi}_k$ converges exponentially fast to $\boldsymbol{\chi}_{eq}$. The following theorem is a suitable modification of Theorem 4.2 for contraction maps.

**Theorem 4.3** *Assume that there exist a closed subset $\mathcal{E}$ of $\Delta_M$ and a constant $\rho \in (0, 1)$ such that (29) holds for any $\mathbf{x}, \mathbf{y} \in \mathcal{E}$. Then the following holds for any $\varepsilon > 0$, $K \in \mathbb{N}$, $N \in \mathbb{N}$ such that $1 - 2M \exp\left(-\frac{(1-\rho)^2 \varepsilon^2 N}{2}\right) > 0$, and $\mathbf{X}_0^{(N)} \in \mathcal{E}$:*

$$P\big(\tau_N(\varepsilon) > K\big) \geq \Big[1 - 2M \exp\Big(-\frac{(1-\rho)^2 \varepsilon^2 N}{2}\Big)\Big]^K.$$

*In particular,*

$$E\big(\tau_N(\varepsilon)\big) = \sum_{K=0}^{\infty} P\big(\tau_N(\varepsilon) > K\big) \geq \frac{1}{2M} \exp\Big(\frac{(1-\rho)^2 \varepsilon^2 N}{2}\Big).$$

The proof of the theorem is given in Sect. 5.3. We remark that relaxing the assumption that $\mathbf{\Gamma}$ is Lipschitz continuous on the whole simplex $\Delta_M$ allows, for instance, to cover the one-dimensional setup of Aydogmus (2016) where $M = 2$, the underlying Markov chain is binomial, and $\mathbf{\Gamma}$ is a contraction only on a subset of $\Delta_2$.

### 4.3 Metastability and elimination of types

In this section, we consider a large population Wright–Fisher model with the update rule $\mathbf{\Gamma}$ satisfying the following conditions:

**Assumption 4.4** Suppose that $\mathbf{\Gamma}$ is defined by (10) and (11). Furthermore, assume that $\mathbf{b} = (1, 1, \ldots, 1)$ and $\mathbf{A}$ is a symmetric invertible $M \times M$ matrix with positive entries such that the following two conditions hold:

$$\mathbf{w}^T \mathbf{A} \mathbf{w} < 0 \qquad \forall \mathbf{w} \in W_M, \ \mathbf{w} \neq \mathbf{0}, \tag{32}$$

and there are $\boldsymbol{\chi}_{eq} \in \Delta_M^{\circ}$ and $c > 0$ such that

$$\mathbf{A} \boldsymbol{\chi}_{eq} = c\mathbf{e}, \tag{33}$$

where $\mathbf{e} := (1, \ldots, 1) \in \mathbb{R}^M$.

Note that $\Gamma_j(\mathbf{x}) = 0$ if and only if $x_j = 0$, and hence the Markov chain $\mathbf{X}^{(N)}$ is absorbing and will eventually become fixed on one of the simplex vertices (cf. "Appendix A"). The assumption is borrowed from Losert and Akin (1983), where the defined in (3) deterministic dynamical system $\boldsymbol{\psi}_k$ is studied. In particular, it is shown in Losert and Akin (1983) that under these conditions the internal equilibrium point of $\boldsymbol{\psi}_k$ is stable (see Theorem 4 in our "Appendix B" for technical details).

The main result of the section is Theorem 4.6, which intuitively says that if the push away from the boundary and toward the equilibrium is strong enough and Assumption 4.4 holds, with a high degree of certainty, the type with the lowest fitness at the equilibrium is the one that goes extinct first. The result provides an interesting link between the interior equilibrium of the mean-field dynamics and the asymptotic behavior of the stochastic system which is eventually absorbed at the boundary. A

similar phenomenon for a different type of stochastic dynamics has been described in physics papers (Assaf and Meerson 2010; Assaf and Mobilia 2011; Park and Traulsen 2017), see also Auletta et al. (2018). Key generic features that are used to prove the result are:

1. The mean-field dynamics is *permanent*, that is (see "Appendix B" below for a brief background discuion): there exists $\delta > 0$ such that

$$\liminf_{k\to\infty} \psi_k(i) > \delta \quad \forall i \in S_M, \boldsymbol{\psi}_0 \in \Delta_M^\circ.$$

   The condition ensures a strong repelling of the mean-field dynamical system from the boundary. Heuristically, for the corresponding Wright–Fisher model and a large population $N$, this condition propagates into the metastable (asymptotically quasi-stationary with a long life-time before extinction of any type) behavior of the stochastic system.
2. There is a unique interior equilibrium $\boldsymbol{\chi}_{eq}$. In addition, $\lim_{k\to\infty} \boldsymbol{\psi}_k = \boldsymbol{\chi}_{eq}$ for any $\boldsymbol{\psi}_0 \in \Delta_M^\circ$.
3. The push toward equilibrium is strong enough to force the stochastic system to follow the deterministic dynamics closely and get trapped for a prolonged time in a "metastable" state near the equilibrium.
4. Finally, the same strong push toward the equilibrium yields a large deviation principle which ensures that the stochastic system is much likely to jump to the boundary directly from the neighborhood of the equilibrium by one giant fluctuation than by a sequence of small consecutive steps.

With these features, one would expect the following. If the starting point $\mathbf{X}_0^{(N)}$ lies within the interior of $\Delta_M$, the mean-field dynamical system $\boldsymbol{\psi}_k$ converges to the unique equilibrium, $\boldsymbol{\chi}_{eq} \in \Delta_M^\circ$. It can be shown that the rate of the convergence is exponential. In view of Theorem 4 in "Appendix B" and Theorem 4.1, one can expect that the trajectory of the stochastic system will stay close to the deterministic mean-field path for a long enough time, so that the stochastic model will follow the mean-field path to a small neighborhood of $\boldsymbol{\chi}_{eq}$. Then the probability of an immediate elimination of a type $i$ will be of order $\left(1 - \Gamma_i(\boldsymbol{\chi}_{eq})\right)^N$. Assuming that

$$\text{"it is easier to return to a small neighborhood of } \boldsymbol{\chi}_{eq} \text{ than to leave it"} \quad (34)$$

we expect that the stochastic system will enter a metastable state governed by its quasi-stationary distribution and will be trapped in a neighborhood of the equilibrium for a very long time. In fact, since by virtue of Theorem 4.1, $\mathbf{X}^{(N)}$ can be considered as a random perturbation of the dynamical system $\boldsymbol{\psi}_k$ with a small noise, the quasi-stationary distribution should converge as $N$ goes to infinity to the degenerate probability measure supported by the single point $\boldsymbol{\chi}_{eq}$. Finally, general results in the theory of large deviations (Faure and Schreiber 2014; Iglehart 1972) suggest that under a suitable condition (34), with an overwhelming probability, the system will eventually jump to the boundary by one giant fluctuation. It follows from the above heuristic argument, that the probability of such an instant jump to the facet where, say, $x(j) = 0$
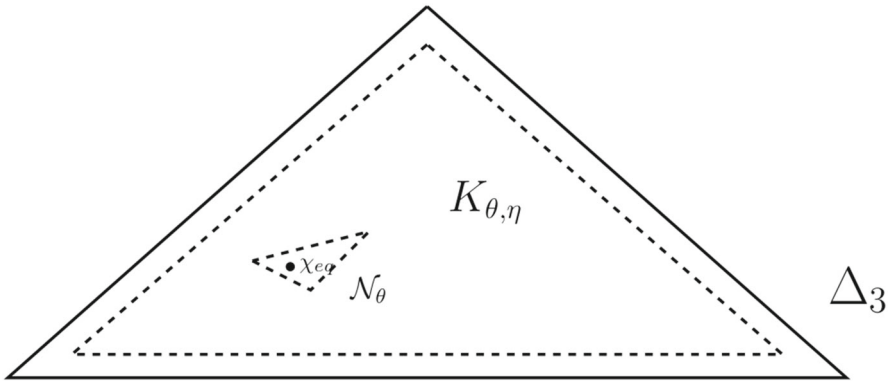
**Fig. 1** Schematic sketch (in the planar barycentric coordinates) of the simplex $\Delta_3$, the interior equilibrium $\chi_{eq}$, its neighborhood $\mathcal{N}_\theta$, and a compact set $K_{\theta,\eta}$

and the rest of coordinates are positive is roughly $\Gamma_j(\chi_{eq})^N$. For large values of $N$ this practically guarantees that the stochastic system will hit the boundary through the interior of $\Delta_{[J^*]}$, where $J^* = S_M \backslash \{j^*\}$, and $j^* \in \arg\min_{i \in S_M} \{\Gamma_i(\chi_{eq}) : i \in S_M\}$. In other words, under Assumption 4.4, when $N$ is large, the elimination of the least fit and survival of the remaining types is almost certain (see results of a numerical simulation supporting the hypothesis below in Tables 1 and 2).

We now turn to the statement of our results. First, we state the following technical result. Recall (3) and Definition 1.

**Proposition 4.5** *Let Assumption 4.4 hold. Then for any neighborhood $\mathcal{N}$ of the interior equilibrium $\chi_{eq}$ and any compact set $K \subset \Delta_M$ there exist a real constant $\varepsilon = \varepsilon(\mathcal{N}, K) > 0$ and an integer $T = T(\mathcal{N}, K) \in \mathbb{N}$ such that if $\delta \in (0, \varepsilon)$, $\mathbf{x}_0 \in K$, and $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_T$ is an $\delta$-chain of length $T$ for the dynamical system $\psi_k$, then $\mathbf{x}_T \in \mathcal{N}$.*

The proof of the proposition is deferred to Sect. 5.4. The proposition is a general property of dynamical system converging to an asymptotically stable equilibrium (see Remark 2 at the end of the proof). Since we were unable to find a reference to this claim in the literature, we provided a short self-contained proof in Sect. 5.4.

We are now ready to state the main result of this section. Suppose $\mathbf{x} \in \Delta_M$ such that at least one component $\Gamma_i(\mathbf{x})$ is not equal to $1/M$. Then we define

$$\alpha_{\mathbf{x}} = \min_{j \in S_M} \Gamma_j(\mathbf{x}), \qquad J_{\mathbf{x}}^* = \{j \in S_M : \Gamma_j(\mathbf{x}) = \alpha_{\mathbf{x}}\}, \qquad \text{and} \qquad eta_{\mathbf{x}} = \min_{j \notin J_{\mathbf{x}}^*} \Gamma_j(\mathbf{x}).$$

We set $\alpha := \alpha_{\chi_{eq}}$, $eta := eta_{\chi_{eq}}$, and $J^* := J_{\chi_{eq}}^*$.

Additionally, for any $\theta \in \left(0, \frac{eta-\alpha}{2}\right)$ and $\eta > 0$ such that

$$1 - \alpha - \theta > (1 - eta + \theta)^{1-\eta}, \tag{35}$$

we define (see Fig. 1):

$$\mathcal{N}_\theta = \left\{\mathbf{x} \in \Delta_M^\circ : \alpha_x < \alpha + \theta, \; eta_x > eta - \theta\right\} \tag{36}$$

and

$$K_{\theta,\eta} := \left\{ \mathbf{x} \in \Delta_M^\circ \backslash \mathcal{N}_\theta : \max_{j \notin J^*} x(j) \geq \eta \right\}. \tag{37}$$

In the infinite population limit we have:

**Theorem 4.6** *Assume that $\mathbf{\Gamma}$ satisfies Assumption 4.4 and $\mathbf{X}_0^{(N)} \xrightarrow{P} \mathbf{x}_0$ for some $\mathbf{x}_0 \in \Delta_M^\circ$. Suppose in addition that there exist $\theta \in \left(0, \frac{\mathbf{e}ta - \alpha}{2}\right)$ and $\eta \in \left(0, 1 - \frac{\log(1 - \alpha - \theta)}{\log(1 - \mathbf{e}ta + \theta)}\right)$ (and, therefore, the inequality in (35) is satisfied) such that*

$$1 - \alpha - \theta > e^{-\frac{\varepsilon_\theta^2 c_T^2}{2}}, \tag{38}$$

*where the constant $c_T$ is introduced in (31), $\varepsilon_\theta = \varepsilon(\mathcal{N}_\theta, K_{\theta,\eta})$, $T = T(\mathcal{N}_\theta, K_{\theta,\eta})$, and $\varepsilon(\mathcal{N}, K)$ and $T(\mathcal{N}_\theta, K_{\theta,\eta})$ are introduced in Proposition 4.5.*

*Let*

$$\nu_N := \inf \left\{ k > 0 : \mathbf{X}_k^{(N)} \in \partial(\Delta_M) \right\} \tag{39}$$

*and $\mathcal{E}_N$ be the event that*

(1) *$\mathbf{X}_{\nu_N}^{(N)}$ has exactly $M - 1$ non-zero components (that is, $|C(\mathbf{X}_{\nu_N}^{(N)})| = M - 1$ in notation of (13)); and*
(2) *$X_{\nu_N}^{(N)}(i) = 0$ for some $i \in J^*$.*

*Then the following holds true:*

(i) *$\lim_{N \to \infty} P(\mathcal{E}_N) = 1$.*
(ii) *$\lim_{N \to \infty} E(\nu_N) = +\infty$.*

The proof of Theorem 4.6 is included in Sect. 5.5. In words, part (i) of the theorem is the claim that with overwhelming probability, at the moment that the Markov chain hits the boundary, exactly one of its components is zero, and, moreover, the index $i$ of the zero component $\mathbf{X}_{\nu_N}^{(N)}(i)$ belongs to the set $J^*$. A similar metastability phenomenon for continuous-time population models has been considered in Assaf and Meerson (2010) and Assaf and Mobilia (2011), Auletta et al. (2018), Park and Traulsen (2017), however their proofs rely on a mixture of rigorous computational and heuristic "physicist arguments".

Condition (38) is a rigorous version of the heuristic condition (34). We note that several more general but less explicit and verifiable conditions of this type can be found in mathematical literature, see, for instance, Faure and Schreiber (2014), Panageas et al. (2016) and Assaf and Meerson (2010), Park and Traulsen (2017). It seems plausible that the negative-definite condition (32) is an artifact of the proof and can be relaxed.

It is shown in Losert and Akin (1983) that in this more general situation, the mean-field dynamical system $\boldsymbol{\psi}_k$ still converges to an equilibrium which might be on the boundary and depends on the initial condition $\boldsymbol{\psi}_0$.

**Remark 1** Recall (13). A classical result of Karlin and McGregor (1964) implies that under the conditions of Theorem 4.6, the types vanish from the population with an exponential rate until only one left. That is, $\lim_{k\to\infty} \frac{1}{k} \log E(|C(\mathbf{X}_k^{(N)})| - 1)$ exists and is finite and strictly negative. However, heuristically, the limit is extremely small under the conditions of the theorem [the logarithm of the expected extinction time is anticipated to be of order $M \cdot N$ (Panageas et al. 2016)], and hence the convergence to boundary states is extremely slow for any practical purpose. The situation when the system escapes from a quasi-equilibrium fast but with a rate of escape quickly converging to zero when the population size increases to infinity, is typical for metastable population systems (Auletta et al. 2018; Broekman et al. 2019; Doorn and Pollett 2013; Méléard and Villemonais 2012).

We will next illustrate the (hypothesized) important role of the threshold condition (38) with a numerical example. To that end, we consider two examples (see matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, bellow) with $N = 500$ and $M = 3$. The choice of the low dimension $M = 3$ is due to the computational intensity involved with tracking the evolution of a metastable stochastic system in a "quasi-equilibrium" state. For each example, we use various sets of initial conditions and run 10000 rounds of simulations. Each simulation is stopped at the random time $T$ defined as follows:

$$T = \min\left\{k \in \mathbb{N} : \min_j X_k^{(N)}(j) \le 0.05\right\}. \tag{40}$$

Using this termination rule, we were able to complete all the simulation runs in a reasonable finite time. In each simulation run, we sample the value of the process at some random time ranging between 1000 and 5000. At this intermediate sampling time, we measure $d_{eq}$, the distance between the stochastic system and the equilibrium $\chi_{eq}$. In addition, we record the composition of the last state of the Markov chain at the end of each simulation run.

**Example 4** Our first example does not satisfy the threshold conditions. More precisely, we choose $\omega$ in (11) in such a way that $\frac{\omega}{1-\omega} = 10^{-3}$ and we consider the following symmetric $3 \times 3$ matrix $\mathbf{A}_1$ :

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 20 & 45 \\ 20 & 21 & 30 \\ 45 & 30 & 1 \end{pmatrix}$$

with the equilibrium $\chi_{eq}^{(1)} = (0.24766355, 0.41121495, 0.3411215)$. Table 1 reports the number of times that each type $j \in S_3$ was the type with abundance less than 0.05 at the time of exiting the simulation runs.

In addition, Fig. 2 below depicts the distribution of $d_{eq}$ different initial values. Evidently, the threshold condition plays an important role as the theory predicts in Theorem 4.6. The results in Table 1 cannot be explained by the equilibrium value $\chi_{eq}^{(1)}$ only, rather they constitute an intricate result of the combined effect of this value and the initial position of the Markov chain.

**Table 1** The least abundant component of the terminal state $\mathbf{X}_T^{(N)}$ by the end of the simulations for $\Gamma_i(\mathbf{x}) = \frac{1000 + A_1 x(i)}{1000 + \mathbf{x}^T \mathbf{A}_1 \mathbf{x}} x(i)$ and the quitting time $T$ introduced in (40)

|  | Number of times $\mathbf{X}_T^{(N)}(i) \leq 0.05$ | | |
|---|---|---|---|
|  | $i = 1$ | $i = 2$ | $i = 3$ |
| For $\mathbf{X}_0^{(N)} = (0.8, 0.1, 0.1)$ | 933 | 6831 | 2248 |
| For $\mathbf{X}_0^{(N)} = (0.1, 0.8, 0.1)$ | 5991 | 164 | 3903 |
| For $\mathbf{X}_0^{(N)} = (0.1, 0.1, 0.8)$ | 3692 | 5940 | 373 |

**Example 5** For the second example, we let $\omega = 1/2$ and use the following matrix $\mathbf{A}_2$:

$$\mathbf{A}_2 = \begin{pmatrix} 1 & 20 & 35 \\ 20 & 21 & 30 \\ 35 & 30 & 1 \end{pmatrix}$$

with the equilibrium $\chi_{eq}^{(2)} = (0.0246913, 0.7345679, 0.2407407)$. This matrix is almost the same as $\mathbf{A}_1$, with the only difference that 45 is changed to 35. The result of simulations for three different initial values are reported in Table 2 and Fig. 3, which are in complete agreement with the prediction of Theorem 4.6.

The difference between the two examples is two-fold: (1) by increasing the value of $\omega$ we increase the influence of the "selection matrix" $\mathbf{A}$ comparing to the neutral selection, and (2) by changing $\mathbf{A}_1$ to $\mathbf{A}_2$ we replace the equilibrium $\chi_{eq}^{(1)}$ with fairly uniform distribution of "types" by the considerably less balanced $\chi_{eq}^{(2)}$, increasing the threshold in (38) by a considerable margin. We note that in a general, and not specifically designed, example, it may be hard to verify conditions of the theorem, particularly (38).

Next we provide an auxiliary "contraction near the equilibrium" type result, further illustrating the mechanism of convergence to equilibrium in Theorem 4.6. Formally, the result is a slight improvement of a classical stability theorem of Losert and Akin (1983) (see Theorem 4 in "Appendix B" of this paper) showing that the unique interior equilibrium of the dynamical system $\psi_k$ introduced in (3) is in fact exponentially stable.

**Proposition 4.7** *Let Assumption* 4.4 *hold. Then the spectral radius of the Jacobian (Frechét derivative in $\Delta_M$) of $\Gamma$ at $\chi_{eq}$ is strictly less than one.*

The proof of the proposition is included in Sect. 5.6. This result is instrumental in understanding various applications including but not limited to mixing times (Panageas et al. 2016) and the random sequence $\mathbf{u}_k$ defined in the statement of Theorem 4.1. In the later case, as we mentioned before, it implies that the limiting sequence $\mathbf{U}_k$ in (27) converges to a stationary distribution.
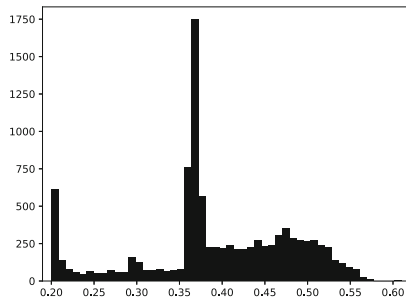
The following corollary is an application directly related to the metastability of the system. Let $\mu_N$ be the *quasi-stationary distribution* of the Markov chain $\mathbf{X}_k^{(N)}$ on
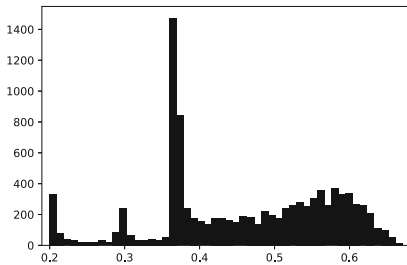
(a) $d_{eq}$ for $\mathbf{X}_0^{(N)} = (0.8, 0.1, 0.1)$       (b) $d_{eq}$ for $\mathbf{X}_0^{(N)} = (0.1, 0.8, 0.1)$
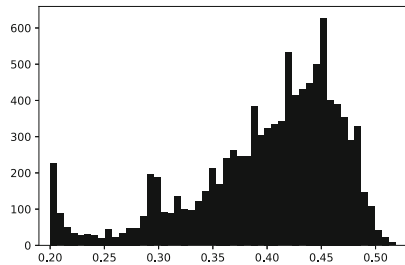


(c) $d_{eq}$ for $\mathbf{X}_0^{(N)} = (0.1, 0.1, 0.8)$

**Fig. 2** Simulations of the trajectories of $\psi_k$ and $\mathbf{X}_k^{(N)}$ for $\Gamma_i(\mathbf{x}) = \frac{1000 + A_1 x(i)}{1000 + \mathbf{x}^T \mathbf{A}_1 \mathbf{x}} x(i)$. The $x$-axis represents $d_{eq}$, and the height of a histogram bar corresponds to the number of occurrences of $d_{eq}$ in a simulation of 10,000 runs

$\Delta_M^\circ$. Then there exists $\lambda_N \in (0, 1)$ such that $\mu_N P_{\circ,N} = \lambda_N \mu_N$, where $P_{\circ,N}$ is the transition matrix of $\mathbf{X}_k^{(N)}$ restricted to $\Delta_{M,N}^\circ$ (Doorn and Pollett 2013; Méléard and Villemonais 2012). Iterating, we obtain that for all $k \in \mathbb{Z}_+$ and $\mathbf{y} \in \Delta_{M,N}^\circ$,

$$\mu_N P_{\circ,N}^k(\mathbf{y}) = \sum_{\mathbf{x} \in \Delta_{M,N}^\circ} \mu_N(\mathbf{x}) P\big(\mathbf{X}_k^{(N)} = \mathbf{y} \mid \mathbf{X}_0^{(N)} = \mathbf{x}\big) = \lambda_N^k \mu_N(\mathbf{y}). \qquad (41)$$

Summing up over all $\mathbf{y} \in \Delta_{M,N}^\circ$, we obtain

$$\lambda_N^k = P_{\mu_N}(\nu_N > k), \qquad (42)$$

where $\nu_N$ is the hitting time introduced in (39). Heuristically, (42) is equivalent to the condition $\mu_N(B) = P_{\mu_N}\big(\mathbf{X}_k^{(N)} \in B \mid \nu_N > k\big) = \frac{P_{\mu_N}(\mathbf{X}_k^{(N)} \in B)}{P_{\mu_N}(\nu_N > k)}$ for all $B \subset \Delta_{M,N}^\circ$, which often serves as an alternative definition of the quasi-stationary distribution.

**Table 2** The least abundant component of the terminal state $\mathbf{X}_T^{(N)}$ by the end of the simulations for $\Gamma_i(\mathbf{x}) = \frac{1+A_2 x(i)}{1+\mathbf{x}^T A_2 \mathbf{x}} x(i)$ and the quitting time $T$ introduced in (40)
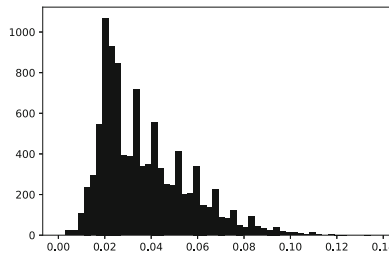
| | Number of times $\mathbf{X}_T^{(N)}(i) \leq 0.05$ | | |
| | --- | --- | --- |
| | $i = 1$ | $i = 2$ | $i = 3$ |
| For $\mathbf{X}_0^{(N)} = (0.8, 0.1, 0.1)$ | 10,000 | 0 | 0 |
| For $\mathbf{X}_0^{(N)} = (0.1, 0.8, 0.1)$ | 10,000 | 0 | 0 |
| For $\mathbf{X}_0^{(N)} = (0.1, 0.1, 0.8)$ | 10,000 | 0 | 0 |
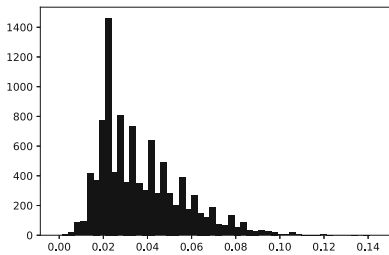


(a) $d_{eq}$ for $\mathbf{X}_0^{(N)} = (0.8, 0.1, 0.1)$
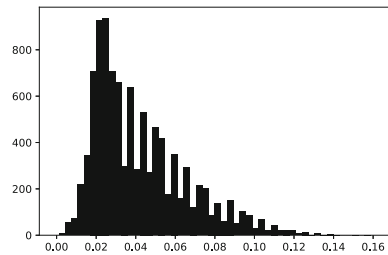
(b) $d_{eq}$ for $\mathbf{X}_0^{(N)} = (0.1, 0.8, 0.1)$



(c) $d_{eq}$ for $\mathbf{X}_0^{(N)} = (0.1, 0.1, 0.8)$

**Fig. 3** Simulations of the trajectories of $\psi_k$ and $\mathbf{X}_k^{(N)}$ for $\Gamma_i(\mathbf{x}) = \frac{1+A_2 x(i)}{1+\mathbf{x}^T A_2 \mathbf{x}} x(i)$. The $x$-axis represents $d_{eq}$, and the height of a histogram bar corresponds to the number of occurrences of $d_{eq}$ in a simulation of 10,000 runs

In particular, (42) implies that

$$1 - \lambda_N = \sum_{\mathbf{x} \in \Delta_{M,N}^\circ, y \in \partial(\Delta_{M,N})} \mu_N(\mathbf{x}) P_N(\mathbf{x}, \mathbf{y}). \tag{43}$$

**Corollary 4.8** *Let Assumption 4.4 hold. Then there exist constants $C_1 > 0$ and $C_2 > 0$ such that $\lambda_N \geq 1 - C_1 e^{-C_2 N}$ for all $N \in \mathbb{N}$.*

A short proof of the corollary is included in Sect. 5.7 below. The result in the corollary and the comparison of (43) with (58) in the proof of Theorem 4.6 given below, motivate the following hypothesis:

**Conjecture 4.9** Let the conditions of Theorem 4.6 hold. Then:

(i) The quasi-stationary distribution $\mu_N$ converges, as $N \to \infty$, to the degenerate distribution supported on the single point $\chi_{eq}$. That is, $\lim_{N \to \infty} \mu_N(\mathcal{N}) = 1$ for any open neighborhood of the equilibrium $\chi_{eq}$.

(ii) $\lim_{N \to \infty} \frac{\alpha^N}{1-\lambda_N} = 1$.

The first part of Conjecture 4.9 is a natural property of stochastic models obtained as a perturbation of asymptotically stable deterministic dynamical system (cf. Faure and Schreiber 2014). The second part of the conjecture is related to the first one as follows. Recall that $\alpha^N = P_{\chi_{eq}}(\mathcal{E}_N)$ and $1 - \lambda_N = P_{\mu_N}(\nu_N = 1)$. If $\mu_N$ indeed converges weakly to the degenerate probability measure supported on $\chi_{eq}$, then one can reasonably expect that $\alpha = P_{\chi_{eq}}(\mathcal{E}_N) \sim P_{\mu_N}(\mathcal{E}_N)$ as $N \to \infty$. The assertion that $P_{\chi_{eq}}(\mathcal{E}_N) \sim P_{\mu_N}(\mathcal{E}_N) \sim P_{\mu_N}(\nu_N = 1)$, if true, would be an expression of the underlying principle that under the conditions of Theorem 4.6 and when $N$ is large, the overwhelming likely scenario for the system to move to the boundary of the simplex $\Delta_M$ is to return to a small neighborhood of $\chi_{eq}$ and move from there directly to one of the boundary sites which realize the event $\mathcal{E}_N$.

A weak form of the hypothesis that $\lim_{N \to \infty} \frac{\alpha^N}{1-\lambda_N} = 1$ is

**Conjecture 4.10** Let the conditions of Theorem 4.6 hold. Then, $\lim_{N \to \infty} \frac{1}{N} \log(1 - \lambda_N) = \alpha$.

The abstract quantitative difference between the two conjectures can be illustrated, for instance, by considering the numerical sequences $a_N = Ne^N$ and $b_N = e^N$. While it is true that $\lim_{N \to \infty} \frac{\log a_N}{\log b_N} = \lim_{N \to \infty} \frac{1}{N} \log(Ne^N) = 1$, the corresponding strong version of this identity $\lim_{N \to \infty} \frac{a_N}{b_N} = \lim_{N \to \infty} \frac{Ne^N}{e^N} = 1$ fails to hold.

# 5 Proofs

This section is devoted to the proof of our main results stated in Sect. 4. As before, transition matrix of the Markov chain $\mathbf{X}^{(N)}$ is denoted by $P_N$. We write $P_{\mu_N}(A_N)$ for the probability of an event $A_N$ measurable with respect to the $\sigma$-algebra generated by $(\mathbf{X}_k^{(N)})_{k \geq 0}$ if $\mu_N$ is the distribution of $\mathbf{X}_0^{(N)}$. If $\mu_N$ is concentrated on a single point $\mathbf{x} \in \Delta_{M,N}$, we simplify this notation to $P_{\mathbf{x}}(A_N)$. We use the standard notations $f_n = o(g_n)$ and $f_n = O(g_n)$ to indicate that, respectively, the sequence $f_n$ is "ultimately smaller" than $g_n$, that is $\lim_{n \to \infty} \frac{f_n}{g_n} = 0$, and $f_n$ and $g_n$ are of the same order, that is $\lim_{n \to \infty} \frac{f_n}{g_n} = $ exists and it is not zero.

### 5.1 Proof of Theorem 4.1

Let $\zeta := \left\{ \zeta_{k,i}^{(\mathbf{x})} : \mathbf{x} \in \Delta_M, i \in \mathbb{N}, k \in \mathbb{Z}_+ \right\}$ be a collection of independent random variables distributed as follows:

$$P\left(\zeta_{k,i}^{(\mathbf{x})} = j\right) = \Gamma_j(\mathbf{x}), \qquad j \in S_M. \tag{44}$$

For $\mathbf{x} \in \Delta_M, i \in \mathbb{N}, k \in \mathbb{Z}_+$, let $\boldsymbol{\xi}_{k,i}^{(\mathbf{x})}$ be a random $M$-dimensional vector such that

$$\xi_{k,i}^{(\mathbf{x})}(j) = \begin{cases} 1 \text{ if } \zeta_{k,i}^{(\mathbf{x})} = j, \\ 0 \text{ otherwise.} \end{cases}$$

For given parameters $k, i$, the family of random vectors $\left\{ \boldsymbol{\xi}_{k,i}^{(\mathbf{x})} : \mathbf{x} \in \Delta_M \right\}$ is a "white noise" random field. Without loss of generality we can assume that

$$\mathbf{Z}_{k+1}^{(N)} = \sum_{i=1}^{N} \boldsymbol{\xi}_{k,i}^{(\mathbf{X}_k^{(N)})}. \tag{45}$$

Thus, by the law of large numbers, the conditional distribution of $\mathbf{X}_{k+1}^{(N)} = \frac{\mathbf{Z}_{k+1}^{(N)}}{N}$, given $\mathbf{X}_k^{(N)}$, converges weakly to a degenerate distribution concentrated on the single point $\boldsymbol{\Gamma}\left(\mathbf{X}_k^{(N)}\right)$. This implies part (a) of the theorem by a result of Karr (1975) adopted to a non-homogeneous Markov chain setting [cf. Remark (i) in Buckley and Pollett (2010, p. 60)].

To prove part (b) of the theorem, write

$$\mathbf{u}_{k+1}^{(N)} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left( \boldsymbol{\xi}_{k,i}^{(\mathbf{x})} - \boldsymbol{\Gamma}(\mathbf{X}_k^{(N)}) \right) + \sqrt{N} \left( \boldsymbol{\Gamma}(\mathbf{X}_k^{(N)}) - \boldsymbol{\psi}_{k+1} \right).$$

The first term converges in distribution, as $N$ goes to infinity, to the degenerate Gaussian distribution $g_k$ defined in the statement of the theorem [see, for instance, Theorem 14.6 in Wasserman (2004)]. Heuristically, the Gaussian limit of a sequence of multinomial random vectors $\mathbf{X_k}$ is degenerate because of the constraint $\|X_k\| = 1$ which propagates into a minus one degree of freedom in the limit.

Furthermore, Taylor expansion for $\boldsymbol{\Gamma}$ gives

$$\sqrt{N}\left(\boldsymbol{\Gamma}(\mathbf{X}_k^{(N)}) - \boldsymbol{\psi}_{k+1}\right) = \sqrt{N}\left(\mathbf{D}_x(\boldsymbol{\psi}_k)(\mathbf{X}_k^{(N)} - \boldsymbol{\psi}_k) + O(1/N)\right)$$
$$= \mathbf{D}_x(\boldsymbol{\psi}_k)\mathbf{u}_k^{(N)} + o(1),$$

which implies the result. □

## 5.2 Proof of Theorem 4.2

We have

$$
\begin{aligned}
\left\| \mathbf{X}_k^{(N)} - \boldsymbol{\psi}_k \right\| &= \left\| \mathbf{X}_k^{(N)} - \boldsymbol{\Gamma}(\mathbf{X}_{k-1}^{(N)}) + \boldsymbol{\Gamma}(\mathbf{X}_{k-1}^{(N)}) - \boldsymbol{\Gamma}(\boldsymbol{\psi}_{k-1}) \right\| \\
&\leq \left\| \mathbf{X}_k^{(N)} - \boldsymbol{\Gamma}(\mathbf{X}_{k-1}^{(N)}) \right\| + \left\| \boldsymbol{\Gamma}(\mathbf{X}_{k-1}^{(N)}) - \boldsymbol{\Gamma}(\boldsymbol{\psi}_{k-1}) \right\| \\
&\leq \left\| \mathbf{X}_k^{(N)} - \boldsymbol{\Gamma}(\mathbf{X}_{k-1}^{(N)}) \right\| + \rho \left\| \mathbf{X}_{k-1}^{(N)} - \boldsymbol{\psi}_{k-1} \right\|.
\end{aligned}
\tag{46}
$$

For $m \in \mathbb{N}$, we set $\mathbf{u}_m^{(N)} := \mathbf{X}_{m+1}^{(N)} - \boldsymbol{\Gamma}(\mathbf{X}_m^{(N)})$. Iterating the above inequality, we obtain that

$$
\left\| \mathbf{X}_k^{(N)} - \boldsymbol{\psi}_k \right\| \leq \sum_{m=0}^{k-1} \rho^m \left\| \mathbf{u}_{k-m-1}^{(N)} \right\| \leq c_k^{-1} \max_{0 \leq m \leq k-1} \left\| \mathbf{u}_m^{(N)} \right\|,
\tag{47}
$$

where $c_k$ is defined in (31). Therefore,

$$
\begin{aligned}
P \left( \max_{1 \leq k \leq K} \left\| \mathbf{X}_k^{(N)} - \boldsymbol{\psi}_k \right\| > \varepsilon \right) &\leq P \left( \max_{0 \leq m \leq K-1} \left\| \mathbf{u}_m^{(N)} \right\| > \varepsilon c_K \right) \\
&\leq \sum_{m=0}^{K-1} P \left( \left\| \mathbf{u}_m^{(N)} \right\| > \varepsilon c_K \right) \\
&= \sum_{m=0}^{K-1} P \left( \left\| \mathbf{X}_{m+1}^{(N)} - \boldsymbol{\Gamma}(\mathbf{X}_m^{(N)}) \right\| > \varepsilon c_K \right).
\end{aligned}
\tag{48}
$$

Given $\mathbf{X}_m^{(N)}$, the conditional distribution of the random vector $\mathbf{Z}_{m+1}^{(N)} = N \mathbf{X}_{m+1}^{(N)}$ is the multinomial $\mathcal{M}(N, \boldsymbol{\Gamma}(\mathbf{X}_m^{(N)}))$. Therefore, applying Hoeffding's inequality (Hoeffding 1963) to conditionally binomial random variables $X_{m+1}(i)$, we obtain that for any $\eta > 0$,

$$
\begin{aligned}
P \left( \left\| \mathbf{X}_{m+1}^{(N)} - \boldsymbol{\Gamma}(\mathbf{X}_m^{(N)}) \right\| > \eta \right) &\leq \sum_{i=1}^{M} P \left( |X_{m+1}^{(N)}(i) - \Gamma_i(\mathbf{X}_m^{(N)})| > \eta \right) \\
&\leq 2M \exp \left( -\frac{N\eta^2}{2} \right).
\end{aligned}
\tag{49}
$$

Therefore,

$$
P \left( \max_{1 \leq k \leq K} \left\| \mathbf{X}_k^{(N)} - \boldsymbol{\psi}_k \right\| > \varepsilon \right) \leq 2KM \exp \left( -\frac{\varepsilon^2 c_K^2 N}{2} \right).
\tag{50}
$$

To complete the proof observe that $c_K = \frac{1-\rho}{1-\rho^K} > 1 - \rho$ if $\rho < 1$.     $\square$

### 5.3 Proof of Theorem 4.3

Recall $\tau_N(\varepsilon)$ from (28). It follows from (46) that for all $k \in \mathbb{Z}_+$ and $\varepsilon > 0$,

$$\left\{\|\mathbf{X}_{k+1}^{(N)} - \boldsymbol{\psi}_{k+1}\| > \varepsilon\right\} \subseteq \left\{\|\mathbf{X}_{k+1}^{(N)} - \Gamma(\mathbf{X}_k^{(N)})\| > (1-\rho)\varepsilon\right\} \bigcup \left\{\|\mathbf{X}_k^{(N)} - \boldsymbol{\psi}_k\| > \varepsilon\right\}.$$

and hence

$$\left\{\tau_N(\varepsilon) = k+1\right\} \subseteq \left\{\|\mathbf{X}_{k+1}^{(N)} - \boldsymbol{\psi}_{k+1}\| > \varepsilon\right\} \bigcap \left\{\|\mathbf{X}_k^{(N)} - \boldsymbol{\psi}_k\| \leq \varepsilon\right\}$$
$$\subseteq \left\{\|\mathbf{X}_{k+1}^{(N)} - \Gamma(\mathbf{X}_k^{(N)})\| > (1-\rho)\varepsilon\right\}. \tag{51}$$

Furthermore, using again (47) and Hoeffding's inequality, one can verify that the analogues of (48) and subsequently of (49) hold almost surely for the conditional probability $P(\cdot \mid \mathbf{X}_k^{(N)})$. It follows then from the latter that

$$P\left(\|\mathbf{X}_{k+1}^{(N)} - \Gamma(\mathbf{X}_k^{(N)})\| > (1-\rho)\varepsilon \mid \mathbf{X}_k^{(N)}\right) \leq 2M \exp\left(-\frac{N(1-\rho)^2\varepsilon^2}{2}\right), \qquad P - a.s.$$

Combining the last inequality with (51) and using the strong Markov property of the Markov chain $\mathbf{X}^{(N)}$, we obtain that with probability one,

$$P\left(\tau_N(\varepsilon) > k+1 \mid \mathbf{X}_k^{(N)}, \tau_N(\varepsilon) > k\right) \geq 1 - 2M \exp\left(-\frac{N(1-\rho)^2\varepsilon^2}{2}\right).$$

Applying the tower property of conditional expectations $E(X) = E\big(E(X \mid Y)\big)$ to the indicator of the event $\{\tau_N(\varepsilon) > k+1\}$ in the role of $X$, $\mathbf{X}_k^{(N)}$ in the role of $Y$, and the conditional expectation $E(\cdot \mid \tau_N(\varepsilon) > k)$ as $E$, we deduce from the last inequality that

$$\begin{aligned}
P(\tau_N(\varepsilon) > k+1) &= P\big(\tau_N(\varepsilon) > k+1, \tau_N(\varepsilon) > k\big) \\
&= P\big(\tau_N(\varepsilon) > k+1 \mid \tau_N(\varepsilon) > k\big) P(\tau_N(\varepsilon) > k) \\
&= E\Big( P\big(\tau_N(\varepsilon) > k+1 \mid \mathbf{X}_k^{(N)}, \tau_N(\varepsilon) > k\big) \mid \tau_N(\varepsilon) > k \Big) P(\tau_N(\varepsilon) > k) \\
&\geq \Big(1 - 2M \exp\Big(-\frac{N(1-\rho)^2\varepsilon^2}{2}\Big)\Big) P(\tau_N(\varepsilon) > k),
\end{aligned}$$

from which the claim of the theorem follows by induction. $\qquad\square$

### 5.4 Proof of Proposition 4.5

Recall $\boldsymbol{\psi}_k$ from (3). For $\mathcal{N} \subset \Delta_M$, denote by $\tau(\mathbf{x}, \mathcal{N})$ the first hitting time of the set $\mathcal{N}$ by $\boldsymbol{\psi}_k$ with $\boldsymbol{\psi}_0 = \mathbf{x}$ :

$$\tau(\mathbf{x}, \mathcal{N}) = \inf\{k \in \mathbb{Z}_+ : \boldsymbol{\psi}_k(\mathbf{x}) \in \mathcal{N}\}.$$

Further, let $t(\mathbf{x}, \mathcal{N})$ be the time when $\boldsymbol{\psi}_k$ becomes trapped within $\mathcal{N}$ :

$$t(\mathbf{x}, \mathcal{N}) = \inf\{n \in \mathbb{Z}_+ : \boldsymbol{\psi}_k(\mathbf{x}) \in \mathcal{N} \ \forall k \geq n\}.$$

In both the definitions we use the regular convention that $\inf \emptyset = +\infty$. Note that, since $\boldsymbol{\psi}_k$ converges to $\boldsymbol{\chi}_{eq}$, both $\tau(\mathbf{x}, \mathcal{N})$ and $t(\mathbf{x}, \mathcal{N})$ are finite for all $\mathbf{x} \in \Delta_M^\circ$.

**Lemma 5.1** *For any open neighborhood $\mathcal{N}$ of $\boldsymbol{\chi}_{eq}$, there is an open neighborhood $U_{\mathbf{x}}$ of $\mathbf{x}$ and a constant $T_{\mathbf{x}} \in \mathbb{N}$ such that $t(\mathbf{y}, \mathcal{N}) \leq T_{\mathbf{x}}$ for all $\mathbf{y} \in U_{\mathbf{x}}$.*

***Proof of Lemma 5.1*** Under the conditions stated in Assumption 4.4, $\boldsymbol{\Gamma}$ is a diffeomorphism (Losert and Akin 1983). In particular, $\boldsymbol{\Gamma}$ is a locally Lipschitz map. Hence, we may apply Theorem 1.2 in Bof et al. (2018) to conclude that the interior equilibrium $\boldsymbol{\chi}_{eq}$ is asymptotically stable for the dynamical system $\boldsymbol{\psi}_k$. That is, for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\|\boldsymbol{\psi}_0 - \boldsymbol{\chi}_{eq}\| \leq \delta \quad \Longrightarrow \quad \|\boldsymbol{\psi}_k - \boldsymbol{\chi}_{eq}\| \leq \varepsilon \ \forall k \geq 0. \tag{52}$$

Equivalently, for any open neighborhood $\mathcal{N}$ of $\boldsymbol{\chi}_{eq}$ there exists an open neighborhood of $\boldsymbol{\chi}_{eq}$, say $\mathcal{U}$, included in $\mathcal{N}$ such that $\boldsymbol{\Gamma}(\mathcal{U}) \subset \mathcal{U}$ (that is $\mathcal{U}$ is forward-invariant for $\boldsymbol{\Gamma}$).

Let $\mathcal{U}$ be a forward-invariant open neighborhood of $\boldsymbol{\chi}_{eq}$ included in $\mathcal{N}$. Since $\mathcal{U}$ is an open set and $\boldsymbol{\psi}_{\tau(\mathbf{x}, \mathcal{U})}(\mathbf{y})$ is a continuous function of the initial state $\mathbf{y}$, one can choose an open neighborhood $V_{\mathbf{x}}$ of $\boldsymbol{\psi}_{\tau(\mathbf{x}, \mathcal{U})}(\mathbf{x})$ and a constant $\varepsilon_{\mathbf{x}} > 0$ such that $V_{\mathbf{x}} \subset \mathcal{U}$ and, furthermore, $\|\mathbf{y} - \mathbf{x}\| < \varepsilon_{\mathbf{x}}$ implies $\boldsymbol{\psi}_{\tau(\mathbf{x}, \mathcal{U})}(\mathbf{y}) \in V_{\mathbf{x}}$. Set $U_{\mathbf{x}} = \left\{ \mathbf{y} \in \Delta_M^\circ : \|\mathbf{y} - \mathbf{x}\| < \varepsilon_{\mathbf{x}} \right\}$. $\qquad\square$

**Corollary 5.2** *Let $\mathcal{N}$ be any open neighborhood of $\chi_{eq}$ and $K$ be a compact subset of $\Delta_M^\circ$. Then there exists $T = T(K, \mathcal{N}) \in \mathbb{N}$ such that $t(\mathbf{x}, \mathcal{N}) \leq T$ for all $\mathbf{x} \in K$.*

***Proof of Corollary 5.2*** Let $U_{\mathbf{x}}$ and $T_{\mathbf{x}}$ be as in Lemma 5.1. Consider a cover of $K$ by the union of open sets $\bigcup_{x \in K} U_{\mathbf{x}}$. Since $K$ is compact, we can choose a finite subcover, say $\bigcup_{j=1}^m U_{\mathbf{x}_j}$. Set $T(K, \mathcal{N}) = \max_{1 \leq j \leq m} t(\mathbf{x}_j, \mathcal{N})$. $\qquad\square$

**Remark 2** The claim is a "uniform shadowing property" of $\boldsymbol{\Gamma}$. Recall that a sequence of points $(\mathbf{x}_k)_{k \leq T}$ where $T \in \mathbb{N} \bigcup \{+\infty\}$ is a $\delta$-pseudo orbit of the dynamical system $\boldsymbol{\psi}_k$ if $\|\boldsymbol{\Gamma}(\mathbf{x}_k) - \mathbf{x}_{k+1}\| \leq \delta$ for $k = 0, \ldots, T - 1$ (if $T$ is finite then the definitions of $\delta$-pseudo orbit and $\delta$-chain are equivalent) and that a $\delta$-pseudo orbit is said to be $\varepsilon$-shadowed by a true orbit $\left(\boldsymbol{\psi}_k(\mathbf{x})\right)_{k \in \mathbb{Z}_+}$ if $\|\boldsymbol{\psi}_k(\mathbf{x}) - \mathbf{x}_k\| \leq \varepsilon$ for all $k \geq 0$. The proposition is basically saying that since $\boldsymbol{\psi}_k$ is asymptotically stable, then 1) any infinite $\delta$-pseudo orbit starting at $\mathbf{x}_0$ is $\varepsilon$-shadowed by the true orbit $\left(\boldsymbol{\psi}_k(\mathbf{x}_0)\right)_{k > 0}$ provided that $\delta > 0$ is small enough; and 2) the property (more precisely, a possible value of the parameter $\delta$ given $\varepsilon > 0$) is uniform over compact sets with respect to the initial point $\mathbf{x}_0$.

### 5.5 Proof of Theorem 4.6

**Proof of (i)** Let $\mathbf{1}_A$ denote the indicator of the event $A$, that is $\mathbf{1}_A = 1$ if $A$ occurs and $\mathbf{1}_A = 0$ otherwise. For $\theta$ introduced in the statement of the theorem and $\mathcal{N}_\theta$ defined in (36), let $L_N$ be the number of times the Markov chain $\mathbf{X}^{(N)}$ visits $\mathcal{N}_\theta$ *after time zero* and before it hits the boundary $\partial(\Delta_M)$ for the first time. That is,

$$L_N = \sum_{k=1}^{\nu_N} \mathbf{1}_{\{\mathbf{X}_k^{(N)} \in \mathcal{N}_\theta\}}, \tag{53}$$

where $\nu_N$ is the first hitting time of the boundary introduced in (39). Notice that time zero is not counted in the above formula regardless of whether or not $\mathbf{X}_0^{(N)} \in \mathcal{N}_\theta$. Since for a fixed $N \in \mathbb{N}$, the state space of the Markov chain $\mathbf{X}^{(N)}$ is finite and

$$P_N\big(\mathbf{x}, \partial(\Delta_{M,N})\big) = \sum_{\mathbf{y} \in \partial(\Delta_{M,N})} P_N(\mathbf{x}, \mathbf{y}) > 0$$

for all $\mathbf{x} \in \Delta_{M,N}$, we have:

$$P_\mathbf{x}(0 \le L_N < \infty) = P_\mathbf{x}(1 \le \nu_N < \infty) = 1 \qquad \forall \mathbf{x} \in \Delta_{M,N}^\circ.$$

Note that in (53) we are counting the number of *visits after time zero* to $\mathcal{N}_\theta$, those excluding the initial state $\mathbf{X}_0^{(N)}$ from the account even if $\mathbf{X}_0^{(N)} \in \mathcal{N}_\theta$.

Denote

$$G_N = \{L_N > 0\}, \qquad \overline{G}_N = \{L_N = 0\},$$

and

$$\xi_N = \inf\{k \ge 0 : \mathbf{X}_k^{(N)} \in \mathcal{N}_\theta\}. \tag{54}$$

That is $G_N$ is the event that $\mathbf{X}^{(N)}$ visits $\mathcal{N}_\theta$ after time zero at least once, $\overline{G}_N$ is its complement, and $\xi_N$ is the time of the first visit (rather than the first return, which means $\mathbf{X}_0^{(N)}$ counts in (54) if it belongs to $\mathcal{N}_\theta$). Note that $P_\mathbf{x}(\xi_N = +\infty) > 0$ for all $\mathbf{x} \in \Delta_{M,N} \backslash \mathcal{N}_\theta$ because the chain can escape to the boundary of the simplex before it visits $\mathcal{N}_\theta$, and the boundary is an absorbing set for $\mathbf{X}^{(N)}$.

Let $\theta > 0$ and $\eta > 0$ be as in the statement of the theorem, and recall $K_{\theta,\eta}$ from (37). Proposition 4.5 implies that there exist (deterministic, non-random) real constant $\varepsilon_\theta > 0$ and integer $T > 0$ such that

$$\mathbf{X}_0^{(N)} \in K_{\theta,\eta} \text{ and } \big\|\mathbf{X}_{k+1}^{(N)} - \Gamma\big(\mathbf{X}_k^{(N)}\big)\big\| < \varepsilon_\theta \text{ for } k \le T-1 \quad \Rightarrow \quad \mathbf{X}_T^{(N)} \in \mathcal{N}_\theta.$$

The vector field $\mathbf{\Gamma}$ satisfies the conditions of Theorem 4.2, and it follows from (30) that

$$P_{\mathbf{x}}(\overline{G}_N) \leq 2T M e^{-\frac{N \varepsilon_\theta^2 c_T^2}{2}} \qquad \forall\, \mathbf{x} \in K_{\theta,\eta}, \tag{55}$$

where $c_T$ is defined in (31). For $\mathbf{x} \in K_{\theta,\eta}$, write

$$P_{\mathbf{x}}(\mathcal{E}_N) = P_{\mathbf{x}}(\mathcal{E}_N; G_N) + P_{\mathbf{x}}(\mathcal{E}_N; \overline{G}_N),$$

and notice that $\lim_{N\to\infty} P_{\mathbf{x}}(\overline{G}_N) = 0$ by virtue of (55). Moreover,

$$P_{\mathbf{x}}(\mathcal{E}_N; G_N) = P_{\mathbf{x}}(\mathcal{E}_N \mid G_N) P_{\mathbf{x}}(G_N) = P_{\phi_{N,\mathbf{x}}}(\mathcal{E}_N) P_{\mathbf{x}}(G_N), \tag{56}$$

where $\phi_{N,\mathbf{x}}$ is the distribution of $\mathbf{X}_{\xi_N}^{(N)}$ under the conditional measure $P_{\mathbf{x}}(\cdot \mid G_N)$, with $\xi_N$ introduced in (54). Hence, in order to prove the first part of the theorem, it suffices to show that for any sequence of (discrete) probability measures $\phi_N$, each supported on $\mathcal{N}_\theta \bigcap \Delta_{M,N}$ with the corresponding $N$, we have

$$\lim_{N\to\infty} P_{\phi_N}(\mathcal{E}_N) = 1.$$

Since

$$P_{\phi_N}(\mathcal{E}_N) = \sum_{\mathbf{y} \in \mathcal{N}_\theta \bigcap \Delta_{M,N}} P_{\mathbf{y}}(\mathcal{E}_N)\, \phi_N(\mathbf{y}),$$

it suffices to show that

$$\lim_{N\to\infty} P_{\mathbf{y}}(\mathcal{E}_N) = 1,$$

and the convergence is uniform over $\mathbf{y} \in \mathcal{N}_\theta$. More precisely,

**Lemma 5.3** *For any $\varepsilon > 0$ there exists an integer $N_0 = N_0(\varepsilon) > 0$ such that*

$$P_{\mathbf{y}}(\mathcal{E}_N) \geq 1 - \varepsilon$$

*for all $N > N_0$ and $\mathbf{y} \in \mathcal{N}_\theta \bigcap \Delta_{M,N}$.*

**Proof of Lemma 5.3** Recall $L_N$ from (53). If $L_N \geq 1$, let $S_{1,N}, \ldots, S_{L_N,N}$ be successive return times of the Markov chain $\mathbf{X}^{(N)}$ to $\mathcal{N}_\theta$. Namely, we set $S_{0,N} = 0$ and define recursively,

$$S_{k,N} = \inf\{j > S_{k-1,N} : \mathbf{X}_j^{(N)} \in \mathcal{N}_\theta\}.$$

As usual, we assume that $\inf \emptyset = +\infty$. Let $S_{L_N+1,N}$ be the first hitting time of the boundary:

$$S_{L_N+1,N} = \nu_N,$$

where $\nu_N$ is defined in (39). Finally, for $k \in \mathbb{Z}_+$ define

$$\mathbf{Y}_k^{(N)} := \begin{cases} \mathbf{X}_{S_{k,N}}^{(N)} & \text{if } k \leq L_N + 1 \\ \mathbf{X}_{S_{L_N+1,N}}^{(N)} & \text{if } k > L_N + 1. \end{cases}$$

Then $\left(\mathbf{Y}_j^{(N)}\right)_{j\in\mathbb{N}}$ is a Markov chain on $\left(\mathcal{N}_\theta \bigcup \partial(\Delta_M)\right) \bigcap \Delta_{M,N}$ with all states on the boundary $\partial\left(\Delta_{M,N}\right)$ being absorbing. Denote by $Q_N$ transition kernel of this chain.

Recall (4) and (13), and define

$$\partial_{M,\Gamma}^* = \left\{\mathbf{x} \in \partial_M : |C(\mathbf{x})| = M - 1 \text{ and } x(i) = 0 \text{ for some } i \in J^*\right\}. \tag{57}$$

In words, $\partial_{M,\Gamma}^*$ is precisely the subset of the boundary such that for all $N \in \mathbb{N}$, the event $\mathcal{E}_N$ occurs if and only if $\mathbf{X}_{\nu_N}^{(N)} \in \partial_{M,\Gamma}^*$, where $\nu_N$ is the first hitting time of the boundary defined in (39). We then have for any $\mathbf{x} \in \mathcal{N}_\theta \bigcap \Delta_{M,N}$ :

$$\begin{aligned} P_{\mathbf{x}}(\mathcal{E}_N) &\geq \sum_{m=0}^{\infty} \sum_{\mathbf{y}\in\mathcal{N}_\theta \bigcap \Delta_{M,N}} \sum_{\mathbf{z}\in\partial_{M,\Gamma}^*} Q_N^m(\mathbf{x}, \mathbf{y}) P_N(\mathbf{y}, \mathbf{z}) \\ &\geq (1 - \alpha - \theta)^N \sum_{m=0}^{\infty} \sum_{\mathbf{y}\in\mathcal{N}_\theta \bigcap \Delta_{M,N}} Q_N^m(\mathbf{x}, \mathbf{y}) \\ &= (1 - \alpha - \theta)^N \sum_{k=0}^{\infty} P_{\mathbf{x}}(L_N \geq k) \\ &= (1 - \alpha - \theta)^N \left(1 + E_{\mathbf{x}}(L_N)\right), \end{aligned} \tag{58}$$

where $L_N$ is the number of visits to $\mathcal{N}_\theta$ after time zero introduced in (53). In the above formula, $P_{\mathbf{x}}(\mathcal{E}_N)$ is first bounded from below by the probability of the event "$\mathbf{X}^{(N)}$ hits the boundary at one of the vertices of the set $\partial_{M,\Gamma}^*$, moreover, the last site visited by $\mathbf{X}^{(N)}$ before passing to the boundary lies within $\mathcal{N}_\theta$". The index $m$ in (58) counts the number of visits to $\mathcal{N}_\theta$ after time zero, $\mathbf{y}$ is the last visited state at $\Delta_M^\circ$, and the estimate

$$P_N(\mathbf{y}, \mathbf{z}) \geq (1 - \alpha - \theta)^N, \qquad \mathbf{y} \in \mathcal{N}_\theta, \; \mathbf{z} \in \partial_{M,\Gamma}^*,$$

is a direct implication of (7) and the definition of $\mathcal{N}_\theta$.

To complete the proof of Lemma 5.3, we will now decompose an (almost surely) certain event "$\mathbf{X}^{(N)}$ hits the boundary eventually" in a way similar to (58) and compare

two estimates. Let (cf. (37))

$$\partial_{\theta,\eta} := \left\{ x \in \Delta_M^\circ \backslash \mathcal{N}_\theta : \max_{j \notin J^*} x(j) < \eta \right\},$$

so that the interior of the simplex partitions as $\Delta_M^\circ = \mathcal{N}_\theta \bigcup K_{\theta,\eta} \bigcup \partial_{\theta,\eta}$. Similarly to (58), for the probability to eventually hit the boundary starting from a point $\mathbf{x} \in \mathcal{N}_\theta \bigcap \Delta_{M,N}$, we have:

$$
\begin{aligned}
1 \le \ & \sum_{m=0}^{\infty} \sum_{\mathbf{y} \in \mathcal{N}_\theta \bigcap \Delta_{M,N}} \sum_{\mathbf{z} \in \partial_{M,\Gamma}^*} Q_N^m(\mathbf{x},\mathbf{y}) P_N(\mathbf{y},\mathbf{z}) \\
& + \sum_{m=0}^{\infty} \sum_{\mathbf{y} \in \mathcal{N}_\theta \bigcap \Delta_{M,N}} \sum_{\mathbf{z} \in \partial_{\theta,\eta} \backslash \partial_{M,\Gamma}^*} Q_N^m(\mathbf{x},\mathbf{y}) P_N(\mathbf{y},\mathbf{z}) \\
& + \sum_{m=0}^{\infty} \sum_{\mathbf{y} \in \mathcal{N}_\theta \bigcap \Delta_{M,N}} \sum_{z \in \Delta_{M,N} \backslash (\mathcal{N}_\theta \bigcup \partial_{\theta,\eta})} Q_N^m(\mathbf{x},\mathbf{y}) P_N(\mathbf{y},\mathbf{z}) P_{\mathbf{z}}(\overline{G}_N) \\
\le \ & \sum_{m=0}^{\infty} \sum_{\mathbf{y} \in \mathcal{N}_\theta \bigcap \Delta_{M,N}} \sum_{\mathbf{z} \in \partial_{M,\Gamma}^*} Q_N^m(\mathbf{x},\mathbf{y}) P_N(\mathbf{y},\mathbf{z}) \\
& + (1 - \mathbf{e}ta + \theta)^{\eta N} \big(1 + E_{\mathbf{x}}(L_N)\big) + \big(1 + E_{\mathbf{x}}(L_N)\big) \cdot 2MT e^{-\frac{N\varepsilon_\theta^2 c_T^2}{2}}, \quad (59)
\end{aligned}
$$

where we used (55) to estimate $P_{\mathbf{z}}(\overline{G}_N)$.

It follows from (58) and the assumptions of the theorem that

$$
\begin{aligned}
& \frac{(1 - \mathbf{e}ta + \theta)^{\eta N} \big(1 + E_{\mathbf{x}}(L_N)\big) + \big(1 + E_{\mathbf{x}}(L_N)\big) \cdot 2MT e^{-\frac{N\varepsilon_\theta^2 c_T^2}{2}}}{\sum_{m=0}^{\infty} \sum_{\mathbf{y} \in \mathcal{N}_\theta \bigcap \Delta_{M,N}} \sum_{\mathbf{z} \in \partial_{M,\Gamma}^*} Q_N^m(\mathbf{x},\mathbf{y}) P_N(\mathbf{y},\mathbf{z})} \\
& \le \frac{(1 - \mathbf{e}ta + \theta)^{\eta N} \big(1 + E_{\mathbf{x}}(L_N)\big) + \big(1 + E_{\mathbf{x}}(L_N)\big) \cdot 2MT e^{-\frac{N\varepsilon_\theta^2 c_T^2}{2}}}{(1 - \alpha - \theta)^N \big(1 + E_{\mathbf{x}}(L_N)\big)} \xrightarrow{N \to \infty} 0,
\end{aligned}
$$

where (58) is used to estimate the denominator and therefore justify the first inequality. Combining the last result with (58) and (59), we obtain that

$$
\begin{aligned}
1 \le \ & \sum_{m=0}^{\infty} \sum_{\mathbf{y} \in \mathcal{N}_\theta \bigcap \Delta_{M,N}} \sum_{\mathbf{z} \in \partial_{M,\Gamma}^*} Q_N^m(\mathbf{x},\mathbf{y}) P_N(\mathbf{y},\mathbf{z}) \\
& + (1 - \mathbf{e}ta + \theta)^{\eta N} \big(1 + E_{\mathbf{x}}(L_N)\big) + \big(1 + E_{\mathbf{x}}(L_N)\big) \cdot 2MT e^{-\frac{N\varepsilon_\theta^2 c_T^2}{2}} \\
& \le P_{\mathbf{x}}(\mathcal{E}_N) + o\big(P_{\mathbf{x}}(\mathcal{E}_N)\big) \text{ as } N \to \infty.
\end{aligned}
$$

Hence,

$$\lim_{N\to\infty} P_{\mathbf{x}}(\mathcal{E}_N) = \lim_{N\to\infty} \left( \sum_{m=0}^{\infty} \sum_{\mathbf{y}\in\mathcal{N}_\theta \cap \Delta_{M,N}} \sum_{\mathbf{z}\in\partial_{M,\Gamma}^*} Q_N^m(\mathbf{x},\mathbf{y}) P_N(\mathbf{y},\mathbf{z}) \right) = 1.$$

The proof of Lemma 5.3 is complete. $\square$

This completes the proof of part (i) of the theorem. $\square$

**Proof of (ii)** To show that $\lim_{N\to\infty} E(\nu_N) = +\infty$, write, similarly to (56),

$$E(\nu_N) \geq E(\nu_N; G_N) = E_{\varsigma_N}(\nu_N) P(G_N),$$

where $\varsigma_N$ is the distribution of $\mathbf{X}_{\xi_N}^{(N)}$ under the conditional measure $P(\cdot \mid G_N)$, with $\xi_N$ introduced in (54). Using Proposition 4.5 with a compact set $K$ that includes $\mathbf{x}_0$ (the limit of the initial state of $\mathbf{X}^{(N)}$ introduced in the conditions of the theorem) and leveraging the same argument that we employed in order to obtain (55), one can show that $\lim_{N\to\infty} P(G_N) = 1$. Furthermore, the estimate in (55) shows that $L_N$, the number of visits to $\mathcal{N}_\theta$ after time zero is stochastically dominated from below under $P_{\varsigma_N}$ by a geometric variable with probability of success $e^{-\gamma N}$ for any $\gamma > 0$ sufficiently small. Thus,

$$E_{\varsigma_N}(\nu_N) \geq E_{\varsigma_N}(L_N) \geq e^{\gamma N}.$$

Therefore, $\lim_{N\to\infty} E_{\varsigma_N}(\nu_N) = +\infty$, and the proof of the second part of the theorem is complete. $\square$

### 5.6 Proof of Proposition 4.7

Write

$$\Gamma_i(\mathbf{x}) = \frac{1 + Bx(i)}{1 + \mathbf{x}^T \mathbf{B}\mathbf{x}} x(i),$$

where $\mathbf{B} = \frac{\omega}{1-\omega}\mathbf{A}$. By the definition of the interior equilibrium, there is a constant $r > 0$ such that

$$B\chi_{eq}(i) = r \quad \forall i \in S_M. \tag{60}$$

Therefore, for $\mathbf{z} = \chi_{eq} + \mathbf{u} \in \Delta_M^\circ$, as $\mathbf{u}$ approaches zero in $\mathbb{R}^M$, we have:

$$\Gamma_i(\mathbf{z}) = \frac{1 + r + Bu(i)}{1 + r + \mathbf{u}^T \mathbf{B}\chi_{eq} + \chi_{eq}{}^T \mathbf{B}\mathbf{u}} \left( \chi_{eq}(i) + u(i) \right) + o(\|\mathbf{u}\|).$$

Since $\mathbf{B}^T = \mathbf{B}$ and $\sum_{i \in S_M} u(i) = 0$, we have:

$$\chi_{eq}{}^T \mathbf{B} \mathbf{u} = \mathbf{u}^T \mathbf{B} \chi_{eq} = r \sum_{i \in S_M} u(i) = 0.$$

Thus, as $\mathbf{u}$ approaches $\mathbf{0} \in \mathbb{R}^M$,

$$\Gamma_i(\mathbf{z}) = \frac{1 + r + Bu(i)}{1 + r} \big( \chi_{eq}(i) + u(i) \big) + o(\|u\|)$$

$$= \big( \chi_{eq}(i) + u(i) \big) + \frac{\chi_{eq}(i) \cdot Bu(i)}{1 + r} + o(\|u\|).$$

Since $\boldsymbol{\Gamma}$ is a diffeomorphism (Losert and Akin 1983), we can compare this formula with the first-order Taylor expansion of $\boldsymbol{\Gamma}$ around $\chi_{eq}$. It follows that the Jacobian matrix of $\boldsymbol{\Gamma}$ evaluated at $\chi_{eq}$ is given by

$$D_{ij} := \frac{\partial \Gamma_i}{\partial x_j}(\chi_{eq}) = \delta_{ij} + \frac{\chi_{eq}(i) B_{i,j}}{1 + r}, \qquad i.j \in S_M, \tag{61}$$

where $\delta_{ij}$ is the Kronecker symbol.

**Remark 3** The partial derivative computed in (61) is a derivative of $\boldsymbol{\Gamma}$ in the probability simplex $\Delta_M$, which can be thought as an "$\mathbb{R}^M$ with additional constraints." In order to take the constraints into consideration, we vary the variables around an inner point of $\Delta_M$ and use directly the definition of the derivative rather than the standard "algebra of derivatives". In comparison to the Fréchet derivative of $\boldsymbol{\Gamma}$ with the domain in the Banach space $\Delta_M$ calculated in (61), a similar expression for the derivative $\frac{\partial \Gamma_i}{\partial x_j}(\chi_{eq})$ in $\mathbb{R}^M$ contains an extra term $-2r \chi_{eq}(i)/(1 + r)$. To illustrate the point, consider as an example $M = 2$ and $\boldsymbol{\Gamma} = \mathbf{I}$, the identity matrix. Then $\chi(1) = \chi(2) = r = 1/2$ and (61) suggests that $\mathbf{D} = \frac{4}{3}\mathbf{I}$. This can be verified directly: for any $t \in (-1/2, 1/2)$ and $\mathbf{u} = (t, -t)^T$, we have

$$\boldsymbol{\Gamma}(\chi_{eq} + \mathbf{u}) - \chi_{eq} - \mathbf{Du} = \frac{1}{3/2 + 2t^2} \Big( \frac{3}{4} + 2t + t^2, \frac{3}{4} - 2t + t^2 \Big)^T$$

$$- \Big( \frac{1}{2} + \frac{4}{3}t, \frac{1}{2} - \frac{4}{3}t \Big)^T$$

$$= \frac{1}{3/2 + 2t^2} \Big( -\frac{8}{3}t^3, \frac{8}{3}t^3 \Big)^T,$$

and hence, when $\boldsymbol{\Gamma}$ is considered as a function with the domain of definition $\Delta_2$ rather than $\mathbb{R}^2$, $\lim_{\|u\| \to 0} \frac{\|\boldsymbol{\Gamma}(\chi_{eq} + u) - \chi_{eq} - \mathbf{Du}\|}{\|u\|} = 0$ as desired.

It is easy to verify that

$$\mathbf{H} := \frac{1 + r}{1 + 2r} \mathbf{D}$$

is a transpose of a strictly positive stochastic matrix, namely $H_{ij} > 0$ for all $i$, $j \in S_M$ and $\sum_{i \in S_M} H_{ij} = 1$ for all $j \in S_M$.

Note that $\chi_{eq}$ is a left eigenvector of $\mathbf{H}^T$ corresponding to the principle eigenvalue 1. Indeed, assuming $i \neq j$, we deduce from (61) and the fact that $\mathbf{B}^T = \mathbf{B}$ :

$$\chi_{eq}(i) H_{ij}^T = \frac{1+r}{1+2r} \frac{\chi_{eq}(i) B_{i,j} \chi_{eq}(j)}{1+r} = \chi_{eq}(j) H_{ji}^T.$$

Thus, stochastic matrix $\mathbf{H}^T$ is transition kernel of a reversible Markov chain with stationary distribution $\chi_{eq}$. From the reversibility of the Markov transition matrix $\mathbf{H}^T$, it follows that all the eigenvalues of $\mathbf{H}$, and hence also of $\mathbf{D}$, are real numbers [see, for instance, Brémaud (1998)].

Recall $W_M$ from (20). The asymptotic stability of the equilibrium (see the paragraph above (52)) implies that the largest absolute value of an eigenvalue of the linear operator $\mathbf{D}$ restricted to the linear space $W_M$ is at most one (Perko 1991). It follows from (61) that if $\lambda$ is this eigenvalue and $|\lambda| = 1$, then there is an eigenvalue $\eta$ of the matrix with entries $\frac{\chi_{eq}(i) B_{i,j}}{1+r}$ such that $\eta \in \{-2, 0\}$. To complete the proof of the proposition it remains to show that this is indeed impossible. Toward this end, denote by $\mathbf{H}_\chi$ the matrix with entries $\frac{\chi_{eq}(i) B_{i,j}}{1+r}$ and observe that

$$\det(\mathbf{H}_\chi) = (1+r)^{-M} \prod_{i=1}^{M} \chi_{eq}(i) \cdot \det \mathbf{B}.$$

Since $\mathbf{B}$ is invertible by our assumptions, this shows that 0 cannot be an eigenvalue of $\mathbf{H}_\chi$. Furthermore, if $\eta = -2$ is an eigenvalue of $\mathbf{H}_\chi$, then there exists $v \in W_M$ such that

$$\chi_{eq} \circ \mathbf{B} v = -2(1+r)v,$$

where $\circ$ denotes the element-wise product introduced in (6). Recall now that $\mathbf{B} = h\mathbf{A}$, where $h = \frac{\omega}{1-\omega}$. Plugging this expression into the previous formula, we obtain

$$\chi_{eq} \circ \mathbf{A} v = -2(1+r)h^{-1}v \tag{62}$$

Now, choose a constant $\eta > 1$ so close to one that $\widetilde{\mathbf{A}} = \eta \mathbf{A}$ satisfies Assumption 4.4, namely $\widetilde{\mathbf{A}}$ is invertible and has exactly one positive eigenvalue.

Denote $\widetilde{\mathbf{B}} := h\widetilde{\mathbf{A}} = \eta \mathbf{B}$. The analogue of (60) for $\widetilde{\mathbf{B}}$ reads

$$\widetilde{\mathbf{B}} \chi_{eq}(i) = \eta \mathbf{B} \chi_{eq}(i) = r\eta \quad \forall i \in S_M.$$

In this sense, $r\eta$ plays the same role for the pair $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}$ as $r$ does for $\mathbf{A}$ and $\mathbf{B}$. Moreover, (62) implies that

$$\frac{1}{1+r\eta} \chi_{eq} \circ \widetilde{\mathbf{B}} v = -2 \frac{\eta(1+r)}{1+r\eta} v.$$

Since $\left|\frac{\eta(1+r)}{1+r\eta}\right| > 1$, it follows that the spectral radius of the Jacobian of the vector field $\widetilde{\boldsymbol{\Gamma}}(\mathbf{x}) := \mathbf{x} \circ \frac{1+h\widetilde{\mathbf{A}}\mathbf{x}}{1+h\mathbf{x}\widetilde{\mathbf{A}}\mathbf{x}}$ evaluated at $\boldsymbol{\chi_{eq}}$ is greater than one. This is a clear contradiction, since matrix $\widetilde{\mathbf{A}}$ is assumed to satisfy the conditions of Theorem 4, and hence the spectral radius must be at most one. □

### 5.7 Proof of Corollary 4.8

For $\varepsilon > 0$, denote $\mathcal{U}_\varepsilon^{(N)} := \left\{\mathbf{x} \in \Delta_{M,N} : \|\mathbf{x} - \boldsymbol{\chi_{eq}}\| < \varepsilon\right\}$. Since $\boldsymbol{\Gamma}$ is a diffeomorphism under the conditions stated in Assumption 4.4 (Losert and Akin 1983), Proposition 4.7 implies that there exist $\varepsilon > 0$ and $\rho \in (0, 1)$ such that the spectral radius of the Jacobian is strictly less than $\rho$ within $\mathcal{U}_\varepsilon^{(N)}$. Thus, $\boldsymbol{\Gamma}(\mathcal{U}_\varepsilon^{(N)}) \subset \mathcal{U}_{\rho\varepsilon}^{(N)}$. Without loss of generality we may assume that $N$ is so large that both $\mathcal{U}_{\rho\varepsilon}^{(N)}$ and $\mathcal{U}_\varepsilon^{(N)} \backslash \mathcal{U}_{\rho\varepsilon}^{(N)}$ are nonempty. Indeed, for any $N_0 \in \mathbb{N}$ and $b > 0$ the claim is trivially true for all $N \leq N_0$ if we choose $C > 0$ in such a way that $1 - Ce^{-bN_0} < 0$.

Since $\boldsymbol{\Gamma}(\mathcal{U}_\varepsilon^{(N)}) \subset \mathcal{U}_{\rho\varepsilon}^{(N)}$ and $P_N(\mathbf{x}, \mathbf{y}) > 0$ for all $\mathbf{x}, \mathbf{y} \in \Delta_{M,N}^\circ$ (and hence, by virtue of (41), the quasi-stationary distribution $\mu_N$ puts strictly positive weights on all $\mathbf{x} \in \Delta_{M,N}^\circ$), for suitable $C > 0$ and $b > 0$ we have:

$$
\begin{aligned}
\lambda_N \mu_N(\mathcal{U}_{\rho\varepsilon}^{(N)}) = \mu_N P_{\circ,N}(\mathcal{U}_{\rho\varepsilon}^{(N)}) &= \sum_{\mathbf{x} \in \Delta_{M,N}^\circ} \sum_{\mathbf{y} \in \mathcal{U}_{\rho\varepsilon}^{(N)}} \mu_N(x) P_N(\mathbf{x}, \mathbf{y}) \\
&\geq \sum_{\mathbf{x} \in \mathcal{U}_\varepsilon^{(N)}} \sum_{\mathbf{y} \in \mathcal{U}_{\rho\varepsilon}^{(N)}} \mu_N(\mathbf{x}) P_N(\mathbf{x}, \mathbf{y}) \geq \mu_N(\mathcal{U}_\varepsilon^{(N)})(1 - Ce^{-bN}) \\
&> \mu_N(\mathcal{U}_{\rho\varepsilon}^{(N)})(1 - Ce^{-bN}),
\end{aligned}
\tag{63}
$$

which implies the claim. To obtain the crucial second inequality in (63)

$$
\sum_{\mathbf{y} \in \mathcal{U}_{\rho\varepsilon}^{(N)}} P_N(\mathbf{x}, \mathbf{y}) = P_N(\mathbf{x}, \mathcal{U}_{\rho\varepsilon}^{(N)}) \geq 1 - Ce^{-bN}
$$

one can rely on the result of Proposition 4.5 and use an argument similar to the one which led us to (55). □

## 6 Conclusion

In this paper we examined the long-term behavior of a general Wright–Fisher model. Our results hold for a broad class of Wright–Fisher models, including several instances that have been employed in experimental biology applications. Our primary motivation came from research conducted on gut gene-microbial ecological networks (Chapkin et al.).

The main results of the paper are:

– A maximization principle for deterministic replicator dynamics stated in Theorem 3.2. In this theorem, we show that the average of a suitably chosen reproductive fitness of the mean-field model is non-decreasing with time, and is strictly increasing outside of a certain chain-recurrent set. The proof relies on a "fundamental theorem of dynamical systems" of Conley (1978).
– Gaussian approximation scheme constructed for the Wright–Fisher model in Theorem 4.1.
– Theorem 4.2, which gives an exponential lower bound on the first time that the error of mean-field approximation exceeds a given threshold.
– Theorem 4.6, which specifies the route to extinction in the case that the following three conditions are satisfied: (1) the mean-field dynamical system has a unique interior equilibrium; (2) boundary of the probability simplex is repelling for the mean-field map; and (3) the stochastic mixing effect (in certain rigorous sense) dominates extinction forces.
– Proposition 4.7 that elaborates on the mechanism of convergence to the equilibrium in the mean-field model of Theorem 4.6 by establishing a stringent contraction near the equilibrium property for the systems considered in the theorem.

We next outline several directions for extensions and future research.

1. The competitive exclusion principle predicts that only the fittest species will survive in a competitive environment. In Hutchinson (1961), Hutchinson addressed the *paradox of the plankton* which is a discrepancy between this principle and the observed diversity of the plankton. One of the suggested resolutions of this paradox is that the environmental variability might promote the diversity. Several mathematical models of species persistence in time-homogeneous environments have been proposed (Benaïm and Schreiber 2019; Carja et al. 2014; Dean and Shnerb 2019; Li and Cheng 2019; Li and Chesson 2016; Mahdipour-Shirayeh et al. 2017). In order to maintain the diversity, environmental factors must fluctuate within a broad range supporting a variety of contradicting trends and thus promoting the survival of various species. The theory of permanence and its stochastic counterpart, the theory of stochastic persistence (Benaïm and Schreiber 2019), is an alternative (but not necessarily mutually excluding) mathematical framework able to explain the coexistence of interacting species in terms of topological properties of the underlying dynamical system. Extension of our results to systems in fluctuating environments and fluctuating size populations (e.g., Park and Traulsen 2017) seems to be a natural direction for future research with important implications regarding the dynamical interactions between microbial communities, host gene expression, and nutrients in the human gut.
2. We believe that Theorem 4.6 can be extended to a broad class of stochastic population models with mean-field vector-field $\mathbf{\Gamma}$ promoting cooperation between the particles (cells/microorganisms in applications). Among specific biological mechanisms promoting forms of cooperation are: public goods (Archetti and Scheuring 2012; Chuang et al. 2009; Constable et al. 2016; Kurokawa and Ihara 2009; McAvoy et al. 2018; Zhang et al. 2013), relative nonlinearity of competition (Arnoldt et al. 2012; Chesson 2018; Harmand et al. 2019), population-level randomness (Constable et al. 2016; Coron et al. 2019; Kroumi and Lessard 2015), and diversification

of cells into distinct ecological niches (Good et al. 2018; Rohlfs and Hoffmeist 2003). Public goods models are of a special interest for us because of a possibility of an application to the study of a gene-microbial-metabolite ecological network in the human gut (Chapkin et al.). A public good is defined in Archetti and Scheuring (2012) as any benefit that is "simultaneously non-excludable (nobody can be excluded from its consumption) and non-rivalrous (use of the benefit by one particle does not diminish its availability to another particle)". For instance, an alarm call against a predator is a pure public good while natural resources (e. g. the air) and enzyme production in bacteria can be approximated to public goods (Archetti and Scheuring 2012). Public goods games in microbial communities have been considered by several authors, see Archetti and Scheuring (2012), Cavaliere et al. (2017), Chuang et al. (2009), Zhang et al. (2013) for a review. An important feature of the public goods game is the intrinsic non-linearity of the fitness (Archetti and Scheuring 2012; Zhang et al. 2013).

3. It seems plausible that Theorem 4.6 can be extended to a multitype Moran model. The mean-field dynamics as well the quasi-stationary distributions for such models are relatively well studied.

4. In contrast to non-rivalrous interactions captured by public goods games, competitive interaction between cells can lead to negative frequency dependence, see Broekman et al. (2019), Broom and Rychtář (2013) and references therein. A classical example of negative frequency dependence is animals foraging on several food patches (Broom and Rychtář 2013). We believe it is of interest to explore the path to extinction of this class of models in the spirit of analysis undertaken in Sect. 4.3.

5. An interesting instance of the Wright–Fisher model with an replicator update rule was introduced in Kroumi and Lessard (2015). This is a mathematical model for cooperation in a phenotypic space, and is endowed with a rich additional structure representing phenotypic traits of the cells in addition to the usual division into two main categories/types, namely cooperators and defectors. In the exact model considered in Kroumi and Lessard (2015), the phenotypic space is modelled as a Euclidean space $\mathbb{Z}^d$. However, a variation with finite population (for instance, restricting the phenotypic space to a cube in $\mathbb{Z}^d$) can be studied in a similar manner. Extending the result of Theorem 4.6 to a variation of this Wright–Fisher model is challenging, but could potentially contribute to a better understanding of mechanisms forming extinction trajectory in stochastic population models.

6. In Sect. 4.3 we investigated the phenomenon of extinction of a long-lived closed stochastic population. The manner in which the population is driven to extinction depends on whether the extinction states are repellers or attractors of the mean-field dynamics (cf. Assaf and Meerson 2010; Assaf and Mobilia 2011; Park and Traulsen 2017). In the latter case the extinction time is short because with a high probability the stochastic trajectory follows the mean-field path closely for a long enough time to get to a close proximity (in a suitable phase space) of the extinction set, and then it becomes fixed there via small noise fluctuations. In contrast, in the former scenario, attractors of the deterministic dynamics act as traps or pseudo-equilibria for the stochastic model and thus create metastable stochastic states. While Theorem 4.6 is concerned with a model of this kind, it seems plausible that the result can be

extended to a situation where all attractors of the mean-field model belong to the boundary of the simplex $\Delta_M$.

7. The proof of Theorem 4.1 relies on a representation of $\mathbf{X}_k^{(N)}$ as partial sums of i. i. d. random indicators (see equations (44) and (45)), and can be extended beyond the i. i. d. setting. In particular, it can be generalized to cover the Cannings exchangeable model and models in fluctuating environments. We believe that with a proper adjustment, the result also holds for a variation of the model with $N$ depending on time (Gompert 2016; Grosjean and Huillet 2017; Iizuka 2010) (this is, for instance, well known for branching processes with immigration). This theorem can also be extended to the *paintbox models* (Boenkost et al. 2019) and to a class of Cannings models in phenotypic spaces with weak and fast decaying dependency between phenotypes (Kroumi and Lessard 2015).

8. Finally, proving Conjecture 4.9 would shed further light on the relations among metastability, quasi-stationary distributions, and the extinction trajectory for the underlying Wright–Fisher model.

# 7 Appendices

## Appendix A. Absorption states and stochastic equilibria

In this appendix we explore the connection between basic geometric properties of the vector field $\boldsymbol{\Gamma}$ and the structure of closed communication classes of the Markov chain $\mathbf{X}^{(N)}$. In other words, we study the configuration of absorbing states and identify possible supports of stationary distributions according to the properties of $\boldsymbol{\Gamma}$. We remark that the results of this section are included for the sake of completeness and are not used anywhere else in the remainder of the paper. We have made the section self-contained and included the necessary background on the Markov chain theory.

The main result of the appendix is stated in Theorem 7.3 where a geometric characterization of recurrent classes is described. The general structure of recurrent classes of $\mathbf{X}^{(N)}$ in this theorem holds universally for the Wright–Fisher model (7) with any update function $\boldsymbol{\Gamma}$. Thus, in Appendix A we deviate from assumption (10) that we have used throughout the paper and to which we will return in the subsequent Appendix B, and adopt instead the following general one:

**Assumption 7.1** $\mathbf{X}^{(N)}$ is a Markov chain on $\Delta_M$ with transition kernel defined by (7) and an arbitrary update rule $\boldsymbol{\Gamma} : \Delta_M \to \Delta_M$.

Recall (5) and (13). The discussion in this section is largely based on the fact that

$$P_N(\mathbf{x}, \mathbf{y}) > 0 \iff \mathcal{C}(\mathbf{y}) \subset \mathcal{C}\big(\boldsymbol{\Gamma}(\mathbf{x})\big) \iff P_N(\mathbf{x}, \mathbf{z}) > 0 \text{ for all } \mathbf{z} \in \Delta_{[\mathcal{C}(\mathbf{y})], M}.$$
(64)

This simple observation allows one to relate the geometry of zero patterns in the vector field to communication properties, and hence asymptotic behavior, of $\mathbf{X}^{(N)}$.

From the point of view of the model's adequacy in potential applications, (64) is a direct consequence of the fundamental assumption of the Wright–Fisher model that particles in the population update their phenotypes independently of each other and follow the same stochastic protocol. Mathematically, this assumption is expressed in identity (45), whose right-hand side is a sum of independent and identically distributed random indicators. To appreciate this feature of the Wright–Fisher model, we remark that while (64) holds true for an arbitrary model (7), that feature does not hold true, in general, for two most natural generalization of the Wright–Fisher model, namely the Cannings exchangeable model (Ewens 2012) and the pure-drift Generalized Wright–Fisher process of Der et al. (2011).

Recall that two states $\mathbf{x}, \mathbf{y} \in \Delta_{M,N}$ of the finite-state Markov chain $\mathbf{X}^{(N)}$ are said to communicate if there exist $i, j \in \mathbb{N}$ such that $P_N^i(\mathbf{x}, \mathbf{y}) > 0$ and $P_N^j(\mathbf{y}, \mathbf{x}) > 0$. If $\mathbf{x}$ and $\mathbf{y}$ communicate, we write $\mathbf{x} \leftrightarrow \mathbf{y}$. The binary relation $\leftrightarrow$ partitions the state space $\Delta_{M,N}$ into a finite number of disjoint equivalence classes, namely $\mathbf{x}, \mathbf{y}$ belong to the same class if and only if $\mathbf{x} \leftrightarrow \mathbf{y}$. A state $\mathbf{x} \in \Delta_{M,N}$ is called absorbing if $P_N(\mathbf{x}, \mathbf{x}) = 1$, recurrent if $\mathbf{X}^{(N)}$ starting at $\mathbf{x}$ returns to $\mathbf{x}$ infinitely often with probability one, and transient if, with probability one, $\mathbf{X}^{(N)}$ starting at $\mathbf{x}$ revisits $\mathbf{x}$ only a finite number of times. It turns out that each communication class consists of either (1) exactly one absorbing state; or (2) recurrent non-absorbing states only; or (3) transient states only. No other behavior is possible due to zero–one laws enforced by the Markov property (Lawler 2006). A communication class is called closed or recurrent if it belongs to the first or second category, and transient if it belongs to the third one. The chain is called irreducible if there is only one (hence, recurrent) communication class, and aperiodic if $P_N^j(\mathbf{x}, \mathbf{y}) > 0$ for some $j \in \mathbb{N}$ and all $\mathbf{x}, \mathbf{y} \in \Delta_{M,N}$. If $C_1, \ldots, C_r$ are disjoint recurrent classes of $\mathbf{X}^{(N)}$, then the general form of its stationary distribution is $\pi(\mathbf{x}) = \sum_{i=1}^r \alpha_i \pi_i(\mathbf{x})$ where $\pi_i$ is the unique stationary distribution supported on $C_i$ (that is $\pi_i(\mathbf{x}) > 0$ if and only if $\mathbf{x} \in C_i$) and $\alpha_i$ are arbitrary non-negative numbers adding up to one. If the chain is irreducible, then the stationary distribution is unique and is strictly positive on the whole state space $\Delta_{M,N}$. An irreducible finite-state Markov chain is aperiodic if and only if $\pi(\mathbf{x}) = \lim_{k \to \infty} P_N^k(\mathbf{y}, \mathbf{x})$ for any $\mathbf{x}, \mathbf{y} \in \Delta_{M,N}$, where $\pi$ is the unique stationary distribution of the chain (Lawler 2006).

Let $\mathcal{A}_N \subset \Delta_M$ denote the set of absorbing states of the Markov chain $\mathbf{X}^{(N)}$. It readily follows from (7) that $\mathcal{A}_N \subset V_M$, and a vertex $\mathbf{e}_j \in V_M$ is an absorbing (fixation) state if and only if $\mathbf{\Gamma}(\mathbf{e}_j) = \mathbf{e}_j$. In particular, $\mathcal{A}_N$ depends only on the update rule $\mathbf{\Gamma}$ and is independent of the population size $N$. We summarize this observation as follows:

$$\mathcal{A}_N = \left\{ \mathbf{e}_j \in V_M : \mathbf{\Gamma}(\mathbf{e}_j) = \mathbf{e}_j \right\} \qquad \forall N \in \mathbb{N}. \tag{65}$$

The following lemma asserts that if $\mathbf{e}_j$ is an absorbing state of $\mathbf{X}^{(N)}$, then the type $j$ cannot be mixed into any stochastic equilibrium of the model, namely any other than $\mathbf{e}_j$ state $\mathbf{x} \in \Delta_{M,N}$ with $x(j) > 0$ is transient.

**Lemma 7.2** *Let Assumption* 7.1 *hold. If* $\mathbf{x} \in \Delta_{M,N}$ *is a recurrent state of the Markov chain* $\mathbf{X}^{(N)}$ *and* $x(j) > 0$, *then* $\mathbf{e}_j$ *is a recurrent state belonging to the same closed communication class as* $\mathbf{x}$.

**Proof** By (65), a state $\mathbf{x} \in \Delta_{M,N}$ cannot be absorbing if $\mathbf{x} \neq \mathbf{e}_j$ and $x(j) > 0$. If $\mathbf{x}$ is recurrent and non-absorbing, then there is a state $\mathbf{y} \in \Delta_{M,N}$ that belongs to the same closed communication class as $\mathbf{x}$ and such that $P_N(\mathbf{y}, \mathbf{x}) > 0$. In view of (7) this implies that $\Gamma_j(\mathbf{y}) > 0$ and hence $P_N(\mathbf{y}, \mathbf{e}_j) > 0$. Since the state $\mathbf{y}$ is recurrent, $\mathbf{e}_j$ must belong to the same communication class as $\mathbf{y}$ and $\mathbf{x}$, and in particular is recurrent. The proof is complete.                                                                                   $\square$

Recall (5). The same argument as in the above lemma shows that if an interior point $\mathbf{x}$ of a simplex $\Delta_{[J],N}$ is recurrent for $\mathbf{X}^{(N)}$, then the whole simplex $\Delta_{[J],N}$ belongs to the (recurrent) closed communication class of $\mathbf{x}$. One can rephrase this observation as follows:

**Theorem 7.3** *Let Assumption* 7.1 *hold. Then, any recurrent class of the Markov chain* $\mathbf{X}^{(N)}$ *has the form of a simplicial complex* $\bigcup_\ell \Delta_{[J_\ell],N}$, *where* $J_\ell$ *are (possibly overlapping) subsets (possibly singletons) of* $S_M$.

With applications in mind, we collect some straightforward implications of this proposition in the next corollary.

**Corollary 7.4** *Let Assumption* 7.1 *hold. Then:*

(i) *If there is a stationary distribution* $\pi$ *of* $\mathbf{X}^{(N)}$ *and* $J \subset S_M$ *such that* $\pi(\mathbf{x}) > 0$ *for some interior point* $\mathbf{x} \in \Delta_{[J],N}^\circ$, *then* $\pi(\mathbf{y}) > 0$ *for all* $\mathbf{y} \in \Delta_{[J],N}$. *Furthermore, the following holds in this case unless* $\Gamma(\mathbf{y}) \cdot \mathbf{y} = 0$ *(that is,* $\mathbf{y}$ *and* $\Gamma(\mathbf{y})$ *are orthogonal) for all* $\mathbf{y} \in \Delta_{[J],N}$ :

   (a) *Let* $C$ *be the (closed) communication class to which the above* $\mathbf{x}$ *(and hence the entire* $\Delta_{[J],N}$) *belongs. Then the Markov chain* $\mathbf{X}^{(N)}$ *restricted to* $C$ *is aperiodic.*

   (b) *Let* $T$ *be the first hitting time of* $C$, *namely* $T = \inf\{k \in \mathbb{Z}_+ : \mathbf{X}_k^{(N)} \in C\}$. *As usual, we use here the convention that* $\inf \emptyset = +\infty$. *Then*

   $$\pi(\mathbf{y}) = \lim_{k \to \infty} P(\mathbf{X}_k^{(N)} = \mathbf{y} \mid \mathbf{X}_0^{(N)} = \mathbf{z}, T < \infty)$$

   *for all* $\mathbf{y} \in C$ *and* $\mathbf{z} \in \Delta_{M,N}$. *In particular,* $\pi(\mathbf{y}) = \lim_{k \to \infty} P_N^k(\mathbf{z}, \mathbf{y})$ *whenever* $\mathbf{y}, \mathbf{z} \in C$.

(ii) *A stationary distribution* $\pi$ *of* $\mathbf{X}^{(N)}$ *such that* $\pi(\mathbf{x}) > 0$ *for some interior point* $\mathbf{x} \in \Delta_{M,N}^\circ$ *exists if and only if* $\mathbf{X}^{(N)}$ *is irreducible and aperiodic, in which case the stationary distribution is unique and* $\pi(\mathbf{y}) = \lim_{k \to \infty} P_N^k(\mathbf{x}, \mathbf{y}) > 0$ *for any* $\mathbf{x}, \mathbf{y} \in \Delta_{M,N}$.

(iii) *Let* $B = \left\{ \exists \, k \in \mathbb{N} : \mathbf{X}_k^{(N)} \in \partial\left(\Delta_{M,N}\right) \text{ for all } m \geq k \right\}$ *be the event that the trajectory of* $\mathbf{X}^{(N)}$ *reach eventually the boundary of the simplex and stays there forever. Then either the Markov chain* $\mathbf{X}^{(N)}$ *is irreducible and aperiodic or* $P(B) = 1$ *for any state* $\mathbf{X}_0^{(N)}$.

(iv) *If* $\Gamma(\mathbf{e}_j) = \mathbf{e}_j$ *for some* $j \in S_M$, *then* $\lim_{k \to \infty} X_k^{(N)}(j) \in \{0, 1\}$, *a. s., for any initial distribution of* $\mathbf{X}^{(N)}$.

We will only prove the claim in (a) of part (i), since the rest of the corollary is an immediate result of a direct combination of Theorem 7.3 and general properties of finite-state Markov chains reviewed in the beginning of the section.

***Proof of Corollary 7.4-(i)-a*** Suppose that $\Gamma(\mathbf{y}) \cdot \mathbf{y} > 0$ for some $\mathbf{y} \in \Delta_{[J],N}$. Let $\mathbf{z}$ be the projection of $\mathbf{y}$ into $\Delta_{[I],M}$, where $I := \{j \in J : y(j)\Gamma_j(\mathbf{y}) > 0\}$. That is,

$$z(i) = \begin{cases} y(i) \text{ if } y(i)\Gamma_i(\mathbf{y}) > 0, \\ 0 \quad \text{otherwise.} \end{cases}$$

Then $\mathbf{z} \in \Delta_{[I],M}^{\circ}$ and $P_N(\mathbf{z}, \mathbf{z}) > 0$. Since $\mathbf{z} \in \Delta_{[J],N}$, by Theorem 7.3 $\mathbf{z}$ belongs to the same closed communication class as $\mathbf{x}$. The asserted aperiodicity of $\mathbf{X}^{(N)}$ restricted to $C$ is a direct consequence of the existence of the 1-step communication loop at $\mathbf{z}$ (Lawler 2006). □

***Example 6*** (Harper and Fryer 2016) If $\Gamma_j(\mathbf{x}) = x(j)^\alpha f(\mathbf{x})$ for some $j \in S_M$, function $f : \Delta_M \to \mathbb{R}_+$ such that $f(\mathbf{e}_j) = 1$, and a constant $\alpha > 0$, then $\Gamma(\mathbf{e}_j) = \mathbf{e}_j$. By part (iv) of Corollary 7.4, this fact alone suffices to conclude that $\lim_{k \to \infty} X_k^{(N)}(j) \in \{0, 1\}$, a. s., for this specific type $j$.

A simple sufficient condition for the Markov chain $\mathbf{X}^{(N)}$ to be irreducible and aperiodic is that $\Gamma(\Delta_M) \subset \Delta_M^{\circ}$, in which case $P_N(\mathbf{x}, \mathbf{y}) > 0$ for any $\mathbf{x}, \mathbf{y} \in \Delta_{M,N}$. If $\Gamma(\Delta_{M,N}^{\circ}) \subset \Delta_{M,N}^{\circ}$ but $\Gamma(\mathbf{e}_j) = \mathbf{e}_j$ for all $j \in S_M$, the model can be viewed as a multivariate analog of what is termed in Chalub and Souza (2017) as a Kimura class of Markov chains. The latter describes evolution dynamics of a population that initially consists of two phenotypes, each one eventually either vanishes or takes over the entire population.

## Appendix B. Permanence of the mean-field dynamics

In this appendix, we focus on Wright–Fisher models with a permanent mean-field dynamics $\Gamma$. The permanence means that the boundary of the state space $\partial(\Delta_M)$ is a repeller for the mean-field dynamics, implying that all types under the mean-field dynamics will ultimately survive (Garay and Hofbauer 2003; Hofbauer and Sigmund 1998; Hutson and Schmitt 1992; Kang and Chesson 2010). This condition is particularly relevant to the stochastic Wright–Fisher model because, as stated in Appendix A, all absorbing states (possibly an empty set) of the Markov chain $\mathbf{X}^{(N)}$ lie within the boundary set $\partial(\Delta_M)$. Of course, in the absence of mutations, any vertex of $\Delta_M$ is an absorbing state for the Markov chain $\mathbf{X}^{(N)}$, and the latter is going to eventually become fixed on one of the absorbing states with probability one (see Appendix A for formal details). However, we showed in Sect. 4 that a permanent mean-field dynamics induces metastability and a prolonged coexistence of types.

**Definition 4** (Hofbauer and Sigmund 1998) $\Gamma$ is permanent if there exists $\delta > 0$ such that

$$\liminf_{k \to \infty} \psi_k(i) > \delta \qquad \forall\, i \in S_M, \boldsymbol{\psi}_0 \in \Delta_M^\circ.$$

See Garay and Hofbauer (2003), Schreiber et al. (2018) and references therein for sufficient conditions for the permanence. It turns out that under reasonable additional assumptions, $\Gamma$ is permanent for the fitness defined in either (11) or (12). More precisely, we have (for a more general version of the theorem, see Garay and Hofbauer (2003, Theorem 11.4) and Hofbauer and Sigmund (2003, Theorem 5):

**Theorem 3** (Garay and Hofbauer 2003) *Let* $\mathbf{A}$ *be an* $M \times M$ *matrix which satisfies the following condition: There exists* $\mathbf{y} \in \Delta_M^\circ$ *such that*

$$\mathbf{y}^T \mathbf{A} \mathbf{z} > \mathbf{z}^T \mathbf{A} \mathbf{z}$$

*for all* $\mathbf{z} \in \partial(\Delta_M)$ *solving the fixed-point equation* $\mathbf{z} = \frac{1}{\mathbf{z}^T \mathbf{A} \mathbf{z}}(\mathbf{z} \circ \mathbf{A} \mathbf{z})$.
*Then:*

(i) *There exists* $\omega_0 > 0$ *such that* $\Gamma$ *given by* (10) *and* (11) *is permanent for all* $\omega \in (0, \omega_0)$.
(ii) $\Gamma$ *given by* (10) *and* (12) *is permanent for all* $\eta > 0$.

Recall Definitions 1 and 2. The setting of the theorem provides a natural within our context example of a non-constant complete Lyapunov function.

**Example 7** Suppose that $\Gamma$ is a replicator function defined in (10). If $\mathbf{A}$ is an $M \times M$ matrix that satisfies the conditions of Theorem 3 and $\boldsymbol{\varphi}$ in (10) is defined by either (11) with small enough $\omega > 0$ or by (12), then $\mathcal{R}(\Gamma) \subset K \bigcup \partial(\Delta_M)$, where $K$ is a compact subset of $\Delta_M^\circ$. In particular, the complete Lyapunov function $h$ constructed in Theorem 2 (p. 14) is strictly increasing on the non-empty open set $\Delta_M^\circ \backslash K$. Notice that while the average $\mathbf{x} \cdot \boldsymbol{\varphi}(\mathbf{x})$ may not be a complete Lyapunov function for $\Gamma$, the average $h(\mathbf{x})$ of the reproductive fitness $\widetilde{\boldsymbol{\varphi}}(\mathbf{x}) := h(\mathbf{x})\boldsymbol{\varphi}(\mathbf{x})$ is.

In Sect. 4.3 we considered a stochastic Wright–Fisher model with a permanent mean-field map $\Gamma$. This model has a global attractor, consisting of a unique equilibrium point. A sufficient condition for such a dynamics is given by the following theorem (Losert and Akin 1983). Recall (20).

**Theorem 4** (Losert and Akin 1983) *Let Assumption 4.4 hold. Then the following holds true:*

(a) *Recall* $\boldsymbol{\psi}_k$ *from* (3). *If* $\boldsymbol{\psi}_0 \in \Delta_M^\circ$, *then* $\lim_{k \to \infty} \boldsymbol{\psi}_k = \boldsymbol{\chi}_{eq}$.
(b) $\Gamma$ *is a diffeomorphism.*

Note that the weak selection condition (cf. Antal et al. 2009; Boenkost et al. 2019; Imhof and Nowak 2006) imposed in part (i) of Theorem 3 is not required in the conditions of Theorem 4.

The evolutionary games with $\mathbf{A} = \mathbf{A}^T$ are called *partnership games* (Hofbauer 2011). We refer to Hummert et al. (2018), Huttegger et al. (2014), Traulsen and Reed

(2012) for applications of partnership games in evolutionary biology. Condition (32) implies the existence of a unique evolutionary stable equilibrium for the evolutionary game defined by the payoff matrix $\mathbf{A}$. For a symmetric reversible matrix $\mathbf{A}$ this condition holds if and only if $\mathbf{A}$ has exactly one positive eigenvalue (Kingman 1961; Mandel 1959). Condition (33) ensures that the equilibrium is an interior point. Note that since the equilibrium is unique, the conditions of Theorem 3 are automatically satisfied, that is $\boldsymbol{\psi}_k$ is permanent. Finally, under the conditions of Theorem 4, the average payoff $\mathbf{x}^T \mathbf{A} \mathbf{x}$ is a Lyapunov function (Kingman 1961; Mandel 1959):

$$\boldsymbol{\psi}_k^T \mathbf{A} \boldsymbol{\psi}_k < \boldsymbol{\psi}_{k+1}^T \mathbf{A} \boldsymbol{\psi}_{k+1} < {\boldsymbol{\chi}_{eq}}^T \mathbf{A} \boldsymbol{\chi}_{eq} \qquad \forall\, k \in \mathbb{Z}_+,$$

as long as $\boldsymbol{\psi}_0 \in \Delta_M^\circ$ and $\boldsymbol{\psi}_0 \neq \boldsymbol{\chi}_{eq}$.

# References

Akhmet M, Fen MO (2016) Input-output mechanism of the discrete chaos extension. In: Afraimovich V, Machado JAT, Zhang J (eds) Complex motions and chaos in nonlinear systems, Nonlinear Syst. Complex., vol 15. Springer, Berlin, pp 203–233

Antal T, Traulsen A, Ohtsuki H, Tarnita CE, Nowak MA (2009) Mutation-selection equilibrium in games with multiple strategies. J Theor Biol 258:614–622

Ao P (2005) Laws in Darwinian evolutionary theory. Phys Life Rev 2:117–156

Archetti M, Scheuring I (2012) Review: game theory of public goods in one-shot social dilemmas without assortment. J Theor Biol 299:9–20

Arnoldt H, Timme M, Grosskinsky S (2012) Frequency-dependent fitness induces multistability in coevolutionary dynamics. J R Soc Interface 9:3387–3396

Ashcroft P, Michor F, Galla T (2015) Stochastic tunneling and metastable states during the somatic evolution of cancer. Genetics 199:1213–1228

Assaf M, Meerson B (2008) Noise enhanced persistence in a biochemical regulatory network with feedback control. Phys Rev Lett 100:058105

Assaf M, Meerson B (2010) Extinction of metastable stochastic populations. Phys Rev E 81:021116

Assaf M, Mobilia M (2011) Fixation of a deleterious allele under mutation pressure and finite selection intensity. Phys Rev E 275:93–103

Auletta V, Ferraioli D, Pasquale F, Persiano G (2018) Metastability of logit dynamics for coordination games. Algorithmica 80:3078–3131

Aydogmus O (2016) On extinction time of a generalized endemic chain-binomial model. Math Biosci 279:38–42

Banks J, Brooks J, Cairns G, Davis G, Stacey P (1992) On Devaney's definition of chaos. Am Math Mon 99:332–334

Barton NH, Coe JB (2009) On the application of statistical physics to evolutionary biology. J Theor Biol 259:317–324

Barton JP, Goonetilleke N, Butler TC, Walker BD, McMichael AJ, Chakraborty AK (2016) Relative rate and location of intra-host HIV evolution to evade cellular immunity are predictable. Nat Commun 7:11660

Basener WF (2013) Limits of chaos and progress in evolutionary dynamics, In: Marks II RJ, Behe MJ, Dembski WA, Gordon BL, Sanford JC (eds) Biological information. New Perspectives, Proceedings of the Symposium (Cornell University, 2011). World Scientific, Singapore, pp 87–104

Basener WF, Sanford JC (2018) The fundamental theorem of natural selection with mutations. J Math Biol 76:1589–1622

Beerenwinkel N, Antal T, Dingli D, Traulsen A, Kinzler KW, Velculescu VE, Vogelstein B, Nowak MA (2007) Genetic progression and the waiting time to cancer. PLoS Comput Biol 3:2239–2246

Benaïm M, Schreiber SJ (2019) Persistence and extinction for stochastic ecological models with internal and external variables. J Math Biol 79:393–431

Bernardi O, Florio A (2019) Existence of Lipschitz continuous Lyapunov functions strict outside the strong chain recurrent set. Dyn Syst 34:71–92

Birch J (2016) Natural selection and the maximization of fitness. Biol Rev 91:712–727

Block GL, Allen LJS (2000) Population extinction and quasi-stationary behavior in stochastic density-dependent structured models. Bull Math Biol 62:199–228

Block L, Franke JE (1985) The chain recurrent set, attractors, and explosions. Ergodic Theory Dyn Syst 5:321–327

Boenkost F, González Casanova A, Pokalyuk C, Wakolbinger A (2019) Haldane's formula in Cannings models: the case of moderately weak selection, preprint is available at arXiv:1907.10049

Bof N, Carli R, Schenato L (2018) Lyapunov theory for discrete time systems, technical report is available at arXiv:1809.05289

Bovier A, den Hollander F (2015) Metastability: a potential-theoretic approach. Springer, Berlin

Bratus AS, Semenov YS, Novozhilov AS (2018) Adaptive fitness landscape for replicator systems: to maximize or not to maximize. Math Model Nat Phenom 13:25

Brémaud P (1998) Markov Chains: Gibbs fields, Monte Carlo simulation, and queues. Springer, Berlin

Broekman MJE, Muller-Landau HC, Visser MD, Jongejans E, Wright SJ, de Kroon H (2019) Signs of stabilisation and stable coexistence. Ecol Lett 22:1957–1975

Broom M, Rychtář J (2013) Game-Theoretical Models in Biology. CRC Press, Cambridge

Buckley FM, Pollett PK (2010) Limit theorems for discrete-time metapopulation models. Probab Surv 7:53–83

Cai J, Tan T, Chan SHJ (2019) Bridging traditional evolutionary game theory and metabolic models for predicting Nash equilibrium of microbial metabolic interactions, preprint is available at https://www.biorxiv.org/content/early/2019/05/14/623173.full.pdf

Carja O, Liberman U, Feldman MW (2014) Evolution in changing environments: modifiers of mutation, recombination, and migration. Proc Natl Acad Sci USA 111:17935–17940

Caulin AF, Graham TA, Wang LS, Maley CC (2015) Solutions to Peto's paradox revealed by mathematical modelling and cross-species cancer gene analysis. Philos Trans R Soc Lond Ser B 370:20140222

Cavaliere M, Feng S, Soyer OS, Jiménez JI (2017) Cooperation in microbial communities and their biotechnological applications. Environ Microbiol 19:2949–2963

Chalub FACC, Souza MO (2014) The frequency-dependent Wright–Fisher model: diffusive and non-diffusive approximations. J Math Biol 68:1089–1133

Chalub FACC, Souza MO (2017) On the stochastic evolution of finite populations. J Math Biol 75:1735–1774

Chapkin RS et al. Propagation of a signal within a gut gene-microbial ecological network with an interface layer, work in progress

Chen H, Kardar M (2019) Mean-field computational approach to HIV dynamics on a fitness landscape, preprint is available at https://www.biorxiv.org/content/10.1101/518704v2

Chesson P (2018) Updates on mechanisms of maintenance of species diversity. J Ecol 106:1773–1794

Christie JR, Beekman M (2017) Selective sweeps of mitochondrial DNA can drive the evolution of uniparental inheritance. Evolution 71:2090–2099

Chuang JS, Rivoire O, Leibler S (2009) Simpson's paradox in a synthetic microbial system. Science (N S) 323:272–275

Conley C (1978) Isolated invariant sets and the Morse index. In: C. B. M. S. Regional Conference Series in Math, vol 38. American Mathematical Society

Constable GWA, Rogers T, McKane AJ, Tarnita CE (2016) Demographic noise can reverse the direction of deterministic selection. Proc Natl Acad Sci USA 113:E4745–E4754

Coron C, Méléard S, Villemonais D (2019) Impact of demography on extinction/fixation events. J Math Biol 78:549–577

Cressman R, Tao Y (2014) The replicator equation and other game dynamics. Proc Natl Acad Sci USA 111:10810–10817

Cvijović I, Good BH, Desai MM (2018) The effect of strong purifying selection on genetic diversity. Genetics 209:1235–1278

Doorn EV, Pollett P (2013) Quasi-stationary distributions for discrete state models. Eur J Oper Res 230:1–14

Datta RS, Gutteridge A, Swanton C, Maley CC, Graham TA (2013) Modelling the evolution of genetic instability during tumour progression. Evol Appl 6:20–33

Dean A, Shnerb NM (2019) Stochasticity-induced stabilization in ecology and evolution, preprint is available at https://www.biorxiv.org/content/10.1101/725341v1

Der R, Epstein CL, Plotkin JB (2011) Generalized population models and the nature of genetic drift. Theor Popul Biol 80:80–99

de Visser JAGM, Krug J (2014) Empirical fitness landscapes and the predictability of evolution. Nat Rev Genet 15:480–490

Dewar RC, Sherwin WB, Thomas E, Holleley CE, Nichols RA (2011) Predictions of single-nucleotide polymorphism differentiation between two populations in terms of mutual information. Mol Ecol 20:3156–66

Doebeli M, Ispolatov Y, Simon B (2017) Towards a mechanistic foundation of evolutionary theory. Elife 6:e23804

dos Reis M (2015) How to calculate the non-synonymous to synonymous rate ratio of protein-coding genes under the Fisher-Wright mutation-selection framework. Biol Lett 11:20141031

Edwards AWF (2016) Analysing nature's experiment: Fisher's inductive theorem of natural selection. Theor Popul Biol 109:1–5

Ewens WJ (2011) What is the gene trying to do? Br J Philos Sci 62:155–176

Ewens WJ (2012) Mathematical population genetics (I. Theoretical introduction), 2nd edn. Interdisciplinary Applied Mathematics Series, vol 27. Springer, Berlin

Ewens WJ, Lessard S (2015) On the interpretation and relevance of the Fundamental Theorem of Natural Selection. Theor Popul Biol 104:59–67

Fathi A, Pageault P (2015) Aubry–Mather theory for homeomorphisms. Ergodic Theory Dyn Syst 35:1187–1207

Fathi A, Pageault P (2019) Smoothing Lyapunov functions. Trans Am Math Soc 371:1677–1700

Faure M, Schreiber S (2014) Quasi-stationary distributions for randomly perturbed dynamical systems. Ann Appl Probab 24:553–598

Ferguson TS (1967) Mathematical statistics: a decision theoretic approach. Academic Press, New York

Franke JE, Selgrade JF (1976) Abstract $\omega$-limit sets, chain recurrent sets, and basic sets for flows. Proc Am Math Soc 60:309–316

Franks J Notes on chain recurrence and Lyapunonv functions, (unpublished) lecture notes, available at https://arxiv.org/pdf/1704.07264.pdf

Garay BM, Hofbauer J (2003) Robust permanence for ecological differential equations, minimax, and discretizations. SIAM J Math Anal 34:1007–1039

Garcia V, Glassberg EC, Harpak A, Feldman MW (2018) Clonal interference can cause wavelet-like oscillations of multilocus linkage disequilibrium. J R Soc Interface 15:20170921

Gavrilets S (2010) High-dimensional fitness landscapes and speciation. In: Pigliucci M, Müller GB (eds) Evolution–the extended synthesis. MIT Press, Cambridge, pp 45–79

Gokhale CS, Traulsen A (2014) Evolutionary multiplayer games. Dyn Games Appl 4:468–488

Gompert Z (2016) Bayesian inference of selection in a heterogeneous environment from genetic time-series data. Mol Ecol 25:121–34

Good BH, Martis S, Hallatschek O (2018) Directional selection limits ecological diversification and promotes ecological tinkering during the competition for substitutable resources, preprint is available at https://www.biorxiv.org/content/10.1101/292821v1

Goudenège L, Zitt P-A (2015) A Wright–Fisher model with indirect selection. J Math Biol 71:1411–1450

Grosjean N, Huillet T (2017) Wright–Fisher-like models with constant population size on average. Int J Biomath 10:1750078

Gyllenberg M, Silvestrov DS (2008) Quasi-stationary phenomena in nonlinearly perturbed stochastic systems, De Gruyter Expositions in Mathematics, vol 44. Walter de Gruyter

Harmand N, Federico V, Hindre T, Lenormand T (2019) Nonlinear frequency-dependent selection promotes long-term coexistence between bacteria species. Ecol Lett 22:1192–1202

Harper M, Fryer D (2016) Stationary stability for evolutionary dynamics in finite populations. Entropy 18:316

Hartle DL, Clark AG (2007) Principles of population genetics, 4th edn. Sinauer, Sunderland

Hintze A, Olson RS, Adami C, Hertwig R (2015) Risk sensitivity as an evolutionary adaptation. Sci Rep 5:8242

Hobolth A, Sireén J (2016) The multivariate Wright-Fisher process with mutation: moment-based analysis and inference using a hierarchical Beta model. Theor Popul Biol 108:36–50

Hoeffding W (1963) Probability inequalities for sums of bounded random variables. J Am Stat Assoc 58:13–30

Hofbauer J (1985) The selection mutation equation. J Math Biol 23:41–53

Hofbauer J (2011) Deterministic evolutionary game dynamics. In: Sigmund K (ed) Evolutionary game dynamics. Proc. Sympos. Appl. Math., vol. 69, AMS Short Course Lecture Notes. Amer. Math. Soc, pp 61–79

Hofbauer J, Sigmund K (1998) Evolutionary games and population dynamics. Cambridge University Press, Cambridge

Hofbauer J, Sigmund K (2003) Evolutionary game dynamics. Bull Am Math Soc (N S) 40:479–519

Hofrichter J, Jost J, Tran T (2017) Information geometry and population genetics. The mathematical structure of the Wright–Fisher model. Springer, Berlin

Houchmandzadeh B (2015) Fluctuation driven fixation of cooperative behavior. Biosystems 127:60–66

Hummert S, Bohl K, Basanta D, Deutsch A, Werner S, Theißen G, Schroeter A, Schuster S (2014) Evolutionary game theory: cells as players. Mol BioSyst 10:3044

Hummert S, Glock C, Lang SN, Hummert C, Skerka C, Zipfel PF, Germerodt S, Schuster S (2018) Playing "hide-and-seek" with factor H: game-theoretical analysis of a single nucleotide polymorphism. J R Soc Interface 15:20170963

Hurley M (1998) Lyapunov functions and attractors in arbitrary metric spaces. Proc Am Math Soc 126:245–256

Hutchinson GE (1961) The paradox of the plankton. Am Nat 95:137–145

Huttegger S, Skyrms B, Tarrès P, Wagner E (2014) Some dynamics of signaling games. Proc Natl Acad Sci USA 111:10873–10880

Hutson V, Schmitt K (1992) Permanence and the dynamics of biological systems. Math Biosci 111:1–71

Iglehart DL (1972) Extreme values in the $GI/G/1$ queue. Ann Math Stat 43:627–635

Iizuka M (2010) Effective population size of a population with stochastically varying size. J Math Biol 61:359–375

Imhof LA, Nowak MA (2006) Evolutionary game dynamics in a Wright–Fisher process. J Math Biol 52:667–681

Iwasa Y, Michor F (2011) Evolutionary dynamics of intratumor heterogeneity. PLoS ONE 6:e17866

John S, Seetharaman S (2016) Exploiting the adaptation dynamics to predict the distribution of beneficial fitness effects. PLoS ONE 11:e0151795

Kang Y, Chesson P (2010) Relative nonlinearity and permanence. Theor Popul Biol 78:26–35

Karlin S, McGregor J (1964) Direct product branching processes and related Markov chains. Proc Natl Acad Sci USA 51:598–602

Karr AF (1975) Weak convergence of a sequence of Markov chains. Z Wahrsch Verw Gebiete 33:41–48

Kingman JFC (1961) A mathematical problem in population genetics. Proc Camb Phil Soc 57:574–582

Klebaner FC, Nerman O (1994) Autoregressive approximation in branching processes with a threshold. Stochastic Process Appl 51:1–7

Koopmann B, Müller J, Tellier A, Živković D (2017) Fisher–Wright model with deterministic seed bank and selection. Theor Popul Biol 114:29–39

Kroumi D, Lessard S (2015) Evolution of cooperation in a multidimensional phenotype space. Theor Popul Biol 102:60–75

Kurokawa S, Ihara Y (2009) Emergence of cooperation in public goods games. Proc R Soc B 276:1379–1384

Landim C (2019) Metastable Markov chains. Probab Surv 16:143–227

Lawler GF (2006) Introduction to stochastic processes, 2nd edn. Chapman & Hall/CRC, Cambridge

Li D, Cheng F (2019) The extinction and persistence of tumor evolution influenced by external fluctuations and periodic treatment. Qual Theory Dyn Syst (to appear)

Li L, Chesson P (2016) The effects of dynamical rates on species coexistence in a variable environment: the paradox of the plankton revisited. Am Nat 188:46–58

Losert V, Akin E (1983) Dynamics of games and genes: discrete versus continuous time. J Math Biol 17:241–251

Mahdipour-Shirayeh A, Darooneh AH, Long AD, Komarova NL, Kohandel M (2017) Genotype by random environmental interactions gives an advantage to non-favored minor alleles. Sci Rep 7:5193

Mandel SPH (1959) The stability of a multiple allelic system. Heredity 13:289–302

McAvoy A, Fraiman N, Hauert C, Wakeley J, Nowak MA (2018) Public goods games in populations with fluctuating size. Theor Popul Biol 121:72–84

Méléard S, Villemonais D (2012) Quasi-stationary distributions and population processes. Probab Surv 9:340–410

Nagylaki T (1986) The Gaussian approximation for random genetic drift. In: Karlin S, Nevo E (eds) Evolutionary processes and theory. Academic Press, New York, pp 629–642

Nagylaki T (1990) Models and approximations for random genetic drift. Theor Popul Biol 37:192–212

Norton DE (1995) The fundamental theorem of dynamical systems. Comment Math Univ Carolin 36:585–597

Nourmohammad A, Otwinowski J, Plotkin JB (2016) Host-pathogen coevolution and the emergence of broadly neutralizing antibodies in chronic infections. PLoS Genet 12:e1006171

Nowak MA, Sigmund K (2004) Evolutionary dynamics of biological games. Science 303:793–799

Obolski U, Ram Y, Hadany L (2018) Key issues review: evolution on rugged adaptive landscapes. Rep Prog Phys 81:012602

Orr HA (2009) Fitness and its role in evolutionary genetics. Nat Rev Genet 10:531–539

Pageault P (2009) Conley barriers and their applications: chain-recurrence and Lyapunov functions. Topol Appl 156:2426–2442

Panageas I, Srivastava P, Vishnoi NK (2016) Evolutionary dynamics in finite populations mix rapidly, In: Krauthgamer R (ed) Proceedings of the twenty-seventh annual ACM-SIAM symposium on discrete algorithms. SIAM, pp 480–497

Papadimitriou C, Piliouras G (2016) From Nash equilibria to chain recurrent sets: solution concepts and topology. In: Sudan M (ed) ITCS'16-Proceedings of the 2016 ACM conference on innovations in theoretical computer science. ACM, New York, pp 227–235

Park S-C, Krug J (2007) Clonal interference in large populations. Proc Natl Acad Sci USA 104:18135

Park HJ, Traulsen A (2017) Extinction dynamics from metastable coexistences in an evolutionary game. Phys Rev E 96:042412

Parra-Rojas C, Challenger JD, Fanelli D, McKane AJ (2014) Intrinsic noise and two-dimensional maps: Quasicycles, quasiperiodicity, and chaos. Phys Rev E 90:032135

Perko L (1991) Differential equations and dynamical systems, 3rd edn. Springer, Berlin

Proulx SR (2011) The rate of multi-step evolution in Moran and Wright–Fisher populations. Theor Popul Biol 80:197–207

Queller DC (2017) Fundamental theorems of evolution. Am Nat 189:345–353

Rao CR (2001) Linear statistical inference and its applications, 2nd edn. Wiley, New York

Rohlfs M, Hoffmeist TS (2003) An evolutionary explanation of the aggregation model of species coexistence. Proc Biol Sci 270:33–35

Ruppin E, Papin JA, de Figueiredo LF, Schuster S (2010) Metabolic reconstruction, constraint-based analysis and game theory to probe genomescale metabolic networks. Curr Opin Biotechnol 21:502–510

Samoilov MS, Arkin AP (2006) Deviant effects in molecular reaction pathways. Nat Biotechnol 24:1235–40

Sandholm WH (2010) Population games and evolutionary dynamics, series on economic learning and social evolution, vol 8. MIT Press, Cambridge

Schneider KA (2010) Maximization principles for frequency-dependent selection II: the one-locus multi-allele case. J Math Biol 61:95–132

Schreiber SJ, Levine JM, Godoy O, Kraft NJB, Hart SP (2018) Does deterministic coexistence theory matter in a finite world? Insights from serpentine annual plants, preprint is available at https://www.biorxiv.org/content/10.1101/290882v1

Sigmund K (1986) A survey of replicator equations. In: Casti JL, Karlqvist A (eds) Complexity, language, and life: mathematical approaches, Biomathematics, vol 16. Springer, Berlin, pp 88–104

Silverman S (1992) On maps with dense orbits and the definition of chaos. Rocky Mountain J Math 22:353–375

Simonsen Speed N, Balding DJ, Hobolth A (2019) A general framework for moment-based analysis of genetic data. J Math Biol 78:1727–1769

Spichtig M, Kawecki TJ (2004) The maintenance (or not) of polygenic variation by soft selection in heterogeneous environments. Am Nat 164:70–84

Steinsaltz D, Evans SN (2004) Markov mortality models: implications of quasistationarity and varying initial distributions. Theor Popul Biol 65:319–337

Stephens C (2014) Selection, drift, and the "forces" of evolution. Philos Sci 71:550–570

Svensson EI, Connallon T (2019) How frequency-dependent selection affects population fitness, maladaptation and evolutionary rescue. Evol Appl 12:1243–1258

Tarnita CE (2017) The ecology and evolution of social behavior in microbes. J Exp Biol 220:18–24

Taylor C, Nowak M (2006) Evolutionary game dynamics with non-uniform interaction rates. Theor Popul Biol 69:243–252

Traulsen A, Claussen JC, Hauert C (2005) Coevolutionary dynamics: from finite to infinite populations. Phys Rev Lett 95:238701

Traulsen A, Reed FA (2012) From genes to games: cooperation and cyclic dominance in meiotic drive. J Theor Biol 299:120–125

Vallier M, Abou Chakra M, Hindersin L, Linnenbrink M, Traulsen A, Baines JF (2017) Evaluating the maintenance of disease-associated variation at the blood group-related gene B4galnt2 in house mice. BMC Evol Biol 17:187

Wasserman L (2004) All of statistics: a concise course in statistical inference. Springer texts in statistics. Springer, Berlin

Waxman D, Loewe L (2010) A stochastic model for a single click of Muller's ratchet. J Theor Biol 264:1120–1132

West SA, Diggle SP, Buckling A, Gardner A, Griffin AS (2007) The social lives of microbes. Annu Rev Ecol Evol Syst 38:53–77

Xu S, Jiao S, Jiang P, Ao P (2014) Two-time-scale population evolution on a singular landscape. Phys Rev E 89:012724

Zeng Q, Wu S, Sukumaran J, Rodrigo A (2017) Models of microbiome evolution incorporating host and microbial selection. Microbiome 5:127

Zhang Y, Fu F, Wu T, Xie G, Wang L (2013) A tale of two contribution mechanisms for nonlinear public goods. Sci Rep 3:2021

Zhou D, Qian H (2011) Fixation, transient landscape, and diffusion dilemma in stochastic evolutionary game dynamics. Phys Rev E 84:31907

Zinger T, Gelbart M, Miller D, Pennings PS, Stern A (2019) Inferring population genetics parameters of evolving viruses using time-series data. Virus Evol 5:vez011

Zomorrodi AR, Segrè D (2017) Genome-driven evolutionary game theory helps understand the rise of metabolic interdependencies in microbial communities. Nat Commun 8:1563