Electronic Journal of Probability

Electron. J. Probab. **27** (2022), article no. 85, 1–23.
ISSN: 1083-6489 https://doi.org/10.1214/22-EJP807

Fixation time of the rock-paper-scissors model: rigorous results in the well-mixed setting

Eric Foxall* Bilal Madani* Adam Roemer*

Abstract

The rock-paper-scissors model simulates the effect of cyclic dominance in a finite population of size N and has received considerable attention in applied literature. In the well-mixed version of the model, population densities fluctuate around periodic orbits of a deterministic ODE approximation, and for large N the time to fixation (complete dominance by one species) has been observed by simulation to be approximately τN where τ is a positive, finite random variable. We give a rigorous proof of this observation by establishing a slow diffusion limit for a conserved quantity of the deterministic approximation, together with a careful analysis of the behaviour at the boundary, both in the limit as $N \to \infty$, and for large but finite N.

Keywords: rock-paper-scissors model; density-dependent Markov chain; diffusion limit; heteroclinic cycle; stochastic averaging.
MSC2020 subject classifications: 60K35; 92D55.
Submitted to EJP on July 2, 2021, final version accepted on June 8, 2022.

1 Introduction and main results

In this article we give a rigorous formulation and proof of some results concerning the asymptotic behaviour of a stochastic neutral model of cyclic dominance in a finite, well-mixed population, that is commonly known as the rock-paper-scissors model. For a review of work on the model, in many settings besides the one considered here, see the survey article [16].

The model has three types, that we denote $0, 1$ and 2, and three positive parameters, r_0, r_1 and r_2. The total population size N is constant over time. There are three possible interactions: for each i, an individual (ind) of type i converts an individual of type $i-1$ (modulo 3) to its own type, with associated rate constant r_i. In the notation of chemical reactions these interactions can be summarized as follows:

$$0+1 \xrightarrow{r_1} 1+1, \quad 1+2 \xrightarrow{r_2} 2+2 \ \text{ and } \ 2+0 \xrightarrow{r_0} 0+0.$$

*University of British Columbia (UBC), Okanagan Campus. E-mail: eric.foxall@ubc.ca

Assuming a well-mixed population, the rate constants are translated into rates as follows: for each i, each ind of type i converts each ind of type $i-1$ to its own type at rate r_i/N. The normalization by $1/N$ ensures that the total rate at which an ind interacts with others depends only on the population density $x=(x_0,x_1,x_2)$, defined by $x_i := X_i/N$. For example, an ind of type i interacts with others at a total rate $r_iX_{i-1}/N + r_{i+1}X_{i+1}/N = r_ix_{i-1}+r_{i+1}x_{i+1}$, from either converting an ind of type $i-1$ or being converted by an ind of type $i+1$. Letting X_i denote the number of inds of type i, the vector $X=(X_0,X_1,X_2)$ is a continuous-time Markov chain on the state space $S_N := \{X \in \mathbb{N}^3 \colon \sum_i X_i = N\}$ with the following transitions:

$$\begin{aligned} &\text{for each } \ i \in \{0,1,2\} \mod 3, \\ &X \to X + e_i - e_{i-1} \ \text{ at rate } \ r_iX_{i-1}X_i/N, \end{aligned} \tag{1.1}$$

where e_i is the i^{th} unit vector in $\mathbb{R}^3$, i.e., $e_0=(1,0,0)$, $e_1=(0,1,0)$ and $e_2=(0,0,1)$, and i is always understood modulo 3. System (1.1) is a *density-dependent Markov chain (DDMC)* in the sense of [15]. Roughly speaking, DDMCs correspond to well-mixed models of interacting populations with a system size parameter (the parameter here is N), and are such that the population density $x := X/N$ has both a deterministic and a diffusive approximation on the original time scale, as described in [15]. Understanding the long-term behaviour of DDMCs typically requires a careful analysis of both approximations, which is the case in this article. For models with absorbing states that correspond to unstable equilibria of the deterministic approximation, large deviations analysis may also be required; see [7] for an example in the case of logistic growth.

Write $X^{(N)}$ for sample paths of (1.1), similarly $x^{(N)} = X^{(N)}/N$, when it is necessary to emphasize the dependence on N. We are interested in the behaviour of $X^{(N)}$ in the limit as $N\to\infty$, and specifically in

(i) the fixation time, i.e., the amount of time until one type has taken over the population, or equivalently, until X reaches one of the "single-type" states Ne_i:

$$\tau_N^o := \inf\{t\colon X^{(N)}(t) = Ne_i \text{ for some } i \in \{0,1,2\}\},$$

(ii) and the fixation probability, i.e., which type is the eventual winner:

$$p_i^{(N)} := \mathbb{P}(X^{(N)}(\tau_N^o) = Ne_i).$$

In this article we study the fixation time, for which the main result is Theorem 1.1 below. The fixation probability will be addressed in a later work, so for now, we will simply state the desired result as Conjecture 1.2. The result of Theorem 1.1 is implicit in [17] and [4], though not formally stated or proved. Similarly, Conjecture 1.2 can be found in [9] and [2]. Let $x^* = (x_0^*, x_1^*, x_2^*)$ be the vector with entries $x_i^* = r_{i-1}/\sum_j r_j$, which as described below is the coexistence equilibrium of the deterministic approximation (3.5) of $x^{(N)}$, and let $\overset{\mathrm{d}}{\to}$ denote convergence in distribution.

Theorem 1.1. *Suppose that $X^{(N)}(0)/N \to x^*$ as $N\to\infty$. Then there is a non-degenerate random variable τ in $(0,\infty)$ such that*

$$\tau_N^o/N \overset{\mathrm{d}}{\to} \tau \ \text{ as } \ N\to\infty,$$

To our knowledge, this is the first result on asymptotic behaviour of a density-dependent Markov chain whose deterministic approximation has a neutrally stable heteroclinic cycle. The hypothesis of Theorem 1.1 can be weakened to $P(x^{(N)}) \to p \in (0,p^*]$ as $N\to\infty$, see Theorem 4.1, and the distribution of the limit τ, which depends on p, is stochastically non-increasing in p.

Conjecture 1.2. *Suppose that $X^{(N)}(0)/N \to x^*$ as $N \to \infty$. Then*

$$\lim_{N\to\infty} p_i^{(N)} = \begin{cases} 1/\#\{i\colon r_i = \min_j r_j\} & \textit{if } r_i = \min_j r_j \\ 0 & \textit{otherwise.} \end{cases}$$

To summarize both statements in a few words: if the system is initialized near the interior equilibrium x^*, then

(i) the fixation time is of order N, with a non-degenerate limiting distribution, and
(ii) the final state $X^{(N)}(\tau_N^o)$ is evenly distributed over types with minimal rates.

In the above references, Theorem 1.1 and Conjecture 1.2 are demonstrated using numerical simulations, and explained with the help of two large-N limit processes:

1. a system of ODEs on $S := \{x \in \mathbb{R}_+^3\colon \sum_i x_i = 1\}$, given in (3.5), that approximates $x^{(N)}$ on the original (fast) time scale t and
2. an SDE on an interval $[0, p^*]$, given in (3.8), that approximates the process $P(x^{(N)})$ on the (slow) time scale Nt,

where $P(x) := \prod_i x_i^{r_{i-1}}$ is constant along solutions of (3.5) and $p^* = P(x^*)$ is the maximal value of P on S. Here is a very brief explanation of the results.

- The phase space S is foliated by the level curves $L_p := \{x \in S\colon P(x) = p\}$ of P, with L_0 comprising the boundary, L_{p^*} the coexistence equilibrium and each $L_p, 0 < p < p^*$ a periodic orbit of (3.5).

- On the fast time scale, $x^{(N)}$ closely follows solutions of (3.5) along the level curves L_p, with small fluctuations of size $1/\sqrt{N}$.

- On the slow time scale, the process $Y^{(N)} := P(x^{(N)}(Nt))$ experiences fluctuations of constant size, with drift and diffusion coefficients given by the orbital average, with respect to (3.5), of the instantaneous coefficients, eventually hitting 0 which corresponds to $x^{(N)}$ hitting $\partial S := \{x \in S\colon x_i = 0 \text{ for some } i\}$, the boundary of S relative to the affine space $\{x \in \mathbb{R}^3\colon \sum_i x_i = 1\}$.

- From the point x_0 where $x^{(N)}$ first hits S^o, within $O(\log N)$ time $x^{(N)}$ reaches the absorbing state prescribed by the solution of (3.5) with initial value x_0.

- Letting Y denote the solution of the SDE approximation of $Y^{(N)}$, $\tau = \inf\{t\colon Y(t) = 0\}$ and $\lim_{N\to\infty} p_i^{(N)}$ is predicted by the relative distance, as $p \to 0$, from L_p to the boundary edges $\{x \in S\colon x_i = 0,\ 0 < x_{i+1} < 1\}$.

The rest of the article is organized as follows. In Section 2 we discuss some related work. In Section 3 we compute the deterministic approximation (3.5) for $x^{(N)}$ and the slow diffusion approximation (3.8) for $Y^{(N)}$ and study their solutions. The deterministic approximation has a single interior equilibrium, a family of interior periodic orbits, and a single heteroclinic cycle, given by the level sets of P. The slow diffusion takes place on an interval $[0, p^*]$, with 0 an accessible boundary and p^* an entrance boundary (Lemma 3.5), and from any initial value in $[0, p^*]$ reaches 0 in finite time (Proposition 3.4). In Section 4 we use stochastic averaging (Lemma 4.6) to show convergence of $Y^{(N)}$ to the limiting diffusion on the interior of $[0, p^*]$ (Proposition 4.2), then analyze the boundary behaviour of $Y^{(N)}$ (Proposition 4.3) and combine the two to obtain convergence of fixation time.

2 Related work

As mentioned in the Introduction, our model has a diffusion approximation, which as shown in (3.4) takes the form

$$dx = F(x)dt + \epsilon_N \sqrt{G(x)}dB, \tag{2.1}$$

where $x \in S$, $\epsilon_N = 1/\sqrt{N}$, $F : S \to \mathbb{R}^3$, $G : S \to M_+(\mathbb{R}, 3)$ are smooth functions computed from the transition rates in (3.3) ($M_+(\mathbb{R}, 3)$ denotes 3×3 positive semidefinite matrices) and B is a standard Brownian motion in $\mathbb{R}^3$. The deterministic part $x' = F(x)$ is sometimes called the mean field equations, or replicator equations, of the model. The diffusion part $dx = \epsilon_N \sqrt{G(x)}dB$ captures the finite-size effects of discernible jumps in the population vector, reflecting the demographic stochasticity inherent in the Markov chain formulation. For each N, both the Markov chain and its diffusion approximation hit the boundary of S in finite time, so it is of interest to know how quickly this happens as a function of N, as we do in this article.

If instead we want to model random fluctuations in the per-capita growth rates, which is known as environmental stochasticity, we can use a model of the form

$$dx = x \circ (R(x)dt + Q(x)dB), \tag{2.2}$$

where $xR(x) = F(x)$, $\circ$ is Hadamard (entrywise) product, Q is any Lipschitz continuous $3 \times d$ matrix-valued function for some integer $d \geq 1$ and B is standard BM on $\mathbb{R}^d$. For each $i \in \{0, 1, 2\}$ the diffusion term in dx_i is $x_i \sum_j Q_{ij}(x)dB_j$ which has diffusion coefficient $\sigma_i^2(x) = x_i^2 \sum_j Q_{ij}(x)^2 = O(x_i^2)$, which implies the set $x_i = 0$ inaccessible for each i, i.e., if x_0 is supported on the interior of S then (x_t) never hits the boundary. This is a property shared by the model's deterministic counterpart $x' = F(x)$ and it means that broadly similar results can be provided, provided that the objects of study are suitably generalized. For example, it has long been known (see [1] for an overview and references) that in the deterministic case, permanence (existence of a globally stable interior attractor) and impermanence (attracting boundary) can be partly characterized by invasibility conditions, essentially, whether or not a new species introduced at low density into an equilibrium population can disturb that equilibrium. In [1] the authors show these results are robust to small enough perturbations of the form (2.2), provided that in the stochastic case we understand an attractor to mean a stationary distribution and global stability of μ to mean that the distribution of x_t converges to μ as $t \to \infty$, for any distribution of x_0 supported on the interior; an irreducibility assumption is needed to obtain palatable results. In [18], invasibility conditions are defined for all systems of the form (2.2), extending the deterministic definition by accounting for the effect of the diffusion term. These conditions are shown to imply persistence, defined as "with probability 1, (x_t) spends $o(1)$ proportion of time within distance $\leq \delta$ of the boundary, as $\delta \to 0^+$", and with an irreducibility assumption are shown to imply permanenence in the previous sense. The results of [18] and other works are generalized in several directions in [11].

In [12] the authors consider three-dimensional systems of the form (2.2) and give a complete classification of the possible dynamics. The novel behaviour in three dimensions (i.e., not occurring in dimensions 1 and 2) is a rock-paper-scissors type of dynamic, characterized by a labelling of the 3 types such that the invasion rate $\lambda(i, j)$ of type i into a type j population is positive whenever $j = i + 1 \mod 3$ and negative whenever $j = i - 1 \mod 3$. Letting $\lambda_\pm(i) = \lambda(i, i \pm 1)$, they find that permanence holds when $\prod_i \lambda_+(i) > \prod_i |\lambda_-(i)|$ and that impermanence holds whenever the inequality is reversed. These results are generalized in [10] to allow parameters to change according to a piecewise deterministic Markov chain, and a fast algorithm for computing the stationary distributions is also given. I expect that for density-dependent Markov chains and SDEs of the form (2.1), since $\epsilon_N \to 0$ as $N \to \infty$ a similar result holds if we use the invasion rates defined for the corresponding deterministic system $x' = F(x)$, and define permanence and impermanence according to whether the time to hit the boundary is long or short as a function of N. Denoting that time by τ_N, typically "long" means that $\liminf_N \mathbb{P}(\tau_N \geq e^{cn}) > 0$ for some $c > 0$ while "short" means that for large enough $C > 0$, $\mathbb{P}(\tau_N \geq C\log(N)) \to 0$ as $N \to \infty$. For an example of both, see [7] where the logistic model is considered. For the model considered in the present article, the invasion rates are $\lambda_+(i-1) = -\lambda_-(i) = r_i$ so $\prod_i \lambda_+(i) = \prod_i |\lambda_-(i)|$ and the above condition is inconclusive, which agrees with our result that τ_N is of order N, an intermediate case between long and short.

3 Limit processes

We begin with an informal discussion of approximation of a Markov chain and translate terms and equations between the conventions of mathematics and physics.

3.1 Generalities

Suppose X is a Markov chain on state space $S \subset \mathbb{R}^d$ with transition rate $q(x,y)$ from x to y, $(x,y) \in S \times S$, and that X changes mostly by small jumps. Define the respective drift and diffusion coefficients μ, σ^2 on S by

$$\mu(x) = \sum_y (y-x)q(x,y) \quad \text{and} \quad \sigma^2(x) = \sum_y (y-x)(y-x)^\top q(x,y). \tag{3.1}$$

For each x, $\sigma^2(x)$ is a positive semidefinite $d \times d$ matrix. We can think of μ as the "average velocity" of X, or more precisely, the instantaneous rate of change of the expected value of X, conditioned on past information. Similarly, we can think of σ^2 as the "rate of accumulation of variance" of X, or more precisely, the instantaneous rate of change of the variance of X, conditioned on past information.

If μ and σ^2 can be extended to an open connected set $D \subset \mathbb{R}^d$ containing S, then a suitable diffusion approximation of X is the SDE on D given by

$$dX = \mu(X)dt + \sigma(X)dB,$$

where B is a d-dimensional standard Brownian motion and for $x \in D$, $\sigma(x)$ is the positive square root of the $d \times d$ matrix $\sigma^2(x)$. If, in addition, σ is small, then a suitable deterministic approximation of X is the ODE

$$X' = \mu(X).$$

To connect these with the references, we make a few linguistic notes. In physics and chemistry, stochastic processes are usually studied indirectly via their probability distributions $p(x,t) := \mathbb{P}(X_t = x)$. For a continuous-time Markov chain on a finite or countable state space S (with some additional requirements when S is infinite), the vector-valued function $t \mapsto (p(x,t))_{x \in S}$ solves a system of ODEs known to physicists as the master equation (see, for example, [19]) and to mathematicians as the forward Kolmogorov equations. The above SDE approximation is known to physicists as the Langevin equation corresponding to the Fokker-Planck (FP) approximation of the master equation. The FP approximation is a PDE which, itself, is the forward Kolmogorov equation of the SDE approximation. The ODE approximation is often referred to as the mean-field equation.

3.2 The population density process $x^{(N)}$

Recall that $x^{(N)} = X^{(N)}/N$, where X satisfies (1.1). Let $\Delta(i) = -e_{i-1} + e_i$ and $q_i(x) = r_i x_{i-1} x_i$, then $x^{(N)}$ has the following transitions:

$$\begin{aligned}&\text{for each } \ i \in \{0,1,2\} \mod 3, \\ &x^{(N)} \to x^{(N)} + \Delta(i)/N \ \text{ at rate } \ N q_i(x). \end{aligned} \tag{3.2}$$

Notice that jumps in $x^{(N)}$ are of size $2/N$ (in the ℓ^1 norm, say), so are small when N is large. Referring to (3.1) and defining

$$F(x) = \sum_{i=0}^{2} \Delta(i) q_i(x) \quad \text{and} \quad G(x) = \sum_{i=0}^{2} \Delta(i)\Delta(i)^\top q_i(x), \tag{3.3}$$

a suitable diffusion approximation of $x^{(N)}$ is the SDE

$$dx = F(x)dt + \epsilon_N \sqrt{G(x)}\, dB, \tag{3.4}$$

where $\epsilon_N = 1/\sqrt{N}$. Since ϵ_N is small when N is large, a suitable deterministic approximation of $x^{(N)}$ is $x' = F(x)$, which when written in components has the following form:

$$\begin{aligned}&\text{for each } \ i \in \{0,1,2\} \mod 3, \\ &x_i' = r_i x_{i-1} x_i - r_{i+1} x_i x_{i+1}. \end{aligned} \tag{3.5}$$

In particular, F is locally Lipschitz, so local existence and uniqueness holds. The behaviour of solutions of (3.5) has already been studied, see for example [3]-[8]. Below, we give a quick and mostly self-explanatory summary.

Recall that $S = \{x \in \mathbb{R}_+^3 \colon \sum_i x_i = 1\}$. If $x(t)$ solves (3.5) then $\frac{d}{dt}\sum_i x_i(t) = 0$, and since $x_i' \geq -r_{i+1} x_i$ on S, $x_i(0) \geq 0$ implies $x_i(t) \geq 0$ for $t \geq 0$, so S is forward invariant and in particular, global existence and uniqueness holds. To verify that $P(x) := \prod_i x_i^{r_{i-1}}$ is constant on solutions of (3.5), let $x(t)$ solve (3.5). If $P(x(0)) = 0$ then $x_i(0) = 0$ for some i, implying $x_i(t) = 0$ and thus $P(x(t)) = 0$ for $t > 0$. If $P(x(0)) > 0$ then in a similar way $P(x(t)) > 0$ for all $t > 0$, and in that case check it's easy to check that $\frac{d}{dt}\log P(x(t)) = 0$.

Let $S^o := \{x \in S \colon \min_i x_i > 0\}$ denote the interior of S relative to the affine space $\{x \in \mathbb{R}^3 \colon \sum_i x_i = 1\}$. Define x^* by $x_i^* = r_{i-1}/\sum_j r_j$ and let $p^* = P(x^*)$. Then x^* is both the unique equilibrium of (3.5) in S^o, and the unique maximizer of P, in S. As noted in the Introduction, the level curves $(L_p \colon p \in [0, p^*])$ are a partition of S, with $L_{p^*} = x^*$, $L_0 = \partial S = \{x \in S \colon x_i = 0 \text{ for some } i\}$ and for $0 < p < p^*$, L_p is a periodic orbit of (3.5).

3.3 Slow diffusion limit for $P(x^{(N)})$

Since P is constant on solutions of (3.5), it should vary slowly on solutions of (3.4), since the term $\epsilon_n \sqrt{G(x)} B$ is small. On the other hand, fluctuations can accumulate over time. To quantify this effect, as in [4], use Itô's formula to find that the instantaneous drift and diffusion coefficients of $P(x)$, with respect to the approximation (3.4), are

$$\begin{aligned}&\nabla P(x) F(x) + \frac{\epsilon_N^2}{2} \sum_{i,j} G_{ij}(x)(D^2 P)_{ij}(x) \quad \text{and} \\ &\epsilon_N^2 \sum_{i,j} \nabla P_i(x) \nabla P_j(x) G_{ij}(x).\end{aligned}$$

Since P is constant on solutions of (3.5), $\nabla P(x) F(x) = 0$ for $x \in S$. Speeding up time by a factor $1/\epsilon_N^2 = N$ multiplies both drift and diffusion by the same factor, so the process

$Y^{(N)}(t) := P(x^{(N)}(Nt))$, with respect to the approximation (3.4), has drift and diffusion coefficients μ_Y, σ_Y^2 given in terms of the value of $x^{(N)}$ by

$$\mu_Y(x) := \frac{1}{2}\sum_{i,j} G_{ij}(x)(D^2P)_{ij}(x) \quad \text{and}$$
$$\sigma_Y^2(x) := \sum_{i,j} \nabla P_i(x)\nabla P_j(x) G_{ij}(x). \tag{3.6}$$

Note that these do not depend on N, but are also not written in terms of the desired state variable Y. To fix this, first observe that in a small time increment on the time scale Nt, the value of $P(x^{(N)})$ should not change very much, whereas solutions of (3.5) initialized on any of the periodic orbits L_p, $0 < p < p^*$ complete many cycles. So, as pointed out in [4], we expect stochastic averaging to occur. In other words, letting $T(p)$ denote the period of the orbit L_p with respect to (3.5), we expect that in the large-N limit, the drift and diffusion coefficients of $Y^{(N)}$ are given in terms of the value of $Y^{(N)}$ by

$$\overline{\mu}(p) := \frac{1}{T(p)}\int_0^{T(p)} \mu_Y(\phi(t, x(p)))dt$$
$$\overline{\sigma}^2(p) := \frac{1}{T(p)}\int_0^{T(p)} \sigma_Y^2(\phi(t, x(p)))dt \tag{3.7}$$

where $t \mapsto \phi(t, x)$ is the solution of (3.5) with $\phi(0, x) = x$, $T(p)$ is the period of the orbit L_p and $x(p) \in L_p$ is arbitrary. In other words, the desired large-N limit of $Y^{(N)}$ is the SDE

$$dY = \overline{\mu}(Y)dt + \overline{\sigma}(Y)dB, \tag{3.8}$$

where $\overline{\sigma}(Y)$ is the unique positive semidefinite square root of $\overline{\sigma}^2(Y)$. To understand solutions of (3.8) we begin by proving existence and uniqueness of solutions on $(0, p^*)$, then afterward discuss behaviour at the boundaries 0 and p^*. As we show in Lemma 3.3, $\overline{\mu}, \overline{\sigma}^2$ are sufficiently regular that (3.8) has a stochastic flow of pathwise unique strong solutions, as follows.

Proposition 3.1. *Let B be a standard Brownian motion on $\mathbb{R}$ with respect to a complete filtration $\mathcal{F}$. There is a continuous function $(s, t, p) \mapsto \psi(s, t, p)$ defined for $p \in (0, p^*)$ and $0 \le s \le t \le \tau_e(s, p) := \inf\{t \ge s \colon \psi(s, t, p) \in \{0, p^*\}\}$ (e for exit) such that for each s and any $\mathcal{F}(s)$-measurable random variable ξ in $(0, p^*)$, $t \mapsto \psi(s, t, \xi)$ is the a.s. unique solution, on $[s, \tau_e(s, \xi))$, to the integral equation*

$$Y(t) = \xi + \int_s^t \overline{\mu}(Y(u))du + \int_s^t \overline{\sigma}(Y(u))dB(u). \tag{3.9}$$

Moreover, ψ has the semigroup property:

$$\text{almost surely, } \forall 0 \le s \le t \le u \le \tau_e(s, p),\ \forall p \in (0, p^*),$$
$$\psi(t, u, \psi(s, t, p)) = \psi(s, u, p). \tag{3.10}$$

The proof has a few steps, so it's included in the Appendix. To establish the desired regularity of $\overline{\mu}$ and $\overline{\sigma}^2$, we first need to do the same for the orbital period $p \mapsto T(p)$.

Lemma 3.2. *The function $p \mapsto T(p)$ is C^∞ on $(0, p^*)$.*

Proof. Recall $u_i := x_i/r_{i-1}$ and let $\Gamma = \{x \in S \colon 0 < u_0 = u_2 < u_1\}$ which is a line segment in $S^o \setminus \{x^*\}$ connecting e_1 to x^*. It is visually obvious, and can be checked with a calculation (omitted), that Γ is transverse to the level curves L_p, so for $p \in (0, p^*)$ let $x(p) \in \Gamma$ be the unique point with $P(x(p)) = p$. Then $T(p)$ solves the equation $\phi_0(T(p), x(p)) - \phi_2(T(p), x(p)) = 0$. Writing x_i' from (3.5) in terms of u,

$$x_i' = r_\pi u_i(u_{i-1} - u_{i+1}),$$

so if $u_0 = u_2 < u_1$ then

$$\begin{aligned} x_0' - x_2' &= r_\pi(u_0u_2 - u_0u_1 - (u_1u_2 - u_2u_0)) \\ &= r_\pi(2u_0u_2 - u_1(u_0 + u_2)) = 2r_\pi u_0(u_0 - u_1) < 0. \end{aligned}$$

In particular, for $p \in (0, p^*)$,

$$\partial_t(\phi_0(T(p), x(p))/r_2 - \phi_2(T(p), x(p))/r_1) \neq 0$$

so by the implicit function theorem (IFT), there is a neighbourhood U_p of $x(p)$ in S and a function $\tau_p(x)$ with $\tau_p(x(p)) = T(p)$ (note p here is still fixed) such that for $x \in U_p$,

$$\phi_0(\tau_p(x), x)/r_2 = \phi_2(\tau_p(x), x)/r_1.$$

Since Γ is an open subset of $\{x \in S\colon u_0 = u_2\}$ and $\phi(\tau_p(x(p)), x(p)) = \phi(T(p), x(p)) = x(p) \in \Gamma$, shrinking U_p if necessary, $U_p \cap \Gamma$ is a non-empty open subset of Γ and $\phi(\tau_p(x), x) \in \Gamma$ for all $x \in U_p \cap \Gamma$. Since the vector field F that defines (3.5) is C^∞, the same is true of ϕ (standard theory of dependence on initial data for ODEs), and the IFT implies the same is true of $x \mapsto \tau_p(x)$. By definition $T(p) = \inf\{t > 0\colon \phi(t, x(p)) \in \Gamma\}$ and, since L_p is a periodic orbit of (3.5) that intersects Γ at the unique point $x(p)$, it is again visually obvious that the same is true of $\tau_p(x)$ for $x \in U_p \cap \Gamma$, from which it follows that $\tau_p(x(y)) = T(y)$ for $x(y) \in U_p \cap \Gamma$. Thus, for any $p \in (0, p^*)$, since $y \mapsto x(y)$ and $x \mapsto \tau_p(x)$ are C^∞, the same is true of $y \mapsto T(y)$ for y in a neighbourhood of p. □

Lemma 3.3. *$\overline{\mu}, \overline{\sigma}^2$ are C^∞ on $(0, p^*)$.*

Proof. Combining (3.3) and (3.6), we can write μ_Y and σ_Y^2 in terms of $\Delta(i)$ and q_i as

$$\begin{aligned} \mu_Y(x) &= \frac{1}{2}\sum_{i=0}^{2} \Delta(i)^\top (D^2P)(x)\Delta(i)q_i(x) \quad \text{and} \\ \sigma_Y^2(x) &= \sum_{i=0}^{2}((\nabla P)(x)\Delta(i))^2 q_i(x). \end{aligned} \tag{3.11}$$

For each $i \in \{0, 1, 2\}$, $x \mapsto q_i(x) = r_i x_{i-1} x_i$ is C^∞ on S, so by (3.3), so are F and G. The function $P(x) = \prod_i x_i^{r_{i-1}}$ is C^∞ on the positive orthant $\{x \in \mathbb{R}^3\colon \min_i x_i > 0\}$, so in particular, ∇P and D^2P are C^∞ on $S^o = \{x \in S\colon P(x) > 0\}$. Combining with (3.11), μ_Y and σ_Y^2 are C^∞ on S^o.

Since F is C^∞ on S^o and the latter is forward invariant for (3.5), $(t, x) \mapsto \phi(t, x)$ is C^∞ on $\mathbb{R}_+ \times S^o$. Let x be a C^∞ function $p \mapsto x(p)$ on $[0, p^*]$, with $x(p) \in L_p$; to obtain x, draw a line segment from the boundary of S to the central equilibrium x^* and parametrize it by the value of P. Using $x(p)$ in (3.7), it follows that $\overline{\mu}, \overline{\sigma}^2$ are C^∞ on $(0, p^*)$ provided that $p \mapsto T(p)$ is smooth on $(0, p^*)$, which was shown in Lemma 3.2. □

3.4 Boundaries of the slow diffusion limit

We next discuss the behaviour of solutions of (3.9) at the boundaries 0 and p^*. First note that if $x_i^{(N)}(t) = 0$ then $x_i^{(N)}(s) = 0$ for $t > s$, which implies that 0 is absorbing for $Y^{(N)}$. To reflect this in the limit process, we impose the same for $Y(t)$: letting

$\tau = \inf\{t\colon Y(t) = 0\}$, we define $Y(t) = 0$ for all $t > \tau$. Let $\mathbb{P}_p$ denote the measure of the solution with initial value p, and $\mathbb{E}_p$ the corresponding expectation. Let $\tau(a) = \inf\{t\colon Y(t) = a\}$, $\tau(a,b) = \tau(a) \wedge \tau(b)$ and $\tau_e = \tau(0,p^*)$. Our main finding settles the boundary behaviour of solutions of (3.9).

Proposition 3.4. *For each $p \in [0,p^*)$, (3.9) has a pathwise unique strong solution $(Y(t))_{t\le\tau}$, where $\tau = \tau(0) = \inf\{t\colon Y(t) = 0\}$, satisfying $Y(0) = p$ and $Y(t) \in (0,p^*)$ for $t \in (0,\tau)$. Moreover, $p \mapsto \mathbb{E}_p[\,\tau\,]$ is bounded on $[0,p^*)$, and as a function of $Y(0)$, τ converges in distribution to an a.s. finite limit as $Y(0) \to p^*$.*

The main tool required for Proposition 3.4 is the following.

Lemma 3.5. *Let $Y(t)$ solve (3.9) with $\xi = p$ and $\tau(p), \mathbb{P}_p$ be as above. Then*

(i) 0 is accessible, i.e., if $p \in (0,p^*)$ then $\mathbb{P}_p(\tau(0) < \infty) > 0$, and

(ii) p^* is an entrance boundary, i.e., $\mathbb{P}_p(\tau(p^*) < \infty) = 0 \ \ \forall\, p \in (0,p^*)$ and

$$\lim_{t\to\infty} \inf_{p\in(y,p^*)} \mathbb{P}_p(\tau(y) \le t) > 0 \ \ \forall\, y \in (0,p^*). \tag{3.12}$$

First we show how this gives Proposition 3.4.

Proof of Proposition 3.4, using Lemma 3.5. The result is trivial if $p = 0$, as then $\tau = 0$ a.s. and $\mathbb{E}_0[\,\tau\,] = 0$. Let ψ be as in Proposition 3.1. The first part of (ii) in Lemma 3.5 implies that for each s, p, $\tau_e(s,p) = \inf\{t\colon \psi(s,t,p) = 0\}$ a.s. Combining with Proposition 3.1, for $p \in (0,p^*)$, $t \mapsto \psi(0,t,p)$ is the desired solution. It remains to do show $p \mapsto \mathbb{E}_p[\,\tau\,]$ is bounded on $(0,p^*]$ and τ has an a.s. finite limit as $Y(0) \to p^*$, beginning with the latter. Let

$$\tau(s,p,y) = \inf\{t\colon \psi(s,t,p) = y\}.$$

Then $p \mapsto \tau(s,p,y)$ is non-decreasing on $[y,p^*)$. To see this, for $s \in \mathbb{R}_+$ and $p_1, p_2 \in (0,p^*)$ define the collision time $\tau_c(s,p_1,p_2) = \inf\{t\colon \psi(s,t,p_1) = \psi(s,t,p_2)\}$ of $\psi(s,\cdot,p_1)$ with $\psi(s,\cdot,p_2)$. Applying the semigroup property (3.10) to each trajectory at time $\tau_c(s,p_1,p_2)$ shows that $\psi(s,t,p_1) = \psi(s,t,p_2)$ for all $t \ge \tau_c(s,p_1,p_2)$. In particular, if $p_1 \le p_2$ then $\psi(s,t,p_1) \le \psi(s,t,p_2)$ for all $t \ge s$, from which it follows that $p \mapsto \tau(s,p,y)$ is non-decreasing on $[y,p^*)$.

Let $\tau(s,p^*,y) = \lim_{p\to p^*} \tau(s,p,y)$. To see that the latter is a.s. finite, first note that the limit in (3.12), when positive, is equal to 1; this is explained below (21) in Chapter 20 of [13]. Since $p \mapsto \tau(s,p,y)$ is non-decreasing, $\{\tau(s,p^*,y) \le t\}$ is the decreasing limit of $\{\tau(s,p,y) \le t\}$ as $p \to p^*$, so $\mathbb{P}(\tau(s,p^*,y) \le t) = \inf_{p\in(y,p^*)} \mathbb{P}(\tau(s,p,y) \le t)$ and $\mathbb{P}(\tau(s,p^*,y) < \infty) = \lim_{t\to\infty} \inf_{p\in(y,p^*)} \mathbb{P}(\tau(s,p,y) \le t) = 1$ from the stronger (3.12).

It remains to show that $\mathbb{E}_p[\,\tau\,]$ is bounded on $[0,p^*)$. Fix $y \in (0,p^*)$ and let t,δ be such that $\mathbb{P}_y(\tau(0) \le t) \ge \delta$ and $\mathbb{P}_p(\tau(y) \le t) \ge \delta$ for $p \in [y,p^*)$. By the Markov property, $\mathbb{P}_p(\tau(y) \le 2t) \ge \delta^2$ for $p \in [y,p^*)$. Since $p \mapsto \tau(s,p,0)$ is non-decreasing, $\mathbb{P}_p(\tau(0) \le t) \ge \delta$ for $p < y$, so in particular $\mathbb{P}_p(\tau(0) \le 2t) \ge \delta^2$ for $p \in [0,p^*)$. By the Markov property $\mathbb{P}_p(\tau(0) \ge 2kt) \le (1-\delta^2)^k$ for $p \in [0,p^*)$ and integer $k \ge 1$ so in particular $\mathbb{E}_p[\,\tau(0)\,]$ is bounded on $[0,p^*)$. $\square$

Next we prove Lemma 3.5, using the boundary point classification from [14]. First we check that the process $(Y(t))$ is a regular diffusion on the interval $I = (0,p^*)$ with infinitesimal mean $\overline{\mu}$ and variance $\overline{\sigma}^2$, i.e., that it satisfies assumptions 1-4 of Ch. 15, Sec. 3 on pg. 191. Assumption 1 is just that the range of Y is an interval. Assumption 2 is regularity, i.e., $\mathbb{P}(\tau(y) < \infty \mid Y(0) = x) > 0$ for x, y in I, that we address in a moment. Assumption 3 is that

$$\begin{aligned} \overline{\mu}(x) &= \lim_{h\to 0^+} \mathbb{E}[Y(h) - Y(0) \mid Y(0) = x] \quad \text{and} \\ \overline{\sigma}^2(x) &= \lim_{h\to 0^+} \mathbb{E}[(Y(h) - Y(0))^2 \mid Y(0) = x] \end{aligned} \tag{3.13}$$

which follows by taking the appropriate limits in (3.9). Assumption 4 is that $\overline{\mu}, \overline{\sigma}^2$ are continuous and $\overline{\sigma}^2$ is positive on $(0, p^*)$. Continuity follows from Lemma 3.3. For positivity, by (3.7), $\overline{\sigma}^2$ is the average value of σ_Y^2 on L_p so it suffices that σ_Y^2 is positive on $\bigcup_{p \in (0,p^*)} L_p = S^o \setminus \{x^*\}$. If $x \in S^o$ then $q_i(x) > 0$ for all i, and using (3.15) below, since $P(x) > 0$, $\nabla P(x)\Delta(i) = 0$ for all i iff $r_{i-1}/x_i = r_{i-2}/x_{i-1}$ for all i, i.e., if $x = x^*$. The claim then follows from the formula (3.11) for σ_Y^2.

Regularity of Y is "obviously" implied by positivity of $\overline{\sigma}^2$, but for some reason I could not find a reference for this, so I'll show it here using the random time-change representation of one-dimensional diffusions, and also introduce a function that we'll need later. By assumption on $\overline{\mu}, \overline{\sigma}^2$ the scale function S ((6.3) in Chapter 15 of [14]) is defined on $(0, p^*)$ by

$$S(x) = \int_c^x s(y)dy \quad \text{where} \quad s(x) = \exp\left(-\int_c^y (2\overline{\mu}(z)/\overline{\sigma}^2(z))dz\right) dy \tag{3.14}$$

and $c \in (0, p^*)$ is arbitrary but fixed. Clearly S is continuous and since $s > 0$ on $(0, p^*)$, S is increasing, so S extends to $[0, p^*]$ by taking the appropriate limits, provided we allow it to take the values $\pm\infty$, and then $S^{-1} : (S(0), S(p^*))$ is defined, continuous and increasing, so regularity holds for Y iff it holds for $Z := S(Y)$. As outlined below (1) in Chapter 20 of [13], using Itô's formula Z takes values in $(S(0), S(p^*))$ and is on natural scale, i.e., is a martingale, and has diffusion coefficient $\tilde{\sigma}^2$ where $\tilde{\sigma} = \overline{\sigma} \circ S^{-1}$. Since $\overline{\mu}, \overline{\sigma}^2$ are C^∞ and σ^2 is positive on $(0, p^*)$, $\tilde{\sigma}$ is C^∞ and positive on $(S(0), S(p^*))$. In particular, distributional uniqueness holds for the solutions of $dZ = \tilde{\sigma}(Z)dB$ on $(S(0), S(p^*))$, so Z is equal in distribution to the time-changed Brownian motion $t \mapsto B(T(t))$ with $B(0) := Z(0)$, defined for $t < \zeta := \inf\{t \colon B(T(t)) \notin (S(0), S(p^*))\}$, where $T(t) = \inf\{s \geq 0 \colon \int_{S(0)}^{S(p^*)} L^x(s)dx/\tilde{\sigma}^2(x) \geq t\}$ is continuous and increasing and L is the local time of B. To see why, show that $B(T(t))$ has coefficients 0 and $\tilde{\sigma}^2$ in the sense of (3.13), using $\mathbb{P}(|B(h) - B(0)| \geq \epsilon) \to 0$ as $h \to 0^+$ for fixed $\epsilon > 0$ together with $\int_{a_t}^{b_t} L^x(t)dx = t$ where $a_t = \inf_{s \leq t} B(s)$ and $b_t = \sup_{s \leq t} B(s)$, together with continuity of $\tilde{\sigma}^2$. Now, suppose $B(0) = z$ and $S(0) < a < z < S(p^*)$, and fix $b \in (z, S(p^*))$. We'll show that $\mathbb{P}(B(T(t)) = a$ for some $t < \zeta \mid B(0) = z) \geq (z - a)/(b - a) > 0$, establishing regularity. From basic properties of B, $T_{a,b} := \inf\{t \colon B(t) \in \{a, b\}\}$ is a.s. finite and $\mathbb{P}(B(T_{a,b}) = a) = (z - a)/(b - a)$. On the other hand, $\delta := \inf\{\tilde{\sigma}^2(x) \colon x \in [a, b]\} > 0$ so for $s \leq T_{a,b}$, $\int_a^b L^x(s)dx/\tilde{\sigma}^2(x) \leq \delta s$, so $T(t) \geq t/\delta$ on the event that $T_{a,b} > T(t)$. Together, this gives the desired estimate.

With the assumptions verified, we can now make use of the theory in Ch. 15, Sec. 6 of [14]. We'll need to do computations involving $\overline{\mu}$ and $\overline{\sigma}^2$ so we'll need formulae for the latter, which in turn requires formulae for μ_Y and σ_Y^2. We begin with σ_Y^2.

Since i indexes the possible interactions, use j, k to index vectors and matrices. For each j,

$$(\nabla P)_j(x) = \frac{\partial}{\partial x_j} P(x) = \frac{r_{j-1}}{x_j} P(x).$$

Since the non-zero entries of $\Delta(i)$ are $\Delta_{i-1}(i) = -1$ and $\Delta_i(i) = 1$,

$$\nabla P(x)\Delta(i) = \left(\frac{r_{i-1}}{x_i} - \frac{r_{i-2}}{x_{i-1}}\right) P(x). \tag{3.15}$$

Referring to (3.6),

$$\begin{aligned}\sigma_Y^2(x) &= \sum_i (\nabla P(x)\Delta(i))^2 q_i(x) \\ &= \sum_i \left(\frac{r_{i-1}}{x_i} - \frac{r_{i-2}}{x_{i-1}}\right)^2 r_i x_{i-1} x_i P(x)^2. \end{aligned} \tag{3.16}$$

Next we compute μ_Y. We have $(D^2P)_{jj}(x) = r_{j-1}(r_{j-1}-1)P(x)/x_j^2$ and for $j \neq k$,

$$(D^2P)_{jk}(x) = \frac{r_{j-1}r_{k-1}}{x_j x_k} P(x),$$

so

$$\begin{aligned}&\Delta(i)^\top (D^2P)(x)\Delta(i) \\ &= ((D^2P)_{i-1,i-1}(x) + (D^2P)_{i,i}(x) - 2(D^2P)_{i-1,i}(x))P(x) \\ &= \left(\frac{r_{i-2}(r_{i-2}-1)}{x_{i-1}^2} + \frac{r_{i-1}(r_{i-1}-1)}{x_i^2} - 2\frac{r_{i-2}r_{i-1}}{x_{i-1}x_i}\right) P(x). \end{aligned} \tag{3.17}$$

and since $q_i(x) = r_i x_{i-1} x_i$,

$$\begin{aligned}&\Delta(i)^\top (D^2P)(x)\Delta(i) q_i(x) \\ &= \left(r_{i-2}(r_{i-2}-1)\frac{x_i}{x_{i-1}} + r_{i-1}(r_{i-1}-1)\frac{x_{i-1}}{x_i} - 2r_0r_1r_2\right) P(x)\end{aligned} \tag{3.18}$$

Letting $\rho_i = r_i(r_i-1)$ and $r_\pi = \prod_i r_i$ for tidiness and summing over i,

$$\mu_Y(x) = \frac{1}{2}\sum_i \left(\rho_{i-2}\frac{x_i}{x_{i-1}} + \rho_{i-1}\frac{x_{i-1}}{x_i}\right) P(x) - 3r_\pi P(x). \tag{3.19}$$

We can now prove Lemma 3.5.

Proof of Lemma 3.5. First note that since $r_i \leq 1$ for all i by assumption, $\mu_Y(x) \leq -3r_\pi P(x)$ for all $x \in S^o$, so $\overline{\mu}(p) \leq -3r_\pi p$ for $p \in (0, p^*]$.

First we check that 0 is accessible. With the scale function S from (3.14), by Lemma 6.2 in Ch. 15 of [14] 0 is accessible (they call it "attainable") if $S(0) > -\infty$ and $\int_0^p (S(y) - S(0))m(y)dy < \infty$ for some $p \in (0, p^*)$, where $m(y) := 1/(\overline{\sigma}^2(y)s(y))$ is the density of the speed measure and s is given by (3.14). By definition

$$S'(p) = s(p) = \exp\left(\int_p^c (2\overline{\mu}(z)/\overline{\sigma}^2(z))dz\right).$$

In particular $S'(p) \geq 0$ and since $\overline{\mu}(p) \leq 0$, $p \mapsto S'(p)$ is increasing, i.e., $p \mapsto S(p)$ is convex, on $(0, p^*)$. In particular, since $S'(c) = 1$, $S'(p) \leq 1$ for $p \leq c$ so $S(0) \geq S(c) - c > -\infty$. Since $y \mapsto S'(y)$ is increasing, $S(y) - S(0) \leq yS'(y) = ys(y)$ so

$$\int_0^p (S(y) - S(0))m(y)dy = \int_0^p \frac{S(y) - S(0)}{s(y)\overline{\sigma}^2(y)}dy \leq \int_0^p \frac{y\,dy}{\overline{\sigma}^2(y)}.$$

It remains to show the integral on the right is finite. Since $\overline{\sigma}^2$ is continuous and positive on $(0, p^*)$ and $\overline{\sigma}^2(p) \geq \inf\{\sigma_Y^2(x) \colon x \in S, P(x) = p\}$ it suffices to find $c, \epsilon, \delta > 0$ such that $\sigma_Y^2(x) \geq cP(x)^{2-\epsilon}$ whenever $P(x) \leq \delta$. Recall $u_i := x_i/r_{i-1}$. Rewriting (3.16),

$$\sigma_Y^2(x) = \sum_i \frac{(u_{i-1} - u_i)^2}{u_{i-1}u_i} P(x)^2 r_\pi. \tag{3.20}$$

Let $u_M = \max_i u_i$ and $u_m = \min_i u_i$. Since $\sum_i x_i = 1$, $\sum_i r_{i-1}u_i = 1$ and with $r_M := \max_i r_i$, $\sum_i u_i \geq 1/r_M$ so $u_M \geq 1/3r_M$. By definition of P, $u_m \leq u_M/2$ if $P(x)$ is small enough, then using both observations, for some $c, \delta > 0$, $\sigma_Y^2(x) \geq cP(x)^2/u_m$ when $P(x) \leq \delta$. Let $r_s = \sum_i r_i$. Since $P(x) = \prod_i x_i^{r_{i-1}} \geq (r_m u_m)^{r_s}$, for some $C > 0$, $u_m \leq CP(x)^{1/r_s}$. Combining with the above $\sigma_Y^2(x) \geq (c/C)P(x)^{2-1/r_s}$ when $P(x) \leq \delta$, as desired.

Next we check that p^* is an entrance boundary. By Lemma 6.2 in Ch. 15 of [14], if $S(p^*) = \infty$ (which we show in a moment) then p^* is inaccessible, i.e., $\mathbb{P}_p(\tau(p^*) < \infty) = 0$ for $p \in (0, p^*)$, which is half of the definition of entrance boundary. The other half, as in Lemma 3.5, is that

$$\lim_{t\to\infty} \inf_{p\in(y,p^*)} \mathbb{P}_p(\tau(y) \leq t) > 0 \tag{3.21}$$

for $y \in (0, p^*)$. To show this, we will make use of the drift of Y. Indeed, since $\overline{\mu}(p) \leq -3r_\pi p$, the process W defined by $W(t) := Y(t \wedge \tau(y, p^*)) + (t \wedge \tau(y, p^*))3r_\pi y$ is a non-negative supermartingale with $W(0) \leq p^*$. Since p^* is inaccessible, $\tau(y) = \tau(y, p^*)$ so if $p \in (y, p^*)$ and $t < \tau(y)$ then $Y(t) > y$ and $t < \tau(y, p^*)$, so

$$\mathbb{P}_p(\tau(y) > t) \leq \mathbb{P}_p(W(t) > y + 3r_\pi yt) \leq \frac{p^*}{y + 3r_\pi t}.$$

Since this tends to 0 uniformly over $p \in (y, p^*)$ as $t \to \infty$ we have proved (3.21). It remains to show that $S(p^*) = \infty$. Given $c \in (0, p^*)$, $\overline{\mu}(y) \leq -3r_\pi c$ for $y \in [c, p^*)$, and since $\nabla P(x^*)\Delta(i) = 0$ for each i, $P(x) = O(|x^* - x|^2)$ as $x \to x^*$ and so $\overline{\sigma}^2(p) = O((p^* - p)^2)$ as $p \to p^*$, so for some $c_1 > 0$, letting $c_2 = \exp(-c_1/(p^* - c))$,

$$\begin{aligned} S(p^*) &\geq \int_c^{p^*} \exp\left(c_1 \int_c^y dz/(p^* - z)^2\right) dy \\ &= \int_c^{p^*} c_2 \exp\left(c_1/(p^* - y)\right) dy. \end{aligned}$$

To see that the integral diverges, note the integrand is at least $c_2 e^{c_1 k}$ for $y \geq p^* - 1/k$ and with k_0 large enough that $p^* - 1/k_0 \geq c$, bound the integral below by the divergent sum

$$\sum_{k\geq k_0} \frac{1}{k(k+1)} c_2 e^{c_1 k}. \qquad \square$$

4 Convergence of fixation time

In this section we prove Theorem 1.1 part (i), using Theorem 4.1 below which establishes convergence of $Y^{(N)}$ to the slow diffusion limit Y. Recall that $\tau^{(N)} = \inf\{t \colon Y^{(N)}(t) = 0\}$ is the hitting time of ∂S on the slow time scale and that $\xrightarrow{\text{d}}$ denotes convergence in distribution. For convergence of processes we use the Skorokhod topology on càdlàg functions.

Theorem 4.1. *Suppose* $P(x^{(N)}(0)) \to p$ *as* $N \to \infty$ *for some* $p \in [0, p^*]$ *and for* $p < p^*$ *let* $\tau = \inf\{t \colon Y(t) = 0\}$ *where* Y *solves* (3.9) *with* $Y(0) = p$*, letting* τ *be the distributional limit given by Proposition* (3.4) *if* $p = p^*$*. Then* $\tau^{(N)} \xrightarrow{\text{d}} \tau$.

Note that Theorem 4.1 with $p = p^*$ is just Theorem 1.1 part (i) with $\tau^{(N)}$ in place of the rescaled fixation time τ_N^o/N. So, to prove Theorem 1.1 part (i) it remains to show that the time to fixation after hitting the boundary of S is $o(N)$; in fact it's $O(\log N)$.

Proof of Theorem 1.1 assuming Theorem 4.1. Recall the fixation time $\tau_N^o = \inf\{t \colon x^{(N)} \in \{e_i\}\}$ where e_i are standard basis vectors in $\mathbb{R}^3$. It suffices to show that $(\tau_N^o - \tau_N)/N \xrightarrow{\text{p}} 0$.

This is implied by $\mathbb{E}[\tau_N^o - \tau_N] = O(\log N)$ which we show now; note that $P(x) = 0$ iff $x_i = 0$ for some i so $\tau_N = \inf\{t\colon x_i^{(N)}(t) = 0 \text{ for some } i\}$.

If $x_i = 0$ then $x_{i+2} = 1 - x_{i+1}$ and $x_{i+2} \to x_{i+2} + 1$ at rate $Nr_{i+2}x_{i+1}x_{i+2} = Nr_{i+2}(1 - x_{i+2})x_{i+2}$. If in addition $X_{i+2} = 0$ then $x = e_{i+1}$, otherwise x_{i+2} has state space $\{0, 1/N, \dots, 1\}$ and transition rates $q(x, x + 1/N) = Nr_{i+2}x(1 - x)$, so from any distribution of $x_{i+2}(0)$ on $\{1/N, \dots, 1\}$,

$$\mathbb{E}[\,\inf\{t\colon x_{i+2}(t) = 1\}\,] \le \sum_{k=1}^{N-1} \frac{1}{Nr_{i+2}(k/N)(1 - k/N)} = \frac{2}{r_{i+2}} \sum_{k=1}^{N-1} \frac{1}{k} = O(\log N)$$

where we used the formulae

$$\frac{1}{x(1-x)} = \frac{1}{x} + \frac{1}{1-x} \quad \text{and} \quad \sum_{k=1}^{N-1} \frac{1}{1 - k/N} = \sum_{k=1}^{N-1} \frac{1}{k/N}. \qquad \square$$

For what follows let

$$\begin{aligned}
&\tau_N^-(a) = \inf\{t\colon P(x^{(N)}(t)) \le a\}, \\
&\tau_N^+(b) = \inf\{t\colon P(x^{(N)}(t)) \ge b\}, \\
&\tau_N(a,b) = \tau_N^-(a) \wedge \tau_N^+(b) \ \text{ and} \\
&\tau_N^e(\epsilon) = \tau_N(\epsilon, p^* - \epsilon)
\end{aligned}$$

and let

$$\tau_-^{(N)}(a) = \tau_N(a^-)/N = \inf\{t\colon Y^{(N)}(t) \le a\},$$

analogously for $\tau_+^{(N)}(b)$ and $\tau^{(N)}(a,b)$, and for $\epsilon \in (0, p^*/2)$ let $\tau_e^{(N)}(\epsilon) = \tau^{(N)}(\epsilon, p^* - \epsilon)$ and for Y solving (3.9) recall $\tau(a) = \inf\{t\colon Y(t) = a\}$ and let $\tau(a,b) = \tau(a) \wedge \tau(b)$ and $\tau_e(\epsilon) = \inf\{t\colon Y(t) \notin (\epsilon, p^* - \epsilon)\}$. Let $\tau_N = \tau_N^-(0)$, similarly for $\tau^{(N)}$ and τ.

The proof strategy for Theorem 4.1 mirrors the development of the previous section: Proposition 4.2 demonstrates convergence on the interior of $(0, p^*)$, then Proposition 4.3 establishes that p^* and 0 are, asymptotically, entrance and accessible boundaries for $Y^{(N)}$, respectively.

Proposition 4.2. *Suppose $P(x^{(N)}(0)) \to p$ as $N \to \infty$ for some $p \in (0, p^*)$ and Y is as in Theorem 4.1. For every $\epsilon \in (0, p^*/2)$, as $N \to \infty$*

$$(Y^{(N)}(\cdot \wedge \tau_e^{(N)}(\epsilon)), \tau_e^{(N)}(\epsilon)) \xrightarrow{\mathrm{d}} (Y(\cdot \wedge \tau_e(\epsilon)), \tau_e(\epsilon)).$$

Proposition 4.3. *For each $\epsilon > 0$ there is $\delta > 0$ so that*

(i) if $Y^{(N)}(0) \ge p^* - \delta$ then $\limsup_{N\to\infty} \mathbb{P}(\tau_-^{(N)}(p^* - \delta) > \epsilon) \le \epsilon$ and
(ii) if $Y^{(N)}(0) \le \delta$ then $\limsup_{N\to\infty} \mathbb{P}(\tau^{(N)} > \epsilon) \le \epsilon$.

Proof of Theorem 4.1 assuming Propositions 4.2 and 4.3. We distinguish two cases according to $p = \lim_{N\to\infty} P(x^{(N)}(0))$, namely, $p < p^*$ and $p = p^*$, beginning with the first case.
Case 1: $p < p^*$. Since, by Proposition 4.2, $\tau_e^{(N)}(\epsilon) \xrightarrow{\mathrm{d}} \tau_e(\epsilon)$ for every $\epsilon \in (0, p^*/2)$ it's enough to show that $\tau - \tau_e(\epsilon) \xrightarrow{\mathrm{p}} 0$ as $\epsilon \to 0$ and that $\tau^{(N)} - \tau_e^{(N)}(\epsilon) \xrightarrow{\mathrm{p}} 0$ if $N \to \infty$ slowly enough as $\epsilon \to 0$.

By Proposition 3.4, $\tau < \infty$ a.s. and $Y(t) < p^*$ for $t \le \tau$, so $\sup_{t\le\tau} Y(t) < p^*$ a.s. by continuity of $t \mapsto Y(t)$. In particular, $\mathbb{P}_p(\tau(\epsilon) = \tau_e(\epsilon)) \to 1$ as $\epsilon \to 0$. Since $t \mapsto Y(t)$ is a.s. continuous, $\tau - \tau(\epsilon) \xrightarrow{\mathrm{p}} 0$ as $\epsilon \to 0$. Combining, $\tau - \tau_e(\epsilon) \xrightarrow{\mathrm{p}} 0$. On the other hand, using $Y^{(N)}(\cdot \wedge \tau_e^{(N)}(\epsilon)) \xrightarrow{\mathrm{d}} Y(\cdot \wedge \tau_e(\epsilon))$ and $\mathbb{P}_p(\tau(\epsilon/2) = \tau_e(\epsilon/2)) \to 1$ as $\epsilon \to 0$, $\mathbb{P}(\tau_-^{(N)}(\epsilon) =$

$\tau_e^{(N)}(\epsilon)) \to 1$ if $N \to \infty$ slowly as $\epsilon \to 0$. By Proposition 4.3 (ii), $\tau^{(N)} - \tau_-^{(N)}(\epsilon) \xrightarrow{\mathrm{p}} 0$ if $N \to \infty$ slowly as $\epsilon \to 0$. Combining, $\tau^{(N)} - \tau_e^{(N)}(\epsilon) \xrightarrow{\mathrm{p}} 0$ as desired.

Case 2: $p = p^*$. $Y^{(N)}(\tau_-^{(N)}(p^* - \epsilon)) \to p^* - \epsilon$ as $N \to \infty$ since $\sup_{t \le \tau_-^{(N)}} |Y^{(N)}(t) - Y^{(N)}(t^-)| \xrightarrow{\mathrm{p}} 0$ as $N \to \infty$; to see this, note that $x^{(N)}$ jumps by $2/N$ and use the fact that P is Lipschitz on $\{x \in S\colon P(x) \ge p^* - 2\epsilon\}$. Observe from Proposition 4.3 (i) that $\tau_-^{(N)}(p^* - \epsilon) \xrightarrow{\mathrm{p}} 0$ if $N \to \infty$ slowly enough as $\epsilon \to 0$, and that for càdlàg f, $f(a + \cdot)$ converges to f in the Skorokhod topology as $a \to 0$. To obtain the result, apply Case 1 to $Y^{(N)}(\tau_-^{(N)}(p^* - \epsilon) + \cdot)$ and $Y(\cdot)$ with $Y(0) = p^* - \epsilon$, letting $N \to \infty$ slowly as $\epsilon \to 0$. □

Remark. To prove Theorem 4.1 we did not need to show that $Y^{(N)} \xrightarrow{\mathrm{d}} Y$, nor would doing so directly imply that $\tau^{(N)} \xrightarrow{\mathrm{d}} \tau$. Based on what we proved, to find that $Y^{(N)} \xrightarrow{\mathrm{d}} Y$ it would suffice to show that for each $\epsilon > 0$, once $Y, Y^{(N)}$ go below δ, with probability $\to 1$ as $\delta \to 0$ (and for $Y^{(N)}$ as $N \to \infty$ slowly enough as $\delta \to 0$) they remain below ϵ until they hit 0.

4.1 Stochastic averaging and convergence on $(0, p^*)$

In this section we prove Proposition 4.2. To do so we'll use a standard limit theorem from the canonical reference [5], which roughly amounts to showing that the jump size of $Y^{(N)}$ tends to 0 and that its linear and quadratic martingales converge to those of the limiting diffusion; we begin with the former, which is not hard.

Lemma 4.4. *For each N, $\sup_t |Y^{(N)}(t) - Y^{(N)}(t^-)| \le (1/N)^{r_m}$ where $r_m = \min_i r_i$.*

Proof. Let $S^{(N)} = \{x \in S\colon Nx \in \mathbb{Z}^3\}$. Jumps in $Y^{(N)}$ are bounded by

$$\sup_{x \in S^{(N)},\ i \in \{0,1,2\}} |P(x + \Delta(i)/N) - P(x)|.$$

If $x \in S$ and $P(x) > 0$, then from (3.17), since $\max_i r_i \le 1$ by assumption,

$$\Delta(i)^\top (D^2 P)(x) \Delta(i) \le 0$$

for each i, i.e., P is concave on S. It follows that for $x \in S^{(N)}$ and k such that $x + k\Delta(i)/N, x + (k+1)\Delta(i)/N \in S^{(N)}$, $k \mapsto P(x + (k+1)\Delta(i)/N) - P(x + k\Delta(i)/N)$ is monotone, so if x maximizes $|P(x + \Delta(i)/N) - P(x)|$ over i and $x, x + \Delta(i)/N \in S^{(N)}$ then either x or $x + \Delta(i)/N$ is on the boundary of S, i.e., $x_j = 0$ or $(x + \Delta(i)/N)_j = 0$ for some j. Suppose $x_j = 0$; the proof is similar in the other case. Then $P(x) = 0$, so

$$|P(x + \Delta(i)/N) - P(x)| = \prod_\ell (x_\ell + \Delta_\ell(i)/N)^{r_{\ell-1}} \le (1/N)^{r_m},$$

since $x_\ell \in [0,1]$ for each i, $x_j + \Delta_j(i)/N \le 1/N$ and $N \ge 1$. □

The generator A_N of $x^{(N)}$ is given for bounded measurable f by

$$(A_N f)(x) = \sum_{i=0}^{2} (f(x + \Delta(i)/N) - f(x)) N q_i(x).$$

Defining the appropriate functions and using Dynkin's formula, then speeding up time by a factor of N,

$$M_N(t) = Y^{(N)}(t) - Y^{(N)}(0) - \int_0^t N\mu_Y^{(N)}(x^{(N)}(Ns))ds \quad \text{and}$$
$$Q_N(t) = M_N(t)^2 - \int_0^t N(\sigma_Y^{(N)})^2(x^{(N)}(Ns))ds \tag{4.1}$$

are martingales, where

$$\mu_Y^{(N)}(x) = \sum_{i=0}^{2}(P(x+\Delta(i)/N) - P(x))Nq_i(x) \text{ and}$$
$$(\sigma_Y^{(N)})^2(x) = \sum_{i=0}^{2}(P(x+\Delta(i)/N) - P(x))^2 Nq_i(x). \tag{4.2}$$

We can think of $N\mu_Y^{(N)}$ and $N(\sigma_Y^{(N)})^2$ as the drift and diffusion coefficients of $Y^{(N)}$. The following result asserts that, away from the boundary of S, they converge uniformly to the drift and diffusion coefficients of Y.

Lemma 4.5. *Let μ_Y, σ_Y^2 as in* (3.11) *and let $S(\epsilon) = \{x \in S \colon P(x) \geq \epsilon\}$. For each $\epsilon > 0$,*

$$\sup_{x\in S(\epsilon)} |N\mu_Y^{(N)}(x) - \mu_Y(x)| + |N(\sigma_Y^{(N)})^2(x) - \sigma_Y^2(x)| = O(1/N).$$

Proof. We begin with $\mu_Y^{(N)}$. Since $P, \nabla P$ and $D^2 P$ are bounded on $S(\epsilon)$, a Taylor expansion implies that

$$\sup_{x\in S(\epsilon)} |P(x+\Delta(i)/N) - P(x)$$
$$- (\nabla P)(x)\Delta(i)/N - \Delta(i)^\top D^2 P(x)\Delta(i)/2N^2| = O(1/N^3).$$

As noted earlier, since P is constant on solutions of (3.5), $(\nabla P)(x)F(x) = 0$ for $x \in S$, which implies that

$$\sum_{i=0}^{2}(\nabla P)(x)(\Delta(i)/N)Nq_i(x) = 0.$$

Using the Taylor expansion in $\mu_Y^{(N)}$ and cancelling the above terms we find that

$$\sup_{x\in S(\epsilon)} |N\mu_Y^{(N)}(x) - \mu_Y(x)| = O(1/N).$$

Let $a = P(x+\Delta(i)/N) - P(x)$ and $b = (\nabla P)(x)\Delta(i)/N$. Then

$$|N(\sigma_Y^{(N)})^2(x) - \sigma_Y^2(x)| \leq \sum_{i=0}^{2} |a^2 - b^2| N^2 q_i(x).$$

Since ∇P is bounded on $S(\epsilon)$, $b = O(1/N)$, and a Taylor approximation gives $a = O(1/N)$ and $|a-b| = O(1/N^2)$, so $|a^2 - b^2| = |a-b|\cdot|a+b| = O(1/N^3)$. Since $q_i = O(1)$,

$$\sup_{x\in S(\epsilon)} |N(\sigma_Y^{(N)})^2(x) - \sigma_Y^2(x)| = O(1/N). \qquad \square$$

The following lemma gives the required stochastic averaging property.

Lemma 4.6. *Given $\epsilon > 0$, let $S_e(\epsilon) = \{x \in S \colon P(x) \in [\epsilon, p^* - \epsilon]\}$. Suppose $f : S_e(\epsilon) \to \mathbb{R}$ is continuous. Let $\overline{f} :\in [\epsilon, p^*] \to \mathbb{R}$ denote the orbital average of f, defined for $p \in [\epsilon, p^* - \epsilon]$ by*

$$\overline{f}(p) = \frac{1}{T(p)}\int_0^{T(p)} f(\phi(t, x_p))dt$$

where $x_p \in L_p$ is arbitrary. Then for $T > 0$,

$$\sup_{t\leq T}\left|\frac{1}{N}\int_0^{Nt\wedge\tau_N^e(\epsilon)} (f(x^{(N)}(s)) - \overline{f}(P(x^{(N)}(s)))ds\right| \xrightarrow{\text{p}} 0 \text{ as } N \to \infty.$$

Lemma 4.6 takes a bit of effort so first we'll use it to prove Proposition 4.2.

Proof of Proposition 4.2 assuming Lemma 4.6. To prove the result we use Theorem 4.1 in Chapter 7 of [5], which is a general diffusion limit theorem that requires us to verify some conditions, listed as (4.1)-(4.7) in the reference. Since $\overline{\mu}, \overline{\sigma}^2$ are not globally Lipschitz, we will localize and work with somewhat more regular processes. It's enough to show that for $\epsilon \in (0, p^*/2)$,

$$(Y_\epsilon^{(N)}, \tau_e^{(N)}(\epsilon)) \xrightarrow{\mathrm{d}} (Y_\epsilon, \tau_e(\epsilon))$$

where, for each N, $Y_\epsilon^{(N)}$ is a process coupled to $Y^{(N)}$ such that $Y_\epsilon^{(N)}(t) = Y_\epsilon(t)$ for $t \le \tau_e^{(N)}(\epsilon)$ and similarly for Y_ϵ and Y.

As in the proof of Proposition 3.1, let $\mu_\epsilon, \sigma_\epsilon^2$ be Lipschitz extensions of $\overline{\mu}, \overline{\sigma}^2$, respectively, from $(\epsilon, p^* - \epsilon)$ to $\mathbb{R}$ and let Y_ϵ solve (3.9) with $\mu_\epsilon, \sigma_\epsilon^2$ in place of $\overline{\mu}, \overline{\sigma}^2$. As explained in Chapter 5, Section 3 of [5], since $\mu_\epsilon, \sigma_\epsilon^2$ are Lipschitz on $\mathbb{R}$, the martingale problem for the diffusion with coefficients $\mu_\epsilon, \sigma_\epsilon^2$ is well-posed – in particular, the solution $t \mapsto Y_\epsilon(t)$ exists for all $t \in \mathbb{R}_+$ and is pathwise unique. Coupling the solutions by using the same Brownian motion to construct them, $Y_\epsilon(t) = Y(t)$ for $t \le \tau_e(\epsilon)$. Let ψ_ϵ be the solution flow of Proposition 3.1 with $\mu_\epsilon, \sigma_\epsilon^2$ in place of $\overline{\mu}, \overline{\sigma}^2$ and let

$$Y_\epsilon^{(N)}(t) = \begin{cases} Y^{(N)}(t) & \text{for } \ t \le \tau_e^{(N)}(\epsilon), \\ \psi_\epsilon(\tau_e^{(N)}(\epsilon), t, Y^{(N)}(\tau_e^{(N)}(\epsilon))) & \text{for } \ t > \tau_e^{(N)}(\epsilon). \end{cases}$$

In order for $Y_\epsilon^{(N)} \xrightarrow{\mathrm{d}} Y_\epsilon$ it's enough that conditions (4.1)-(4.7) of the limit theorem are satisfied by $Y^{(N)}(t)$ for $t \le \tau_e^{(N)}(\epsilon)$, since by modifying the processes A_n, B_n and M_n described below in the obvious way, i.e., taking for $t > \tau_e^{(N)}(\epsilon)$

$$A_n(\tau_e^{(N)}(\epsilon)) + \int_{\tau_e^{(N)}(\epsilon)}^t \sigma_\epsilon^2(Y_\epsilon^{(N)}(s))ds$$

in place of $A_n(t)$ and analogously for B_n and M_n, the conditions are trivially satisfied for $t > \tau_e^{(N)}(\epsilon)$. With respect to their notation with a, b and X_n, A_n, B_n, M_n, we have $a = \overline{\sigma}^2$, $b = \overline{\mu}$, $X_n = Y^{(n)}$,

$$A_n(t) = \int_0^t (\sigma_Y^{(N)})^2(x^{(N)}(Ns))ds, \quad B_n(t) = \int_0^t \mu_Y^{(N)}(x^{(N)}(Ns))ds,$$

and M_n as in (4.1).

(i) Conditions (4.1)-(4.2) are implied by (4.1) above.
(ii) Condition (4.3): this follows from Lemma 4.4.
(iii) Conditions (4.4)-(4.5) are trivial, as A_n, B_n are continuous.
(iv) Recall $\tau_e^{(N)}(\epsilon) = N\tau_N^e(\epsilon)$, where both are defined at the top of this section. Recalling that $Y^{(N)}(t) = P(x^{(N)}(Nt))$, condition (4.6) is equivalent to the following estimate: for $t > 0$,

$$\frac{1}{N}\int_0^{Nt \wedge \tau_N^e(\epsilon)} |\mu_Y^{(N)}(x^{(N)}(Ns)) - \overline{\mu}(P(x^{(N)}(Ns)))|ds \xrightarrow{\mathrm{p}} 0 \ \text{ as } \ N \to \infty.$$

Using Lemma 4.5, we can replace $\mu_Y^{(N)}$ with μ_Y. Then, noting $\overline{\mu}$ is by definition the orbital average of μ_Y and using Lemma 4.6, the estimate follows. Condition (4.7) is analogous, so we omit it.

Lastly it remains to show that $\tau_e^{(N)}(\epsilon) \xrightarrow{\mathrm{d}} \tau_e(\epsilon)$. Define the map $T : D(\mathbb{R}_+) \to \mathbb{R}$ on càdlàg functions $X : \mathbb{R}_+ \to \mathbb{R}$ by $T(X) := \inf\{t \colon X_t \le 0\}$, so that $\tau_e^{(N)}(\epsilon) = T(Y^{(N)} - \epsilon) \wedge$

$T(p^* - \epsilon - Y^{(N)})$ and similarly for $\tau_e(\epsilon)$. Since σ_ϵ^2 can be taken continuous and positive on $(0, p^*)$, for fixed $\epsilon \in (0, p^*/2)$, continuity of sample paths and Blumenthal's 0-1 law imply that $X \mapsto T(X - \epsilon)$ and $X \mapsto T(p^* - \epsilon - X)$ are continuous on a set $A \subset D(\mathbb{R}_+)$ that has $\mathbb{P}(Y_\epsilon \in A) = 1$. Convergence of $Y_\epsilon^{(N)}$ to Y_ϵ and the continuous mapping theorem then imply the desired result. □

It remains to prove Lemma 4.6. To do so we require the following uniform approximation by solutions of (3.5). Let $S^{(N)} = \{x \in S \colon Nx \in \mathbb{Z}^3\}$.

Lemma 4.7. *Recall* $\tau_N^e(\epsilon) = \inf\{t \colon P(x^{(N)}(t)) \notin (\epsilon, p^* - \epsilon)\}$*. Let* ϕ *denote the solution flow of* (3.5)*. For each* $T > 0$ *there is* $\gamma > 0$ *so that if* $\epsilon \in [0, 1]$ *and* $x \in S^{(N)}$ *then*

$$\mathbb{P}_x(\sup_{t \le T} |x^{(N)}(t) - \phi(t, x)| > \epsilon) \le 6e^{-\gamma N \epsilon^2}.$$

Proof. From Lemma 3.3 in [6] we have the following general estimate: if X is an $\mathbb{R}$-valued semimartingale with $|\Delta X| \le c$, $a > 0$ and λ is such that $0 < \lambda c \le 1/2$, then

$$\mathbb{P}(\sup_{t \ge 0} |X_t^m| - \lambda \langle X \rangle_t \ge a) \le 2e^{-\lambda a}, \tag{4.3}$$

where X^m is the martingale part and $\langle X \rangle$ is the predictable quadratic variation (pqv). For each i the stopped coordinate process $x_i^{(N)}(\cdot \wedge \tau_N^e(\epsilon))$ has jumps bounded by $c = 1/N$ and has respective martingale part and pqv

$$(I_i x^{(N)})(t \wedge \tau_N^e(\epsilon)) \quad \text{and} \quad \frac{1}{N}(J_i x^{(N)})(t \wedge \tau_N^e(\epsilon))$$

where I_i, J_i are defined for measurable $x : \mathbb{R}_+ \to S$ by

$$\begin{aligned} (I_i x)(t) &= x(t) - x(0) - \int_0^t F_i(x(s))ds \quad \text{and} \\ (J_i x)(t) &= \int_0^t G_{ii}(x(s))ds. \end{aligned}$$

By (3.3) and since $x \mapsto q_i(x)$ are quadratic in x, F, G are bounded and Lipschitz on S. Letting C_1 be a Lipschitz constant, for $T > 0$ and measurable $x : \mathbb{R}_+ \to S$, an application of Gronwall's inequality shows that

$$\sup_{t \in [0,T]} \sum_i |x_i(t) - \phi_i(t, x(0))| \le e^{C_1 t} \sup_{t \in [0,T]} \sum_i |(I_i x_i)(t)|. \tag{4.4}$$

Let C_2 be a bound on F, G on S. Since G_{ii} is non-negative, if $x : \mathbb{R}_+ \to S$ then

$$\sup_{t \in [0,T]} (J_i x)(t) = (J_i x)(T) \le C_2 T,$$

Applying (4.3) to each process $x_i^{(N)}$, if $\lambda c \le 1/2$, i.e., $\lambda \le N/2$, then

$$\mathbb{P}(\sup_{t \le T} \sum_i |(I_i x^{(N)})(t)| \ge 3a + 3\lambda C_2 T/N) \le 6e^{-\lambda a}.$$

Let $a = \epsilon\, e^{-C_1 T}/6$ and $\lambda = \epsilon\, e^{-C_1 T} N/6C_2 T$. If $C_2 \ge e^{-C_1 t}/3T$, which can be chosen as such, and $\epsilon \in [0, 1]$ then $\lambda \le N/2$. Combining the above estimate with (4.4) and taking $\gamma = e^{-2C_1 T}/36 C_2 T$ then gives the result. □

Proof of Lemma 4.6. Fix $\epsilon, T > 0$ and define the sequence (s_i) by $s_0 = 0$ and $s_{i+1} = (s_i + T(p_i)) \wedge NT \wedge \tau_N^e(\epsilon)$, where $p_i = P(x^{(N)}(s_i))$ (no relation to the fixation probability $p_i^{(N)}$ defined in the introduction). For $t \le T$ let $I(t) = \inf\{i\colon s_i \ge Nt \wedge \tau_N^e(\epsilon)\}$. To set up the estimate we'll use the fact that for $i \le I(T) - 2$,

$$\overline{f}(p_i)T(p_i) = \int_{s_i}^{s_{i+1}} f(\phi(t - s_i, x^{(N)}(s_i))ds = \int_0^t \overline{f}(P(x^{(N)}(s_i)))ds.$$

Then,

$$\int_0^{Nt\wedge\tau_N^e(\epsilon)} f(x^{(N)}(s))ds = \sum_{i=0}^{I(t)-2} \overline{f}(p_i)T(p_i) + \sum_{i=0}^{I(t)-2} R_i^{(1)} + R^{(2)}(t)$$

where $R^{(2)}(t) = \int_{s_{I(t)-1}}^{Nt\wedge\tau_N^e(\epsilon)} f(x^{(N)}(s))ds$ and

$$R_i^{(1)} = \int_{s_i}^{s_{i+1}} (f(x^{(N)}(s)) - f(\phi(t - s_i, x^{(N)}(s_i)))ds.$$

Similarly,

$$\sum_{i=0}^{I(t)-2} \overline{f}(p_i)T(p_i) = \int_0^{Nt\wedge\tau_N^e(\epsilon)} \overline{f}(P(x^{(N)}(s))ds + \sum_{i=0}^{I(t)-2} R_i^{(3)} + R^{(4)}(t)$$

where $R^{(4)}(t) = \int_{s_{I(t)-1}}^{Nt\wedge\tau_N^e(\epsilon)} \overline{f}(P(x^{(N)}(s)))ds$ and

$$R_i^{(3)} = \int_{s_i}^{s_{i+1}} (\overline{f}(P(x^{(N)}(s))) - \overline{f}(P(x^{(N)}(s_i)))ds.$$

In particular,

$$\int_0^{Nt\wedge\tau_N^e(\epsilon)} (f(x^{(N)}(s)) - \overline{f}(P(x^{(N)}(s)))ds = \sum_{i=0}^{I(t)-2} (R_i^{(1)} + R_i^{(3)}) + R^{(2)}(t) + R^{(4)}(t).$$

By assumption, f is continuous on the compact set $S_e(\epsilon)$ so $C_1 := \sup_{x\in S_e(\epsilon)} |f(x)| < \infty$. Using Lemma 3.2, $p \mapsto T(p)$ is continuous so similarly $C_2 := \sup_{p\in[\epsilon,p^*-\epsilon]} T(p) < \infty$. By definition $s_{i+1} - s_i \le C_2$ so $\sup_{t\le T} R^{(i)}(t) \le C_1C_2$ for $i \in \{2, 4\}$. Thus, the desired result follows if we can show that for each $T > 0$,

$$\sup_{t\le T} \frac{1}{N} \sum_{i=0}^{I(t)-2} (R_i^{(1)} + R_i^{(3)}) \xrightarrow{\text{p}} 0 \ \text{ as } \ N \to \infty,$$

which, in turn, follows if we can show that for each $T > 0$,

$$\frac{1}{N} \sum_{i=0}^{I(T)-2} \left(|R_i^{(1)}| + |R_i^{(3)}|\right) \xrightarrow{\text{p}} 0 \ \text{ as } \ N \to \infty. \tag{4.5}$$

For $\delta > 0$ let

$$E_i(\delta) = \{\sup_{s_i\le t<s_{i+1}} |x^{(N)}(t) - \phi(t - s_i, x^{(N)}(s_i))| > \delta\}.$$

Since f is continuous on the compact set $S_e(\epsilon)$, f is uniformly continuous on $S_e(\epsilon)$. Let m be a modulus of continuity for f and given $c > 0$ let $\delta > 0$ small enough that $m(\delta) \le c$. Then, on the complement of $E_i(\delta)$,

$$|R_i^{(1)}| \le (s_{i+1} - s_i)\,\epsilon \le c\,T(p_i).$$

Using just the bound on f, $R_i^{(1)} \le 2C_1 T(p_i)$, so in general,

$$|R_i^{(1)}| \le c\,T(p_i) + 2C_1 T(p_i)\,\mathbf{1}(E_i(\delta)).$$

For $R_i^{(3)}$ let C_3 be a Lipschitz constant for P on $S_e(\epsilon)$. A similar argument as in the proof of Lemma 3.3 shows that $\overline{f}$ is continuous on $[\epsilon, p^* - \epsilon]$. Let $\overline{m}$ be a modulus of continuity for $\overline{f}$ and $\delta > 0$ small enough that $\overline{m}(C_3\delta) \le c$. Note that $\sup_{p \in [\epsilon, p^*-\epsilon]} \overline{f}(p) \le C_1$. Using the fact that $P(\phi(t,x))$ does not depend on t and arguing in the same way as for $R_i^{(1)}$,

$$|R_i^{(3)}| \le c\,T(p_i) + 2C_1 T(p_i)\,\mathbf{1}(E_i(\delta)).$$

It's clear that $T(p) > 0$ for $p \in (0, p^*)$. Since T is continuous, $c_1 := \inf_{p \in [\epsilon, p^*-\epsilon]} T(p) > 0$. By definition, if $i \le I(T) - 2$ then $p_i \in (\epsilon, p^* - \epsilon)$ and $s_{i+1} - s_i = T(p_i) \ge c_1$. In particular, for each $T > 0$, $NT \ge s_{I(T)-1} \ge c_1(I(T) - 1)$ so $I(T) = O(N)$. By Lemma 4.7 with δ in place of ϵ, there is $\gamma > 0$ so that for each i, $\mathbb{P}(E_i) \le 6e^{-\gamma N \delta^2}$, so $\mathbb{P}(\sum_{i=0}^{I(t)-2} \mathbf{1}(E_i) \ne 0) = O(Ne^{-\gamma N\delta^2}) \to 0$ as $N \to \infty$. It follows that with probability tending to 1 as $N \to \infty$,

$$\sum_{i=0}^{I(T)-2} \left(|R_i^{(1)}| + |R_i^{(3)}|\right) \le 2c \sum_{i=0}^{I(T)-2} T(p_i) = 2c\, s_{I(T)-1} \le 2cNT.$$

Since T is fixed and c is arbitrary, (4.5) follows. □

4.2 Entrance at p^*

Proof of Proposition 4.3, part (i). We use the same approach as in the proof of Lemma 3.5. Using Lemma 4.5 and the estimate $\mu_Y(x) \le -3r_\pi P(x)$, for large N and $p \in (y, p^*)$, $N\mu_Y^{(N)}(x) \le -2r_\pi y$. Fix $c > 0$ and $\delta > 0$ and let $y = p^* - \delta$. If $Y^{(N)}(0) \in (y, p^*)$ and N is large enough that $Y^{(N)}$ jumps by at most c on (y, p^*) then from (4.1) it follows that the process $W^{(N)}$ defined by

$$W^{(N)}(t) := Y^{(N)}(t \wedge \tau_-^{(N)}(y)) - (y + c) + (t \wedge \tau_-^{(N)}(y))2r_\pi y$$

is a non-negative supermartingale with $W^{(N)}(0) \le \delta$. So, if $t < \tau_-^{(N)}(y)$ then $Y^{(N)}(t) > y$ and $t < \tau(y, p^*)$, so

$$\mathbb{P}(\tau_-^{(N)}(y) > t) \le \mathbb{P}_p(W(t) > 2r_\pi yt - c) \le \frac{\delta}{2r_\pi t - c}.$$

Let $t = \epsilon$, then let $c = r_\pi \epsilon$ and $\delta = r_\pi \epsilon^2$ to obtain the desired result. □

4.3 Accessibility at 0

Proof of Proposition 4.3, part (ii). Recall $\mu_Y^{(N)}$ and $(\sigma_Y^{(N)})^2$ from (4.2) that we think of as drift and diffusion coefficients of $Y^{(N)}$. We'll show that $Y^{(N)}$ has non-positive drift and that, if $\sum_i r_i$ is small enough, its diffusion coefficient is bounded below by a positive constant, at which point we can use optional stopping arguments to obtain the result.

First, we show that $Y^{(N)}$ itself has non-positive drift for each N, i.e., $\mu_Y^{(N)}(x) \le 0$ for $x \in S$ where $\mu_Y^{(N)}$ is given in (4.2). This is clear if $P(x) = 0$ since 0 is absorbing for $Y^{(N)}$: if $P(x) = 0$ and $P(x + \Delta(i)/N) > 0$ then $q_i(x) = 0$. As noted in the proof of Lemma 4.4, P is concave on S. In particular, $P(x + \Delta(i)/N) - P(x) \le \nabla P(x)\Delta(i)/N$ for all $x \in S^o$ (note that $\nabla P(x)$ is undefined if $x_i = 0$ and $r_i < 1$). From (4.2) and using $\nabla PF = 0$,

$$\mu_Y^{(N)}(x) \le \sum_i \nabla P(x)\Delta(i) q_i(x) = \nabla P(x) F(x) = 0.$$

Next, recall from (3.15) that

$$\nabla P(x)\Delta(i) = \left(\frac{r_{i-1}}{x_i} - \frac{r_{i-2}}{x_{i-1}}\right) P(x).$$

Suppose $x \in S_o^{(N)} := \{x \in S_o \colon Nx \in \mathbb{Z}^3\}$. If $P(x+\Delta(i)/N) \le P(x)$ then by concavity $P(x) - P(x+\Delta(i)/N) \ge \nabla P(x)\Delta(i)/N$. Otherwise, $P(x+\Delta(i)/N) - P(x) \ge \nabla P(x+\Delta(i)/N)\Delta(i)/N$. Since $x \in S_o^{(N)}$, $(x+\Delta(i)/N)_i \le 2x_i$ for each i. Using this, (3.15) and $P(x+\Delta(i)/N) \ge P(x)$,

$$\nabla P(x+\Delta(i)/N)\Delta(i)/N \ge \left(\frac{r_{i-1}}{2x_i} - \frac{r_{i-2}}{x_{i-1}}\right) P(x).$$

In either case, if $x \in S_o^{(N)}$ then

$$|P(x+\Delta(i)/N) - P(x)| \ge \left(\frac{r_{i-1}}{2x_i} - \frac{r_{i-2}}{x_{i-1}}\right) P(x).$$

It follows that

$$(\sigma_Y^{(N)})^2(x) \ge \sum_i \frac{(u_{i-1} - 2u_i)^2}{2u_{i-1}u_i} P(x)^2 r_\pi. \tag{4.6}$$

Recall $r_s = \sum_i r_i$. Arguing as we did below (3.20), from (4.6) we find $c, \delta > 0$ such that $(\sigma_Y^{(N)})^2(x) \ge cP(x)^{2-1/r_s}$ when $P(x) \le \delta$. Rescaling time in (3.2) by a constant factor, we may assume that $r_s = 1/2$, and the rescaling does not affect the conclusion of the lemma. With this assumption, $(\sigma_Y^{(N)})^2(x) \ge c$ when $P(x) \le \delta$.

With $M^{(N)}$ as in (4.1) define $W^{(N)}$ by

$$W^{(N)}(t) = \begin{cases} M^{(N)}(t) & t \le \tau^{(N)} \\ W^{(N)}(\tau^{(N)}) + cB(t - \tau^{(N)}) & t > \tau^{(N)} \end{cases}$$

where B is a standard Brownian motion independent of $Y^{(N)}$. Since $\mu_Y^{(N)} \le 0$, $W^{(N)}$ is a martingale and $W^{(N)}(t) \le Y^{(N)}(t) - Y^{(N)}(0)$ for $t \le \tau^{(N)}$, and in particular, if $W^{(N)}(t) \le -Y^{(N)}(0)$ then $t > \tau^{(N)}$. By construction and the lower bound on $(\sigma_Y^{(N)})^2$, $Q^{(N)}(t) := ct - W^{(N)}(t)^2$ is a supermartingale with $Q^{(N)}(0) = 0$. Fix $C > 0$ and let $\tau_*^{(N)} = \inf\{t \colon W^{(N)}(t) \notin (-\delta, C\delta)\}$. If $Y^{(N)}(0) \le \delta$ and $W^{(N)}(\tau_*^{(N)}) \le -\delta$ then $\tau^{(N)} \le \tau_*^{(N)}$.

By definition, $W^{(N)}(t) - W^{(N)}(t^-) = Y^{(N)}(t) - Y^{(N)}(t^-)$ for $t \le \tau^{(N)}$ and $t \mapsto W^{(N)}(t)$ is a.s. continuous for $t > \tau^{(N)}$ so by Lemma 4.4, jumps in $W^{(N)}$ are bounded by $(1/N)^{r_m}$. If $Y^{(N)}(0) \le \delta$ and $N \ge \delta^{1/r_m}$ then $W^{(N)}(\tau_*^{(N)}) \in [-2\delta, -\delta] \cup [C\delta, (C+1)\delta]$ and $-\delta \le (W^{(N)})(t) \le (C+1)\delta$ for $t \in [0, \tau_*^{(N)}]$. Since $\mathbb{E}[Q^{(N)}(t \wedge \tau_*^{(N)})] \le 0$ for each t, using the definition of $Q^{(N)}$ and Fatou's lemma implies $\mathbb{E}[\tau_*^{(N)}] < \infty$.

If $W^{(N)}(0) \in [-\delta, 0]$ then applying optional sampling to $W^{(N)}$, $\mathbb{P}(W^{(N)}(\tau_*^{(N)}) \le -\delta) \ge C/(C+2)$. Applying optional sampling to the supermartingale $Q^{(N)}$,

$$\mathbb{E}[\tau_*^{(N)}] = c^{-1}\mathbb{E}[W^{(N)}(\tau_*^{(N)})^2 + Q^{(N)}(\tau_*^{(N)})] = c^{-1}\mathbb{E}[W^{(N)}(\tau_*^{(N)})^2] \le (C+1)^2\delta^2/c.$$

For $\epsilon > 0$, if $Y^{(N)}(0) \le \delta$ then using Markov's inequality on the second term,

$$\begin{aligned} \mathbb{P}(\tau^{(N)} > \epsilon) &\le \mathbb{P}(\tau^{(N)} \ne \tau_*^{(N)}) + \mathbb{P}(\tau_*^{(N)} > \epsilon) \\ &= \mathbb{P}(W^{(N)}(\tau_*^{(N)}) \ge C\delta) + \mathbb{P}(\tau_*^{(N)} > \epsilon) \\ &\le 2/(C+2) + (C+1)^2\delta^2/(c\epsilon). \end{aligned}$$

Given $\epsilon > 0$, the desired result follows if $C + 2 \ge 4/\epsilon$ and $\delta^2 \le c\epsilon^2/(2(C+1)^2)$. ☐

Appendix

Proof of Proposition 3.1. We begin by solving, for each $\epsilon > 0$, the equation (3.9) with modified coefficients $\mu_\epsilon, \sigma_\epsilon$ that are bounded, Lipschitz, and agree with $\overline{\mu}, \overline{\sigma}$ on the interval $[\epsilon, p^* - \epsilon]$, denoting the solution $\psi_\epsilon(s,t,p)$, with domain $D := \{(s,t,p) \in \mathbb{R}_+^2 \times \mathbb{R} \colon s \le t\}$. We prove a weak semigroup property for ψ_ϵ:

$$\begin{gathered}\text{for } 0 \le s \le t \le u \ \text{ and } \ \mathcal{F}(s) \text{ measurable } \xi, \\ \psi_\epsilon(t,u,\psi_\epsilon(s,t,\xi)) = \psi_\epsilon(s,u,\xi) \ \text{ almost surely.}\end{gathered} \tag{4.7}$$

We extract a continuous version of ψ_ϵ, which has the semigroup property. We show that $(\psi_{1/k})_{k>0}$ are consistent and from this deduce that $\lim_{k\to\infty} \psi_{1/k}$ exists, denoting it ψ. We show that ψ solves (3.9) and inherits continuity and the semigroup property from $(\psi_{1/k})$, completing the proof.

First we obtain ψ_ϵ. Lemma 3.2 below implies that $\overline{\mu}, \overline{\sigma}$ are Lipschitz on compact subsets of $(0,p^*)$. Fix $0 < \epsilon < p^*/2$ and let $\mu_\epsilon, \sigma_\epsilon$ be a Lipschitz extension of $\overline{\mu}, \overline{\sigma}$ from $(\epsilon, p^* - \epsilon)$ to $\mathbb{R}$; for example, let $\mu_\epsilon = \overline{\mu}$ and $\sigma_\epsilon = \overline{\sigma}$ on $[\epsilon, p^* - \epsilon]$, set them equal to 0 outside $(0,p^*)$, and interpolate linearly on $(0,\epsilon)$ and $(p^* - \epsilon, p^*)$. For each $s \in \mathbb{R}_+^2$, Theorem 18.3 in [13] then gives a random function $(t,p) \mapsto \psi_\epsilon(s,t,p)$, defined for $t \ge s$ and $p \in \mathbb{R}$, such that for fixed (s,p) and $\mathcal{F}(s)$ measurable ξ, $t \mapsto \psi_\epsilon(s,t,\xi)$ is the a.s. unique solution to (3.9) with $\mu_\epsilon, \sigma_\epsilon$ in place of $\overline{\mu}, \overline{\sigma}$, which we'll notate in shorthand as (3.9)ϵ.

To prove the weak semigroup property (4.7) for ψ_ϵ, let $\eta = \psi_\epsilon(s,t,\xi)$. Then η is $\mathcal{F}(t)$-measurable so $u \mapsto \psi_\epsilon(t,u,\eta)$ is the a.s. unique solution to (3.9)ϵ for $u \ge t$, with t,u,η in place of s,t,ξ. Using $\eta = \psi_\epsilon(s,t,\xi)$ and (3.9)ϵ for $\psi_\epsilon(s,t,\xi)$ in the corresponding (3.9)ϵ for $\psi_\epsilon(t,u,\eta)$ shows that $u \mapsto \psi_\epsilon(t,u,\eta)$ solves (3.9)ϵ with u in place of t, and a dummy variable other than u, for $u \ge t$; (4.7) then follows from uniqueness.

Theorem 18.3 in [13] already implies that for fixed s, $(t,p) \mapsto \psi_\epsilon(s,t,p)$ is a.s. continuous. With a bit more work we obtain a version continuous in all three variables. The required estimate is given by Theorem 2.23 in [13]: for some $a,b,C > 0$ and all $(s_1,t_1,p_1),(s_2,t_2,p_2)$ in $\mathbb{R}^3$ such that $s_1 \le t_1$ and $s_2 \le t_2$,

$$\mathbb{E}[\,|\psi_\epsilon(s_1,t_1,p_1) - \psi_\epsilon(s_2,t_2,p_2)|^a\,] \le C(|s_1 - s_2|^{3+b} + |t_1 - t_2|^{3+b} + |p_1 - p_2|^{3+d}).$$

In the statement of Theorem 2.23 the domain is assumed to be $\mathbb{R}^d$ for some d, however the proof is unaffected if the domain is D. We'll prove the above estimate with $a = 2q$ and $3 + b = q$ for arbitrary $q > 3$, which suffices. Since $|x_1 + x_2 + x_3|^{2q} \le (3\max_i |x_i|)^{2q} \le 3^{2q}(|x_1|^{2q} + |x_2|^{2q} + |x_3|^{2q})$, it suffices to show that

(i) $\mathbb{E}[\,|\psi_\epsilon(s_1,t,p) - \psi_\epsilon(s_2,t,p)|^{2q}\,] \le C|s_1 - s_2|^q$ for $0 \le s_1 \le s_2 \le t$ and $p \in \mathbb{R}$,
(ii) $\mathbb{E}[\,|\psi_\epsilon(s,t_1,p) - \psi_\epsilon(s,t_2,p)|^{2q}\,] \le C|t_1 - t_2|^q$ for $0 \le s \le t_1 \le t_2$ and $p \in \mathbb{R}$ and
(iii) $\mathbb{E}[\,|\psi_\epsilon(s,t,p_1) - \psi_\epsilon(s,t,p_2)|^{2q}\,] \le C|p_1 - p_2|^q$ for $0 \le s \le t$ and $p_1,p_2 \in \mathbb{R}$.

As noted in the proof of Theorem 2.23 in [13], it suffices to establish the estimates on compact subsets of D. So, we'll fix $T > 0$ and assume $0 \le s \le t \le T$ and $|p| \le T$. For statement (iii), use the the display equation below (12) in the proof of Theorem 18.3 in [13]: there is a positive, non-decreasing process $(c(t))$ such that for $0 \le s \le t$ and $p_1, p_2 \in \mathbb{R}$,

$$\mathbb{E}[\,|\psi_\epsilon(s,t,p_1) - \psi_\epsilon(s,t,p_2)|^{2q}\,] \le 2|p_1 - p_2|^{2q}\exp((t-s)c(t-s)). \tag{4.8}$$

If $|p_1|, |p_2| \le T$ then $|p_1 - p_2|^{2q} \le |p_1 - p_2|^q (2T)^q$, so letting $C = 2\exp(Tc(T))(2T)^q$ gives (iii). Statements (i)-(ii) will follow from the weak semigroup property and the following short-time estimate: for some $C > 0$ and any $(s,t,p) \in \mathbb{R}_+^2 \times \mathbb{R}$ such that $0 \le t - s \le T$,

$$\mathbb{E}[\,|\psi_\epsilon(s,t,p) - p|^{2q}\,] \le C|t-s|^q. \tag{4.9}$$

First we use (4.9) to obtain (i)-(ii) above, then we prove it. We begin with (ii). For fixed $s \le t_1 \le t_2$, by the semigroup property,

$$\psi_\epsilon(s, t_2, p) = \psi(t_1, t_2, \psi(s, t_1, p)) \text{ a.s.}$$

Then, (ii) is obtained by combining this with the short-time estimate, using $t_1, t_2, \psi(s, t_1, p)$ in place of s, t, p in the latter. Next we show (i). Use the weak semigroup property to write

$$\psi_\epsilon(s_1, t, p) = \psi(s_2, t, \psi(s_1, s_2, p)) \text{ a.s.}$$

Combine this with (4.8), with $s_2, t, \psi_\epsilon(s_1, s_2, p)$ in place of s, t, p, to find that with $C = 2\exp(Tc(T))$, for $0 \le s_1 \le s_2 \le T$ and any $p \in \mathbb{R}$,

$$\mathbb{E}[\,|\psi_\epsilon(s_1, t, p) - \psi_\epsilon(s_2, t, p)|^{2q}\,] \le C|\psi_\epsilon(s_1, s_2, p) - p|^{2q}$$

Combine with the short-time estimate to obtain (i). We now prove the short-time estimate (4.9). By construction, $\mu_\epsilon, \sigma_\epsilon$ are continuous and compactly supported, so bounded. So, if $|t - s| \le T$ then for some $C > 0$

$$\left|\int_s^t \mu_\epsilon(\psi_\epsilon(s, u, p))du\right|^{2q} \le (C(t-s))^{2q} \le CT^q(t-s)^q$$

and the quadratic variation of $\int_s^t \sigma_\epsilon(\psi_\epsilon(s, u, p))dB(u)$ is at most $C(t - s)$. Using the BDG inequality (see for example Proposition 15.7 in [13]), for some $c_q > 0$,

$$\mathbb{E}\left|\int_s^t \sigma_\epsilon(\psi_\epsilon(s, u, p))dB(u)\right|^{2q} \le c_q C^q (t-s)^q.$$

Since $|x_1 + x_2|^{2q} \le 2^{2q}(|x_1|^{2q} + |x_2|^{2q})$, using (3.9)$\epsilon$ and the above gives the short-time estimate. This completes the proof that ψ_ϵ has a continuous version. Let $D_1 = \{(s, t, u, p) \in \mathbb{R}^4 : 0 \le s \le t \le u\}$. The weak semigroup property (with $\xi = p$) holds simultaneously for all $(s, t, u, p) \in D_1 \cap \mathbb{Q}^4$ since the latter is countable. Since $D_1 \cap \mathbb{Q}^4$ is dense in D_1, by continuity

$$\begin{aligned}&\text{almost surely, } \forall 0 \le s \le t \le u,\ \forall p \in \mathbb{R},\\ &\psi_\epsilon(t, u, \psi_\epsilon(s, t, p)) = \psi_\epsilon(s, u, p).\end{aligned} \tag{4.10}$$

Discarding a null set we may assume that $\psi_{1/k}$ is continuous, and satisfies (4.10), for every integer $k > 0$. Since $\mu_\epsilon, \sigma_\epsilon$ agrees with $\overline{\mu}, \overline{\sigma}$ on $(\epsilon, p^* - \epsilon)$, for each $(s, p) \in \mathbb{R}_+ \times (0, p^*)$, $t \mapsto \psi_\epsilon(s, t, p)$ is the a.s. unique solution to (3.9) on $s \le t \le \tau_\epsilon(s, p) := \inf\{t \colon \psi_\epsilon(s, t, p) \notin (\epsilon, p^* - \epsilon)\}$. In particular, a.s. for every $j < k \in \mathbb{Z}_{>0}$, $\psi_{1/k}(s, t, p) = \psi_{1/j}(s, t, p)$ for all $s \le t \le \tau_{1/j}(s, t, p)$, and $j \mapsto \tau_{1/j}(s, t, p)$ is non-decreasing. Let $\tau(s, p) = \lim_{j\to\infty} \tau_{1/j}(s, t, p)$ and for $t < \tau(s, p)$ let $\psi(s, t, p) = \lim_{k\to\infty} \psi_{1/k}(s, t, p)$. Then ψ is a.s. continuous on $\{(s, t, p) \in \mathbb{R}_+^2 \times (0, p^*) \colon 0 \le s \le t < \tau(s, p)\}$, $\tau(s, p) > 0$ for $(s, p) \in \mathbb{R}_+ \times (0, p^*)$ and $\tau = \tau_e$, where τ_e is given in the proposition statement, and for each $(s, p) \in \mathbb{R}_+ \times (0, p^*)$, $t \mapsto \psi(s, t, p)$ is the a.s. unique solution to (3.9) on $s \le t \le \tau_e(s, p)$. The semigroup property is inherited, with the restriction $u \le \tau_e(s, p)$, from $(\psi_{1/k})_{k>0}$. □

References

[1] Benaïm, M., Hofbauer, J. and Sandholm, W.H. (2008) Robust permanence and impermanence for stochastic replicator dynamics. *J. Biol. Dyn.*, 180–195. https://doi.org/10.1080/17513750801915269 MR2427526

[2] Berr, M., Reichenbach, T., Schottenloher, M. and Frey, E. (2009) Zero-One Survival Behavior of Cyclically Competing Species *PRL* 102, 048102

[3] Diekmann, O. and van Gils, S. A. (2009) On the cyclic replicator equation and the dynamics of semelparous populations. *SIAM J Appl Dyn Sys*, 8, 1160–1189 MR2551259

[4] Dobrinevski, A. and Frey, E. (2012) Extinction in neutrally stable stochastic Lotka-Volterra models. *Phys Rev E* 85, 051903.

[5] Ethier, S.N., and Kurtz, T.G. (1986) *Markov Processes: Characterization and Convergence.* John Wiley and Sons, New York. MR0838085

[6] Foxall, E. (2018) The naming game on the complete graph *Elec. J. Prob.* 23, no. 126. MR3896863

[7] Foxall, E. (2021) Extinction time of the logistic process. *J Appl Prob, 58(3).* MR4313023

[8] Foxall, E. (2021) Boundary dynamics of the replicator equations for neutral models of cyclic dominance. *DCDS-B*, 26(2), 1061–1082. MR4219121

[9] Frean, M. and Abraham, E.R. (2001) Rock-scissors-paper and the survival of the weakest. *Proc Royal Soc B* 268(1474) 1323–1327

[10] Hening, A. and Li, Yao (2021). Stationary distributions of persistent ecological systems. *J. Math. Biol.* 82:64 MR4266165

[11] Hening, A., Nguyen, D. and Chesson, P. (2021) A general theory of coexistence and extinction for stochastic ecological communities. *J. Math. Biol.* 82(6) MR4257839

[12] Hening, A., Nguyen, D. and Schreiber, S.J. (2021) A classification of the dynamics of three-dimensional stochastic ecological systems. *To appear in Ann Appl Probab.* MR4414698

[13] Kallenberg, O. (1997) *Foundations of modern probability.* Springer. MR1464694

[14] Karlin, S. and Taylor, H.M. (1981) *A second course in stochastic processes.* Academic Press. MR0611513

[15] Kurtz, T.G. (1978) Strong approximation theorems for density dependent Markov chains. *Stoch Proc Appl*, 6(3) 223–240. MR0464414

[16] Mobilia, M. and Rucklidge, A.M. (2014) Cyclic dominance in evolutionary games: a review. *JR Soc Interface.* https://doi.org/10.1098/rsif.2014.0735

[17] Reichenbach, T., Mobilia, M., and Frey, E. (2006) Coexistence versus extinction in the stochastic cyclic Lotka-Volterra model *Phys. Rev. E* 74, 051907 MR2293733

[18] Schreiber, S.J., Benaïm, M. and Atchadé, K.A.S. (2011) Persistence in fluctuating environments. *J. Math. Biol.* 62, 655–683. https://doi.org/10.1007/s00285-010-0349-5 MR2786721

[19] Van Kampen, N.G. (2007) *Stochastic Processes in Physics and Chemistry,* Third edition, Elsevier MR0648937

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS[1])
- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm