QUASI-STATIONARY DISTRIBUTIONS FOR RANDOMLY PERTURBED DYNAMICAL SYSTEMS

BY MATHIEU FAURE AND SEBASTIAN J. SCHREIBER[1]

Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS and University of California, Davis

We analyze quasi-stationary distributions $\{\mu^\varepsilon\}_{\varepsilon>0}$ of a family of Markov chains $\{X^\varepsilon\}_{\varepsilon>0}$ that are random perturbations of a bounded, continuous map $F: M \to M$, where M is a closed subset of $\mathbb{R}^k$. Consistent with many models in biology, these Markov chains have a closed absorbing set $M_0 \subset M$ such that $F(M_0) = M_0$ and $F(M \setminus M_0) = M \setminus M_0$. Under some large deviations assumptions on the random perturbations, we show that, if there exists a positive attractor for F (i.e., an attractor for F in $M \setminus M_0$), then the weak* limit points of μ_ε are supported by the positive attractors of F. To illustrate the broad applicability of these results, we apply them to nonlinear branching process models of metapopulations, competing species, host-parasitoid interactions and evolutionary games.

1. Introduction. A fundamental issue in biology is what are the minimal conditions to ensure the long-term survivorship for all of the interacting components, whether they be viral particles, bio-chemicals, plants or animals. When these conditions are met the interacting populations are said to persist or coexist. Since the pioneering work of Lotka (1925) and Volterra (1926) on competitive and predator-prey interactions, Thompson (1924), Nicholson and Bailey (1935) on host-parasite interactions and Kermack and McKendrick (1927) on disease outbreaks, nonlinear difference and differential equations have been used to understand conditions for persistence of interacting populations. For these deterministic models, persistence is often equated with an attractor bounded away from the extinction states in which case persistence holds over an infinite time horizon [Schreiber (2006b)]. In reality, extinction in finite time is inevitable due to finite population sizes and mortality events occurring with positive probability. However, for systems with a large number of individuals or particles, these times to extinction may be sufficiently long so that the system remains in a "metastable state," bounded away from extinction for a long time. To provide a rigorous mathematical basis for this intuition, we develop a general theory for randomly perturbed dynamical systems with absorbing states.

Under the appropriate assumptions about the random perturbations, we show that the existence of a positive attractor (i.e., an attractor which is bounded away from extinction states) for the unperturbed system implies two things as the number of individuals or particles gets large. First, when they exist, quasi-stationary distributions concentrate on the positive attractors of the unperturbed system. Second, the expected time to extinction for systems starting according to this quasi-stationary distribution grows exponentially with the system size. In particular, we generalize earlier related work for one-dimensional randomly perturbed dynamical systems [Högnäs (1997), Klebaner, Lazar and Zeitouni (1998), Ramanan and Zeitouni (1999)] to higher dimensional systems by extending a general theory of randomly perturbed systems without absorbing states [Kifer (1988, 1989, 1990)] to a general theory of randomly perturbed systems with absorbing states.

For the unperturbed, deterministic dynamics, we consider a bounded continuous map $F : M \to M$, where M is a closed subset of $\mathbb{R}^d$. *A random perturbation of F is a family of homogeneous Markov chains* $\{X^\varepsilon\}_{\varepsilon > 0}$ on M, whose transition kernels

$$p^\varepsilon(x, \Gamma) = \mathbb{P}\big[X^\varepsilon_{n+1} \in \Gamma \mid X^\varepsilon_n = x\big], \qquad x \in M, \Gamma \text{ Borel subset of } M$$

enjoy the following standing hypothesis.

STANDING HYPOTHESIS 1.1.   *For any $\delta > 0$, $\lim_{\varepsilon \to 0} \beta_\delta(\varepsilon) = 0$ where*

$$\beta_\delta(\varepsilon) = \sup_{x \in M} p^\varepsilon\big(x, M \setminus N^\delta(F(x))\big)$$

and $N^\delta(A) := \{x \in M : \inf_{y \in A} \|x - y\| < \delta\}$ denotes the δ-neighborhood of A.

All standing hypotheses are assumed to hold throughout the paper. Standing Hypothesis 1.1 implies that $p^\varepsilon(x, \cdot)$ converges uniformly to the Dirac measure $\delta_{F(x)}$ at $F(x)$ for the vague convergence of measures, that is, for any continuous function $g : M \to \mathbb{R}$ with compact support,

$$\lim_{\varepsilon \to 0} \sup_{x \in M} \left| \int_M g(y) p^\varepsilon(x, dy) - g\big(F(x)\big) \right| = 0.$$

Consequently, for small $\varepsilon > 0$, the asymptotic behavior of the Markov chain $\{X^\varepsilon_n\}^\infty_{n=1}$ should be related to the dynamics of iterating the map F.

When the perturbed system admits an invariant measure (e.g., the Markov chains are irreducible), the correspondence between the deterministic dynamics and the perturbed counterpart was initially studied by Andronov, Pontryagin and Witt (1933), and more recently by Freidlin and Wentzell (1984), Ruelle (1981), Sinaĭ (1972) and Kifer (1988, 1989, 1990). Recall that a Borel probability measure μ_ε on M is called *a stationary distribution for p^ε* if

$$\int_M p^\varepsilon(x, \Gamma) \mu_\varepsilon(dx) = \mu_\varepsilon(\Gamma) \qquad \text{for any Borel set } \Gamma \subset M.$$

These invariant measures describe the long-term statistical behavior of the perturbed system. Let us assume, for a moment, that for all $\varepsilon > 0$, the Markov chain X^ε admits (at least) one invariant measure μ_ε. We call μ a *limiting measure* of the family of measure $\{\mu_\varepsilon\}_{\varepsilon>0}$ if μ is the weak* limit of some sequence $\{\mu_{\varepsilon_n}\}_{n=1}^\infty$, where ε_n decreases to zero. Natural questions about these limiting measures include: Are the limiting measures invariant for the deterministic dynamics? If so, what can be said about their support?

Kifer (1988, 1989, 1990) considered these questions under the assumption that the transition kernels p^ε satisfied the following large deviation assumption: there exists a continuous, nonnegative rate function ρ such that

$$(1) \qquad \lim_{\varepsilon \to 0} \varepsilon \log p^\varepsilon(x, U) = -\inf_{y \in U} \rho(x, y)$$

for any open set $U \subset M$ and uniformly in $x \in M$. Under suitable assumptions, Kifer proved that limiting measures are invariant for F [i.e., $\mu(\Gamma) = \mu(F^{-1}(\Gamma))$ for all Borel set Γ] and are supported by the attractors of the deterministic dynamics. In particular, Kifer's approach allowed him to derive some of Freidlin and Wentzell's results on the asymptotic behavior of invariant measures for diffusion processes X_t^ε generated by operators of the form $L^\varepsilon = \varepsilon L + b$ where L is a "good" second-order elliptic differential operator and b a vector field [Freidlin and Wentzell (1984), Chapter 6].

While Kifer's results are applicable to a wide class of stochastic models for the physical sciences, they are not applicable to many models in ecology, epidemiology, immunology and evolution. These stochastic models often have absorbing states $M_0 \subset M$ corresponding to the loss of one or more populations that satisfy the following standing hypothesis.

STANDING HYPOTHESIS 1.2.   *The state space M can be written $M = M_0 \cup M_1$, where*:

- *M_0 is a closed subset of M*;
- *M_0 and M_1 are positively F-invariant, that is, $F(M_0) \subseteq M_0$ and $F(M_1) \subseteq M_1$*;
- *the set M_0 is assumed to be absorbing for the random perturbations*

$$(2) \qquad p^\varepsilon(x, M_1) = 0 \qquad \textit{for all } \varepsilon > 0, x \in M_0.$$

For many of these biological models, the set M_0 of absorbing states is reached in finite time almost surely for any $\varepsilon > 0$. Despite this eventual absorption, the process $\{X_n^\varepsilon\}_{n=1}^\infty$ may spend an exceptionally long period of time in the set M_1 of transient states provided that $\varepsilon > 0$ is sufficiently small. In applications, this "metastable" behavior may correspond to long-term persistence of an endemic disease, coexistence of interacting species, or maintenance of a genetic polymorphism. One approach to examining metastable behavior are quasi-stationary distributions which are invariant distributions when the perturbed process is conditioned on nonabsorption. More specifically, we have the following.

DEFINITION 1.3.   A probability measure μ_ε on M_1 is a *quasi-stationary distribution* (QSD) for p^ε provided there exists $\lambda_\varepsilon \in (0, 1)$ such that

$$\int_{M_1} p^\varepsilon(x, \Gamma)\mu_\varepsilon(dx) = \lambda_\varepsilon \mu_\varepsilon(\Gamma) \qquad \text{for all Borel } \Gamma \subset M_1.$$

Equivalently, dropping the index ε, a QSD μ satisfies the identity

$$\mu(\Gamma) = \mathbb{P}_\mu(X_n \in \Gamma \mid X_n \in M_1) \qquad \forall n,$$

where $\mathbb{P}_\mu$ denotes the law of the Markov chain $\{X_n\}_{n=0}^\infty$, conditional to X_0 being distributed according to μ. Quasi-stationary distributions can sometimes be defined through the so-called *Yaglom limit*,

$$\mu(\Gamma) = \lim_{n \to +\infty} \mathbb{P}_x(X_n \in \Gamma \mid X_n \in M_1),$$

when the limit exists and is independent of the initial state $x \in M_1$. When $\mathbb{P}(X_1 \in \cdot) = \mu(\cdot)$, λ is the probability of not being absorbed in the next time step. The existence of QSDs has been studied extensively [Arjas, Nummelin and Tweedie (1980), Barbour (1976), Buckley and Pollett (2010), Chan (1998), Coolen-Schrijner and van Doorn (2006), Darroch and Seneta (1965), Ferrari et al. (1995), Gosselin (2001), Kijima (1992), Lasserre and Pearce (2001), Nummelin and Arjas (1976), Seneta and Vere-Jones (1966), Tweedie (1974)].

Högnäs (1997), Klebaner, Lazar and Zeitouni (1998), Ramanan and Zeitouni (1999) studied weak* limit points μ of QSDs μ_ε as $\varepsilon \to 0$ for maps of the interval, that is, $M = [0, 1]$ and $M_0 = \{0\}$. Under suitable assumptions, these authors have shown that if F admits an attractor in $(0, 1)$, then the limiting measure μ is F-invariant and concentrated on the attractors of F in $(0, 1)$. Moreover, $\lambda_\varepsilon \geq 1 - e^{-c/\varepsilon}$ for an appropriate constant $c > 0$. This final assertion implies that if the perturbed processes is initiated in the quasi-stationary state, then the expected time to absorption increases exponentially with exponent $1/\varepsilon$ as ε decreases to zero.

Here, we extend these types of results to higher dimensional systems where M is a subset of $\mathbb{R}^d$. The two main results of the paper are stated in Section 2. First, we state a general result that ensures that the QSDs concentrate on attractors of F restricted to M_1. This result requires conditions on the topological dynamics and the rate at which $\beta_\delta(\varepsilon)$ in Standing Hypothesis 1.1 goes to zero. Second, for many applications, the randomly perturbed Markov chains satisfy large deviation assumptions. We present a result that guarantees the conditions of the general theorem are satisfied. Proofs of these two results are presented in Sections 3 and 4, respectively. In Section 3, we also show how the main result of Klebaner, Lazar and Zeitouni (1998) can be derived from our general theorem. In Section 5, we apply our results to stochastic models of metapopulation dynamics, competing species, host-parasitoid interactions and evolutionary games. In Section 6, we conclude by verifying the large deviation assumptions for the examples in Section 5.

2. Statement of the main results. Let $\{X^\varepsilon\}_{\varepsilon>0}$ be a family of Markov chains on a closed set $M \subset \mathbb{R}^d$, which satisfies Standing Hypothesis 1.2. Since M is assumed to be a closed subset of $\mathbb{R}^d$, every topological concept must be understood in terms of the topology induced in M. In particular, in the following, a compact set K will always be a closed (in M) bounded set $K \subset M$.

We assume that, for each $\varepsilon > 0$, there exists at least one QSD μ_ε: there exists $1 > \lambda_\varepsilon > 0$ such that $\lambda_\varepsilon \mu_\varepsilon = Q^\varepsilon \mu_\varepsilon$ where Q^ε is the operator defined on the set of finite Borel measures on M_1 by

$$Q^\varepsilon(\mu)(\Gamma) = \int_{M_1} p^\varepsilon(x, \Gamma)\mu(dx) \qquad \text{for every Borel } \Gamma \subset M_1.$$

Our main results concern characterizing the support of weak* limit points μ of the μ_ε as $\varepsilon \downarrow 0$. Under suitable assumptions, we show that these weak* limit points are supported by attractors of the map F that lie in M_1; see Section 2.1 for a definition of an attractor. In Section 2.1, we describe sufficient conditions for this result with suitable assumptions about the topological dynamics of F and $\beta_\delta(\varepsilon)$ in Standing Hypothesis 1.1 goes to zero. In Section 2.2, we describe large deviation assumptions which satisfy the assumptions presented Section 2.1 and which are easier to verify for applications presented in Section 5.

2.1. *Absorption-preserving chain recurrence and convergence to attractors.* We begin by recalling a few definitions from dynamical system theory. Let F^n be the n-iterate of F. A set $B \subset M$ is *invariant* for F if $F(B) = B$. A compact set A is an *attractor for F* provided there exists an open neighborhood U of A such that $\bigcap_{n \geq 1} F^n(U) = A$ and, for any open set $V \supset A$, there exists $n(V)$ such that $F^n(U) \subset V$ for all $n \geq n(V)$.

The key notions needed for our main result is *absorption preserving pseudoorbits and chain recurrence* introduced in Jacobs and Schreiber (2006). These definitions generalize Conley's (1978) notion of pseudoorbits and chain recurrence. Given $\delta > 0$, a family of points $\xi = (\xi_0, \ldots, \xi_n) \in M^{n+1}$ is called an *absorption preserving δ-pseudoorbit joining x to y* (ap δ-pseudoorbit for short) provided that:

(a) $x = \xi_0, y = \xi_n$,
(b) $\xi_i \in M_0 \Rightarrow \xi_{i+1} \in M_0$ and
(c) $d(\xi_{i+1}, F(\xi_i)) < \delta, i = 0, \ldots, n - 1$.

One can think of ap δ-pseudoorbits as approximations of actual orbits of the dynamics of F with an error no greater than δ and that preserve the absorbing set M_0. For readers unfamiliar with these concepts, consider F to be the identity map on the interval $[0, 1]$. Then any two points on the interval are connected by δ-pseudoorbits, for any $\delta > 0$. However, as every point is a fixed point, none of the points are connected by interating the map F.

Given $x, y \in M$, we say that x *ap-chains to y* and write $x <_{\text{ap}} y$ if for any $\delta > 0$, there exists an ap δ-pseudoorbit joining x to y. Notice that no point in M_0

ap-chains to any point in M_1. If $x <_{\mathrm{ap}} y$ and $y <_{\mathrm{ap}} x$, we shall write $x \sim_{\mathrm{ap}} y$. If $x \sim_{\mathrm{ap}} x$, then x is an *ap-chain recurrent* point. The set $\mathcal{R}_{\mathrm{ap}}$ of ap-chain recurrent points is F-invariant. The relation $\sim_{\mathrm{ap}}$, restricted to this set defines an equivalence relation. The equivalent classes, $[x]_{\mathrm{ap}}$ with $x \in \mathcal{R}_{\mathrm{ap}}$, are called *ap-basic classes*. In Section 3.1, we prove various properties of these equivalence classes, for example, $\omega(x) \subset \mathcal{R}_{\mathrm{ap}}$ whenever $\omega(x) \subset M_0$ or $\omega(x) \subset M_1$.

Let $[x]_{\mathrm{ap}}$ and $[y]_{\mathrm{ap}}$ be two distinct ap-basic classes. We write $[x]_{\mathrm{ap}} <_{\mathrm{ap}} [y]_{\mathrm{ap}}$ if $x <_{\mathrm{ap}} y$. A maximal basic class $[x]_{\mathrm{ap}}$ (i.e., $[x]_{\mathrm{ap}} <_{\mathrm{ap}} [y]_{\mathrm{ap}}$ implies that $[x]_{\mathrm{ap}} = [y]_{\mathrm{ap}}$) is called an *ap-quasiattractor*. In general, an ap-quasiattractor need not be an attractor for F. A simple example is an increasing function $F : [0, 1] \to [0, 1]$ with $F(x) = x$ for $x = 1 - 1/n$ for all natural numbers n and $F(x) \neq x$ otherwise. If $M_0 = \{0\}$, then $x = 1$ is a quasi-attractor but not an attractor for F.

We need three hypotheses in order to state the first main result. The first hypothesis requires that there is a finite number of ap-basic classes including at least one ap-quasiattractor in M_1. This assumption is satisfied for many important classes of mappings, including gradient-like systems and Axiom A systems. When this hypothesis is satisfied, we prove in Section 3.2 that all the ap-quasiattractors are in fact attractors.

HYPOTHESIS 2.1.   *There exists only a finite number of ap-basic classes in* $M_1 : \{K_i\}_{i=1,\dots,v}$. *Moreover, we assume that they are closed sets and* $\{K_i\}_{i=1,\dots,\ell}$ *with* $\ell \geq 1$ *are the ap-quasiattractors and* $\{K_i\}_{i=\ell+1,\dots,v}$ *are the nonap-quasiattractors.*

Our second hypothesis ensures the time spent near nonap-quasiattractors is not too long relative the $\beta_\delta(\varepsilon)$ described in Standing Hypothesis 1.1. For any Borel set V, we define the first passage time $\tau_V^\varepsilon = \min\{n : X_n^\varepsilon \notin V\}$.

HYPOTHESIS 2.2.   *Given any* $\delta > 0$, *there exist neighborhoods* $V_i \subset N^\delta(K_i)$ *of* K_i *for* $\ell + 1 \leq i \leq v$ *and quantity* $\delta_1 \in (0, 1)$ *such that*

$$\sup_{x \in V_j} \mathbb{P}_x\big[\tau_{V_j}^\varepsilon > h(\varepsilon)\big] \leq \zeta(\varepsilon) \quad and \quad \lim_{\varepsilon \to 0} \zeta(\varepsilon) = 0$$

for a function h satisfying

$$\lim_{\varepsilon \to 0} h(\varepsilon)\beta_{\delta_1}(\varepsilon) = 0.$$

Our final hypothesis provides a lower bound on the probability of absorption on the event X_n^ε is sufficiently close to M_0.

HYPOTHESIS 2.3.   *There exists a neighborhood V_0 of M_0 such that*

$$\lim_{\varepsilon \to 0} \frac{\beta_{\delta_0}(\varepsilon)}{\inf_{x \in V_0} p^\varepsilon(x, M_0)} = 0.$$

We prove in Section 3 that, if M_0 is a global attractor, then μ is supported by M_0; see Theorem 3.12. The main result of this section is the following theorem. A proof is given in Section 3.

THEOREM 2.4. *Assume that Hypotheses* 2.1 *and* 2.2 *hold. Then any weak* limit point μ of $\{\mu^\varepsilon\}_{\varepsilon>0}$ satifies $\mu(V_j) = 0$ for $j = \ell + 1, \ldots, v$. In addition, if Hypothesis* 2.3 *holds, then μ is supported by the union of the attractors $\bigcup_{i=1}^\ell K_i$. Moreover, there exists a $\delta > 0$ such that $\lambda_\varepsilon \geq 1 - \beta_\delta(\varepsilon)$ for all $\varepsilon > 0$ sufficiently small.*

2.2. *Large deviation hypotheses.* For applications, it is often easier to verify certain large deviation hypotheses rather than Hypotheses 2.2 and 2.3. To this end we consider the following large deviation hypothesis.

HYPOTHESIS 2.5. *There exists a function $\rho : M \times M \to [0, +\infty]$ such that*:

  (i) *ρ is continuous on $M_1 \times M$,*
  (ii) *$\rho(x, y) = 0$ if and only if $y = F(x)$,*
  (iii) *for any $\beta > 0$,*

$$(3) \qquad \inf\{\rho(x, y) : x \in M, y \in M, d(F(x), y) > \beta\} > 0,$$

where $d(x, y) = \max_i |x_i - y_i|$,
  (iv) *for any open set U, we have the large deviations lower bound*

$$(4) \qquad \liminf_{\varepsilon \to 0} \varepsilon \log p^\varepsilon(x, U) \geq - \inf_{y \in U} \rho(x, y)$$

that holds uniformly for x in compact subsets of M_1 whenever U is an open ball in M. Additionally, for any closed set C, we have the uniform upper bound

$$(5) \qquad \limsup_{\varepsilon \to 0} \sup_{x \in M} \varepsilon \log p^\varepsilon(x, C) \leq - \inf_{y \in C} \rho(x, y).$$

Equations (3) and (5) imply that Standing Hypothesis 1.1 holds. Additionally, since M_0 is absorbing, (4) implies that $\rho(x, y) = +\infty$ for all $x \in M_0$, $y \in M_1$. The upper bound (5) can be weakened as a uniform bound on compact subsets of M_1. In that case, Hypothesis 1.1 is no longer implied by Hypothesis 2.5.

We also make the following assumption that ensures absorption is reasonably likely when the process is near the absorbing states.

HYPOTHESIS 2.6. *For any $c > 0$, there exists an open neighborhood V_0 of M_0 such that*

$$(6) \qquad \lim_{\varepsilon \to 0} \inf_{x \in V_0} \varepsilon \log p^\varepsilon(x, M_0) \geq -c.$$

To state our main result under these large deviation assumptions, we need to introduce an alternative notion of chain recurrence. Given $n \in \mathbb{N}^* = \{0, 1, 2, 3, \ldots\}$, define the function A_n on $M^{n+1} = \underbrace{M \times \cdots \times M}_{n+1 \text{ times}}$ by

$$\xi = (\xi_0, \ldots, \xi_n) \mapsto A_n(\xi) = \sum_{i=0}^{n-1} \rho(\xi_i, \xi_{i+1}).$$

A_n measures the "cost" of X_n^ε following the partial trajectory ξ where the cost is measured in terms of how much "noise" is required to move along this partial trajectory. For any x, y in M, we define

$$B_\rho(x, y) = \inf\{A_n(\xi) \mid n \geq 1, \xi \in M^{n+1}, \xi_0 = x, \xi_n = y\}.$$

The function $B_\rho(x, y)$ represents the minimal cost in going from x to y. B_ρ induces a partial order on M by writing $x <_\rho y$ (i.e., "x ρ-chains to y") if $B_\rho(x, y) = 0$. Roughly, x ρ-chains to y if the,re exist paths joining x to y with arbitrarily low costs. If $x <_\rho y$ and $y <_\rho x$, we write $x \sim_\rho y$.

We define the *set of ρ-chain recurrent points* $\mathcal{R}_\rho$ to be the set of points $x \in M$ such that $x \sim_\rho x$. The *ρ-basic classes* are the equivalence classes for $\sim_\rho$ restricted to the ρ-chain recurrent set. Since a point in M_0 never ρ-chains to a point in M_1, the ρ-basic classes are included either in M_0 or in M_1. In general, the ρ-basic classes and the ap-basic classes introduced in Section 2.1 need not be equivalent. For example, consider $F : [0, 1] \to [0, 1]$ given by the identity map $F(x) = x$ for all x and $M_0 = \varnothing$. Let $\rho(x, y) = |x - y|$. Then every point $\{x\}$ is a ρ-basic class, but the only ap-basic class is $[0, 1]$. However, unlike this example, if there is a finite number of ρ-basic classes, then we prove in Section 4 (see Theorem 4.12) that the ap-basic classes and ρ-basic classes agree.

Given a ρ-chain recurrent point x, let $[x]_\rho$ denote its ρ-basic class. We say that $[x]_\rho <_\rho [y]_\rho$ if $x <_\rho y$ and call *ρ-quasiattractors* the maximal ρ-equivalence classes. When a ρ-quasiattractor A is isolated (i.e., there is a neighborhood of the quasi-attractor containing no other ρ-chain recurrent point), we prove in Section 4 that A is an *attractor* for F; see Proposition 4.6.

In Section 4, we use Theorem 2.4 to prove the following result. Applications of Theorem 2.7 are given in Section 5.

THEOREM 2.7.   *Assume that Hypotheses 2.5 and 2.6 hold and that there exists a finite number of ρ-basic classes in M_1, which are closed. If:*

- *there is at least one ρ-quasiattractor A among the ρ-basic classes in M_1, and*
- *$\mu_\varepsilon(U) > 0$ for any neighborhood U of A and $\varepsilon > 0$,*

then any weak*-limit point of $\{\mu_\varepsilon\}_{\varepsilon > 0}$ is F-invariant and is supported by the union of ρ-quasiattractors in M_1. Moreover, there exists $c > 0$ such that $\lambda_\varepsilon \geq 1 - e^{-c/\varepsilon}$ for all $\varepsilon > 0$.

REMARK 2.8. Assume that there is a finite number of closed nonquasiattractors $[x_1]_\rho, \ldots, [x_N]_\rho$ in M_1 and $A = (\mathcal{R}_\rho \cap M_1) \setminus \bigcup_{i=1}^N [x_i]_\rho$ is an attractor for F. Then the main result still holds: if $\mu_\varepsilon(U) > 0$ for any neighborhood U of A and $\varepsilon > 0$, then any weak*-limit point of $\{\mu_\varepsilon\}_{\varepsilon > 0}$ is F-invariant and is supported by A. Moreover, there exists $C > 0$ such that $\lambda_\varepsilon \geq 1 - e^{-C/\varepsilon}$ for all $\varepsilon > 0$.

3. Proof of Theorem 2.4. In this section, we prove Theorem 2.4. We begin by proving several key results about ap-chain recurrence in Sections 3.1 and 3.2. In Section 3.3, we prove some key properties of limiting quasi-stationary distributions. A proof of Theorem 2.4 is given in Section 3.4. In Section 3.5, we show how our proof of Theorem 2.4 provides an alternative proof of the main result of Klebaner, Lazar and Zeitouni (1998). In addition to the Standing Hypotheses, the results in Section 3.2 require Hypothesis 2.1, and the proofs in Sections 3.4 and 3.5 require Hypotheses 2.1, 2.2 and 2.3.

3.1. *Absorption preserving chain recurrence.* We recall a few definitions and facts from dynamical systems. The *ω-limit set* of $B \subset M$ is given by $\omega(B) = \bigcap_{n \geq 1} \overline{\bigcup_{p \geq n} F^p(B)}$. It is the maximal invariant set in the closure of $\bigcup_{n \geq 0} F^n(B)$. An equivalent definition of an attractor presented in Section 2.1 is that a compact set A is an attractor for F provided it admits an open neighborhood U such that $\omega(U) = A$; the open set $\{x \in M : \omega(x) \subset A\}$ is then called the *basin of attraction* of A. By a classical result [see Conley (1978)], a compact set A is an attractor for F if and only if there exists an open set V which contains A and such that

$$(7) \qquad F(\overline{V}) \subset V, \qquad \bigcap_{n \in \mathbb{N}} F^n(V) = A.$$

Our assumption that $\|F\| = \sup_{x \in M} \|F(x)\| < \infty$ implies that the set $\mathcal{R}_{ap}$ of ap-chain recurrent points is included in $N^{\|F\|}(0)$. The relation $\sim_{ap}$, restricted to this set defines an equivalence relation. Unlike the ap-basic classes lying in M_0, the ap-basic classes lying in M_1 may not be closed. However, we have the following:

LEMMA 3.1. *Let x be an ap-chain recurrent point in M_1. Then $\overline{[x]_{ap}} \subset M_0 \cup [x]_{ap}$. In particular,*

$$\overline{[x]_{ap}} \subset M_1 \Rightarrow [x]_{ap} \ closed.$$

PROOF. Let $y \in \overline{[x]_{ap}}$. There exists a sequence $\{y_k\}$ in $[x]_{ap}$ which converges to y. Any ap δ-pseudoorbit from x to y_k is an ap 2δ-pseudoorbit from x to y, provided k is chosen large enough. Hence $x <_{ap} y$. On the other hand, assume that $y \notin M_0$ and consider an ap δ-pseudoorbit $(\xi_0, \ldots, \xi_n)$ chaining y_k to x. We have

$$d(F(y), \xi_1) \leq \delta + d(F(y), F(y_k)) \leq 2\delta$$

by continuity of F provided k is large enough. Consequently, $(y, \xi_1, \ldots, \xi_n)$ is an ap 2δ-pseudoorbit chaining y to x and $y \in [x]_{ap}$. $\square$

The following lemma shows that ap-basic classes are invariant.

LEMMA 3.2.   *Any ap-basic class* $[x]_{ap}$ *is positively* F*-invariant:* $F([x]_{ap}) \subseteq [x]_{ap}$. *If* $\overline{[x]_{ap}} \subset M_1$ *(which implies that* $[x]_{ap}$ *is closed), it is* F*-invariant,* $F([x]_{ap}) = [x]_{ap}$.

PROOF.   If $[x]_{ap}$ is a singleton, then $F(x) = x$, and there is nothing to prove. Assume that there exists $y \neq x$ such that $y \in [x]_{ap}$. For any $\delta > 0$, continuity and boundedness of $F(M)$ imply that there exists a $\delta/2 > \delta' > 0$ such that

$$d(z, F(x)) < \delta' \Rightarrow d(F(z), F^2(x)) < \delta/2 \qquad \text{for all } z \in M.$$

Pick an ap δ'-pseudoorbit $(x = \xi_0, \xi_1, \ldots, \xi_n = y)$ joining x to y. Since $d(\xi_1, F(x)) \leq \delta'$

$$d(F^2(x), \xi_2) \leq d(F^2(x), F(\xi_1)) + d(F(\xi_1), \xi_2) \leq \delta/2 + \delta' < \delta$$

and $(F(x), \xi_2, \ldots, \xi_n)$ is an ap δ-pseudoorbit joining $F(x)$ to y. Hence, $F([x]_{ap}) \subseteq [x]_{ap}$.

Next, let us assume that $[x]_{ap}$ is closed in M_1. For every $y \in [x]_{ap}$, we need to prove that $y = F(y')$ for some $y' \in [x]_{ap}$. For any $\delta > 0$, choose an ap δ-pseudoorbit $(\xi_i^\delta)_{i=0,\ldots,n(\delta)}$ joining y to itself. Now choose a compact set $K \subset M_1$ containing an open neighborhood of $[x]_{ap}$. We prove in the next section (see Remark 3.7) that the families ξ^δ can be chosen in such a way that they are contained in K. In particular the family $\xi_{n(\delta)-1}^\delta$ admits an accumulation point y' as $\delta \to 0$. By continuity of F, $F(y') = y$ and, therefore, $y' \sim_{ap} y \sim_{ap} x$.   □

For classical chain recurrence, $\omega(x)$ is contained in the chain recurrent set. While ap-chain recurrence shares this property whenever $\omega(x) \subset M_0$ or $\omega(x) \subset M_1$, in general it only satisfies a weaker property.

LEMMA 3.3.   *For* $x \in M$, $\omega(x) \cap \mathcal{R}_{ap} \neq \varnothing$.

PROOF.   If $x \in M_0$ or $\omega(x) \subset M_1$, then the classical result for chain recurrence implies $\omega(x) \subset \mathcal{R}_{ap}$. Suppose $x \in M_1$ and $y \in \omega(x) \cap M_0$. Then $\omega(y) \subset \mathcal{R}_{ap}$. Since $\omega(y) \subseteq \omega(x)$, the result follows.   □

LEMMA 3.4.   *If* $[x]_{ap}$ *is maximal, then* $x <_{ap} z$ *if and only if* $z \in [x]_{ap}$. *In particular, any ap-quasiattractor* $[x]_{ap}$ *is compact.*

PROOF.   Let z be such that $x <_{ap} z$. To prove that $z \in [x]_{ap}$, we need to show that $z <_{ap} x$. By Lemma 3.3, $\omega(z) \cap \mathcal{R}_{ap} \neq \varnothing$. Hence there exists $z' \in \omega(z) \cap \mathcal{R}_{ap}$. In particular, $x <_{ap} z <_{ap} z'$. As $z' \in \mathcal{R}_{ap}$, maximality of $[x]_{ap}$ implies that $z' \in [x]_{ap}$. Thus $z <_{ap} x$. In particular, if $y \in \overline{[x]_{ap}}$, then the proof of Lemma 4.1 implies that $y >_{ap} x$. Hence, $y \in [x]_{ap}$ and $[x]_{ap}$ is closed.   □

The next result is an easy consequence of Proposition 4.2 in Kifer (1988). A closed ap basic set is said to be isolated in M_1 if it admits an open neighborhood which is disjoint from any other ap basic class:

THEOREM 3.5. *Let $[x]_{\mathrm{ap}}$ be an isolated ap-quasiattractor in M_1. Then $[x]_{\mathrm{ap}}$ is an attractor.*

3.2. *Finiteness of the ap-basic classes.* Throughout this subsection, we require Hypothesis 2.1. Namely, there exists a finite number of ap-basic classes $\{K_i\}_{i=1}^{v}$ where the K_i are closed sets, $\{K_i\}_{i=1}^{\ell}$ are ap-quasiattractors and $\{K_i\}_{i=\ell+1}^{v}$ are nonap-quasiattractors.

The following result is proved in Kifer [(1988), pages 217–218] for classical chain recurrence. We give a proof adapted to our settings for the convenience of the reader.

LEMMA 3.6. (a) *For any $\theta > 0$ sufficiently small, there exists a quantity $0 < \delta(\theta) < \theta$ such that, if there is an ap $\delta(\theta)$-pseudoorbit $(\xi_0, \ldots, \xi_n)$ satisfying*

$$
\begin{aligned}
&d(\xi_0, K_i) < \delta(\theta), \qquad d(\xi_j, K_i) > \theta \quad and \\
&d(\xi_n, K_{i'}) < \delta(\theta) \qquad for\ some\ i, i' \in \{1, \ldots, v\}, j \in \{1, \ldots, n\},
\end{aligned}
\tag{8}
$$

then $i \neq i'$ and $K_{i'} >_{\mathrm{ap}} K_i$.

(b) *For any $\delta' > 0$, there exist $\delta \in (0, \delta')$ and $n_0 \geq 1$ such that any ap δ-pseudoorbit of length greater than n_0 must pass through $N^{\delta'}(\mathcal{R}_{\mathrm{ap}})$.*

PROOF. Assume that, for any $\delta > 0$, there exists an ap δ-pseudoorbit $(\xi_0^\delta, \ldots, \xi_{n(\delta)}^\delta)$ such that

$$
d(\xi_0^\delta, K_i) \leq \delta \quad and \quad d(\xi_{n(\delta)}^\delta, K_{i'}) \leq \delta.
\tag{9}
$$

Then there exists $\delta_k \downarrow 0$, $y \in K_i$ and $y' \in K_{i'}$ such that $\lim_{k\to\infty} \xi_0^{\delta_k} = y$ and $\lim_{k\to\infty} \xi_{n(\delta_k)}^{\delta_k} = y'$. Hence $d(F(y), \xi_1^{\delta_k}) < \delta_k + d(F(y), F(\xi_0^{\delta_k}))$ and $d(F(\xi_{n(\delta_k)-1}^{\delta_k}), y') \leq \delta_k + d(\xi_{n(\delta_k)}^{\delta_k}, y')$. Therefore, for any $\delta > 0$, $(y, \xi_1^{\delta_k}, \ldots, \xi_{n(\delta_k)-1}^{\delta_k}, y')$ is an ap δ-pseudoorbit provided that k is large enough. This proves that $K_{i'} >_{\mathrm{ap}} K_i$. As a consequence, if $K_{i'} >_{\mathrm{ap}} K_i$ does not hold, this means that there exists some quantity $\bar\delta > 0$ such that, for any $0 < \delta < \bar\delta$, we cant have (9). Now pick $\theta > 0$ smaller than $\bar\delta$.

Now assume that $i = i'$. Choose θ small enough such that $\overline{N^\theta(K_i)} \subset M_1$. Assume that there exist a decreasing sequence $\delta_k \downarrow 0$ and ap δ_k-pseudoorbits $(\xi_0^k, \ldots, \xi_{n_k}^k)$ such that (8) holds with $\xi = \xi^k$, $\delta = \delta_k$ and $j = j_k$. Without loss of generality, we may assume that $\lim_{k\to\infty} \xi_0^k = x \in K_i$, $\lim_{k\to\infty} \xi_{j_k}^k = y \in K \setminus N^\theta(K_i)$ and $\lim_{k\to\infty} \xi_{n_k}^k = z \in K_i$, where $K \subset M_1$ is a compact set

such that $F(N^\theta(K_i)) \subsetneq K$. We then have $d(\xi_1^k, F(x)) \leq \delta_k + d(F(\xi_0^k), F(x))$, $d(y, F(\xi_{j_k-1}^k)) \leq d(y, \xi_{j_k}^k) + \delta_k$, $d(F(y), \xi_{j_k+1}^k) \leq d(F(y), F(\xi_{j_k}^k)) + \delta_k$ and $d(z, F(\xi_{n_k-1}^k)) \leq d(y, \xi_{j_k}^k) + \delta_k$. By continuity of F, this implies that, for any $\delta > 0$, the sequence $(x, \xi_1^k, \ldots, \xi_{j_k-1}^k, y, \xi_{j_k+1}^k, \ldots, \xi_{n_k-1}^k, z)$ is an ap δ-pseudoorbit, provided k is large enough. Consequently, $x <_{\mathrm{ap}} y <_{\mathrm{ap}} z$ contradicting the fact that K_i is an ap-basic class.

We now prove point (b). For any $x \in M$ and $\gamma > 0$, Lemma 3.3 implies that the quantity

$$n^\gamma(x) := \inf\{n \in \mathbb{N} : F^n(x) \in N^\gamma(\mathcal{R}_{\mathrm{ap}})\}$$

is finite. By continuity of F, n^γ is upper-semicontinuous. Compactness of $\overline{F(M)}$ and upper semicontinuity imply that

$$n^\gamma := \max_{x \in M} n^\gamma(x) \leq \max_{y \in \overline{F(M)}} n^\gamma(y) + 1 < \infty.$$

Now assume that there exists $\delta' > 0$ such that the statement of (b) is not true. In particular, for each k there exists an ap $\delta_k = \delta'/k$-pseudoorbit of length $n^{\delta'/2}$, $\xi^k = (\xi_0^k, \ldots, \xi_{n^{\delta'/2}}^k)$, which does not enter $N^{\delta'}(\mathcal{R}_{\mathrm{ap}})$. Passing to a subsequence if necessary, we may assume that $\lim_{k \to \infty} \xi_j^k = \xi_j \in M$ for any $j = 1, \ldots, n^{\delta'/2}$. The sequence ξ is a partial solution of the discrete dynamical system induced by F, that is, $F(\xi_i) = \xi_{i+1}$ for $i = 0, \ldots, n^{\delta'/2} - 1$. The definition of $n^{\delta'/2}$ implies that there exists j_0 such that $d(\xi_{j_0}, \mathcal{R}_{\mathrm{ap}}) \leq \delta'/2$. Hence, $\xi_{j_0}^k \in N^{\delta'}(\mathcal{R}_{\mathrm{ap}})$ for k large enough which contradicts the choice of ξ^k. $\square$

REMARK 3.7.   Notice that, even without the finiteness assumption, the following statement still holds: given an ap-basic class $[x]_{\mathrm{ap}}$ in M_1 and $\theta > 0$, there exists a quantity $\delta > 0$ such that any ap δ-pseudoorbit joining $[x]_{\mathrm{ap}}$ to itself remains into $N^\theta([x]_{\mathrm{ap}})$.

COROLLARY 3.8.   *Given $\delta' > 0$, there exist isolating open neighborhoods $\{V_i\}_{i=1,\ldots,v}$ of the ap-basic classes $\{K_i\}_{i=1,\ldots,v}$, and positive constants δ_1 and n_0 such that*:

(a) *$\overline{N^{\delta_1}(K_i)} \subset V_i$ for $1 \leq i \leq v$;*

(b) *any ap δ_1-pseudoorbit starting in V_i remains in V_i for $i = 1, \ldots, \ell$;*

(c) *if there exists an ap δ_1-pseudoorbit $(\xi_0, \ldots, \xi_n)$ such that $\xi_0 \in N^{\delta_1}(K_i)$, $\xi_n \in N^{\delta_1}(K_{i'})$ and $\xi_k \notin V_i$ for some $2 \leq k \leq n - 1$, then $i \neq i'$ and $K_{i'} > K_i$.*

(d) *any ap δ_1-pseudoorbit of length greater than n_0 must pass through $N^{\delta'}(\mathcal{R}_{\mathrm{ap}})$.*

PROOF.   Choose $\theta \in (0, \delta')$ sufficiently small so that Lemma 3.6(a) holds, and let $\delta(\theta) > 0$ be as given by Lemma 3.6(a). Choose neighborhoods V_i of K_i such

that $\overline{N^\theta(K_i)} \subset V_i$ for $i = 1, \ldots, v$ and $F(\overline{V_i}) \subset V_i$ for $i = 1, \ldots, k$. The latter choice is possible as Lemma 3.1 implies that the ap-basic sets K_i are compact for $i = 1, \ldots, v$, and Theorem 3.5 implies that K_i is an attractor for $i = 1, \ldots, \ell$. Choose $\delta_1 \in (0, \delta(\theta))$ such that δ_1 is less than the δ given by Lemma 3.6(b) and such that any ap δ_1-pseudoorbit starting in V_i for $i = 1, \ldots, \ell$ remains in V_i. This latter choice is possible as $F(\overline{V_i}) \subset V_i$ for $i = 1, \ldots, \ell$. $\square$

3.3. *Limit behavior of quasi-stationary distributions.* Throughout this section, we assume that there exists a decreasing sequence $\varepsilon_n \downarrow 0$ such that, for every $n \in \mathbb{N}$, μ_n is a quasi stationary probability measure for p^{ε_n} with associated eigenvalue λ_n. Additionally, we assume that μ_n converges weakly to a Borel probability measure μ. We note that the results in this subsection do not require Hypotheses 2.1 or 2.3. Recall from Standing Hypothesis 1.1 that $\beta_\delta(\varepsilon) = \sup_{x \in M} p^\varepsilon(x, M \setminus N^\delta(F(x)))$.

LEMMA 3.9. *We have the following*:

(a) $\liminf_{n \to \infty} \lambda_n \geq \mu(M_1)$. *In particular, if μ is supported by M_1, then* $\lim_{n \to \infty} \lambda_n = 1$. *Alternatively, if $\lim_{n \to \infty} \lambda_n = 0$, then μ is supported by M_0.*

(b) *If there exists an attractor $A \subset M_1$ such that $\mu_n(U) > 0$ for every n and every open neighborhood U of A, then there exists $\delta > 0$ such that*

$$\lambda_n \geq 1 - \beta_\delta(\varepsilon_n)$$

for all n.

(b′) *If, in addition to the assumption of* (b), *there exists some neighborhood V_0 of M_0 such that*

$$\lim_{n \to \infty} \frac{\beta_\delta(\varepsilon_n)}{\inf_{x \in V_0} p^{\varepsilon_n}(x, M_0)} = 0,$$

then $\mu(V_0) = 0$.

PROOF. (a) Let $(\delta_k)_k$ be a positive sequence, decreasing to zero, and define

$$V_k := \{x \in M_1 : d(x, M_0) > \delta_k\}, \qquad U_k := F^{-1}(V_k).$$

Notice that $(U_k)_k$ is an increasing sequence of open sets such that $U_k \subset M_1$ (by F-invariance of M_0) and $\bigcup_k (U_k) = M_1$ (by closedness of M_0). We have

$$\lambda_n \geq \int_{U_k} \mu_n(dx) p^{\varepsilon_n}(x, M_1)$$

$$\geq \mu_n(U_k) \inf_{x \in U_k} p^{\varepsilon_n}(x, M_1)$$

$$= \mu_n(U_k)\Big(1 - \sup_{x \in U_k} p^{\varepsilon_n}(x, M_0)\Big).$$

Since $F(U_k) \subset V_k$, we have $N^{\delta_k}(F(U_k)) \subset M_1$. Thus

$$\lambda_n \geq \mu_n(U_k)\Big(1 - \sup_{x \in U_k} p^{\varepsilon_n}\big(x, N^{\delta_k}(F(x))^c\big)\Big).$$

By weak* convergence, the definition of λ_n and Standing Hypothesis 1.1,

$$\liminf_n \lambda_n \geq \liminf_n \mu_n(U_k) \geq \mu(U_k)$$

for all k. Point (a) follows since $\lim_k \mu(U_k) = \mu(M_1)$.

(b) Choose an open neighborhood U of A such that $F(\overline{U}) \subset U$ and $\delta > 0$ such that $N^{\delta}(F(\overline{U})) \subset U$. We have

$$\lambda_n \mu_n(U) \geq \mu_n(U)\Big(1 - \sup_{x \in \overline{U}} p^{\varepsilon_n}\big(x, U^c\big)\Big).$$

Since $p^{\varepsilon_n}(x, U^c) \leq p^{\varepsilon_n}(x, (N^{\delta}(F(\overline{U})))^c)$, and $\mu_n(U) > 0$, we get that $\lambda_n \geq 1 - \beta_{\delta}(\varepsilon_n)$.

(b′). By assumption, we have

$$1 - \beta_{\delta}(\varepsilon_n) \leq \lambda_n$$
$$= \int_M (1 - p^{\varepsilon_n}(x, M_0))\mu_n(dx)$$
$$\leq \mu_n(M \setminus V_0) + \mu_n(V_0)\Big(1 - \inf_{x \in V_0} p^{\varepsilon_n}(x, M_0)\Big),$$

which gives

$$\mu_n(V_0) \leq \frac{\beta_{\delta}(\varepsilon_n)}{\inf_{x \in V_0} p^{\varepsilon_n}(x, M_0)}.$$

Since V_0 is open and $\lim_{n \to \infty} \mu_n = \mu$ in the weak* topology, the result follows. $\square$

REMARK 3.10. Notice that we actually have a better result, as the quantity $\beta_{\delta}(\varepsilon_n)$ could be replaced by the smallest quantity

$$\sup_{x \in \overline{U}} p^{\varepsilon_n}\big(x, \big(N^{\delta}(F(\overline{U}))\big)^c\big).$$

PROPOSITION 3.11. *If* $\lim_{n \to \infty} \lambda_n = 1$, *then the probability measure* μ *is F-invariant. In particular, μ is supported by the closure of $\mathcal{R}_{\mathrm{ap}}$.*

PROOF. It suffices to verify that

$$(10) \qquad \int_M g(x)\mu(dx) = \int_M g(F(x))\mu(dx)$$

for any bounded continuous function $g : M \to \mathbb{R}$. Uniform continuity of g on $N^{\|F\|+\delta}(0)$ and Hypothesis 1.1 imply

$$\lim_{n \to \infty} \sup_x \left| \int_M (g(y) - g(F(x))) p^{\varepsilon_n}(x, dy) \right| = 0.$$

Therefore,

$$\left| \int_M (g(x) - g(F(x))) \mu_n(dx) \right|$$

$$= \left| \int_M \left(\lambda_n \int_M g(y) p^{\varepsilon_n}(x, dy) - g(F(x)) \right) \mu_n(dx) \right|$$

$$\leq 2(1 - \lambda_n) \|g\| + \left| \int_M \left(\int_M (g(y) - g(F(x))) p^{\varepsilon_n}(x, dy) \right) \mu_n(dx) \right|$$

$$\leq 2(1 - \lambda_n) \|g\| + \sup_x \left| \int_M (g(y) - g(F(x))) p^{\varepsilon_n}(x, dy) \right|.$$

Sending n to infinity implies (10) for any continuous bounded g. Hence, μ is F-invariant. F-invariance of μ implies that the support of μ is contained in the Birkhoff center of F, that is, the closure of recurrent points of F, $\{x \in M : x \in \omega(x)\}$, which is in turn included in the closure of $\mathcal{R}_{\text{ap}}$. $\quad\square$

The following theorem provides a sufficient condition for the support of the limiting measure μ to lie on the absorbing set M_0.

THEOREM 3.12. *Assume that M_0 is a global attractor. Then μ is supported by M_0.*

PROOF. If $\liminf_{n \to \infty} \lambda_n = 0$, Lemma 3.9 implies that $\mu(M_0) = 1$. Assume that $\liminf_{n \to \infty} \lambda_n > 0$, and let $c = \inf_n \lambda_n > 0$. Given $\alpha > 0$, pick an open neighborhood U of M_0, $\delta_1 > 0$ and $n_0 \in \mathbb{N}$ such that $U \subset N^\alpha(M_0)$, $F(\overline{U}) \subset U$, any ap δ_1-pseudoorbit starting in U remains in U and any ap δ_1-pseudoorbit of length at least n_0 eventually enters U; see Corollary 3.8.

Let $\mathcal{E}_{n,k}$ be the event $\{(X_j^{\varepsilon_n})_{j=0,\ldots,k}$ is an ap δ_1-pseudoorbit$\}$. Since a δ_1-pseudoorbit of length at least n_0 ends in U, we have

$$\mathbb{P}_x[X_k^{\varepsilon_n} \in U^c]$$

$$\leq \mathbb{P}_x[\mathcal{E}_{n,k}^c] + \mathbb{P}[\mathcal{E}_{n,k} \text{ and } X_k^{\varepsilon_n} \in U^c]$$

$$\leq \sum_{j=0}^{k-1} \mathbb{P}_x(d(X_{j+1}^{\varepsilon_n}, F(X_j^{\varepsilon_n})) > \delta_1) + 0$$

$$\leq k \beta_{\delta_1}(\varepsilon_n)$$

for $k \geq n_0$ and $x \in M$. The last inequality follows from the definition of $\beta_\delta(\varepsilon)$, and the second inequality from the fact that the event $\mathcal{E}_{n,k}^c$ is included in the union of the k events $\{d(X_{j+1}^{\varepsilon_n}, F(X_j^{\varepsilon_n})) > \delta_1\}$. By the definition of μ_n,

$$\mu_n(U^c) \leq \frac{1}{\lambda_n^{n_0}} \int_M \mu_n(dx) \mathbb{P}_x[X_{n_0}^{\varepsilon_n} \in U^c] \leq \frac{n_0 \beta_{\delta_1}(\varepsilon_n)}{c^{n_0}}$$

and the last quantity goes to zero as n goes to infinity. Since $\alpha > 0$ was arbitrary, $\mu(M_1) = 0$. $\square$

REMARK 3.13.   The proof is not needed in the particular case where λ_n goes to one since μ is then F-invariant and the Birkhoff center is contained in M_0.

3.4. *Proof of Theorem* 2.4.   We assume Hypotheses 2.1, 2.2 and 2.3 hold. Recall that, under the finiteness assumption, the ap-quasiattractors $\{K_i\}_{i=1,\ldots,\ell}$ are actually attractors; see Theorem 3.5. Also, there exists $\delta_0 > 0$ such that $\lambda_n \geq 1 - \beta_{\delta_0}(\varepsilon_n)$; see Lemma 3.9(b). Let $\{V_i\}_{i=1,\ldots,v}$ and $\delta_1 \leq \delta_0$ be chosen as in Corollary 3.8. Given a Borel set V define $\tau_V^n = \inf\{j \geq 0 : X_j^{\varepsilon_n} \notin V\}$.

Call $b = v - \ell$ the number of nonap-quasiattractors in M_1 and $K = \bigcup_{i=1}^v K_i$. Choose sequences $\{m_n\}_{n \geq 1}$ and $\{m_n'\}_{n \geq 1}$ such that

$$\lim_{n \to \infty} \beta_{\delta_1}(\varepsilon_n) m_n = 0, \qquad \lim_{n \to \infty} \frac{m_n'}{m_n} = 0 \quad \text{and} \quad \lim_{n \to \infty} \frac{h(\varepsilon_n)}{m_n'} = 0.$$

The presence of an attractor inside M_1 such that $\mu_n(U) > 0$ for any n, and any open neighborhood U implies that $\lim_{n \to \infty} \lambda_n = 1$, by Lemma 3.9(b). Proposition 3.11 implies that μ is F-invariant and supported by the closure of $\mathcal{R}_{\mathrm{ap}}$.

Let us prove the first statement of Theorem 2.4. Let $j \in \{\ell + 1, \ldots, v\}$ be fixed. By definition of λ_n,

$$\mu_n(V_j) = \frac{1}{\lambda_n^r} \int_{x \in M} \mu_n(dx) \mathbb{P}_x[X_r^{\varepsilon_n} \in V_j] \qquad \forall r \in \mathbb{N}^*.$$

For $i = \{1, \ldots, b\}$, call t_n^i the integer $\lfloor m_n/i \rfloor$. Let $\mathcal{E}_n$ and $\mathcal{E}_n'$ be the events, respectively, defined by

$$\mathcal{E}_n = \{(X_i^{\varepsilon_n})_{i=1,\ldots,m_n} \text{ is a } \delta_1\text{-pseudoorbit}\}$$

and

$$\mathcal{E}_n' = \{\forall i \in \{\ell + 1, \ldots, v\}, \ \forall q \geq m_n', X_p^{\varepsilon_n} \in N^{\delta_1}(K_i) \Rightarrow X_{p+q}^{\varepsilon_n} \notin N^{\delta_1}(K_i)\}.$$

The set $\mathcal{E}_n'$ is the event "after its first entry in any $N^{\delta_1}(K_i)$, the Markov chain will have escaped from this set after m_n' steps and will never come back."

On the event $\mathcal{E}_n \cap \mathcal{E}_n'$, the process $(X_1^{\varepsilon_n}, \ldots, X_{m_n}^{\varepsilon})$ is an ap δ_1-pseudoorbit and therefore gets trapped in $\bigcup_{i=1}^\ell V_i$ if it enters in this set. Corollary 3.8 implies that it cannot spend more than b blocks of length at most m_n' in $\bigcup_{i=\ell+1}^v N^{\delta_1}(K_i)$. In

particular, for n large enough, $X_{m_n}^{\varepsilon_n}$ is in V_j only if $X_{t_n^i}^{\varepsilon} \in (N^{\delta_1}(K))^c$ for some $i \in \{1, \ldots, b\}$. Therefore,

$$\mathbb{P}_x\big[\{X_{m_n}^{\varepsilon_n} \in V_j\} \cap \mathcal{E}_n \cap \mathcal{E}_n'\big] \le \sum_{i=1}^{b} \mathbb{P}_x\big[X_{t_n^i}^{\varepsilon_n} \notin N^{\delta_1}(K)\big].$$

On the other hand on the event $\mathcal{E}_n$, starting from $N^{\delta_1}(K_i)$, the chain cannot enter back into $N^{\delta_1}(K_i)$ once it exited V_i (by Corollary 3.8(c)). Hypothesis 2.2 implies

$$\mathbb{P}_x\big[(\mathcal{E}_n')^c \cap \mathcal{E}_n\big] \le \sum_{i=\ell+1}^{v} \sup_{y \in V_i} \mathbb{P}_y\big[\tau_{V_i}^n \ge m_n'\big] \le b\zeta(\varepsilon_n)$$

as $m_n' > h(\varepsilon_n)$ for n sufficiently large. Consequently,

$$\mathbb{P}_x\big[X_{m_n}^{\varepsilon_n} \in V_j\big]$$

$$\le \mathbb{P}_x\big[(\mathcal{E}_n)^c\big] + \mathbb{P}_x\big[(\mathcal{E}_n')^c \cap \mathcal{E}_n\big] + \sum_{i=1}^{b} \mathbb{P}_x\big[X_{t_n^i}^{\varepsilon_n} \notin N^{\delta_1}(K)\big]$$

$$\le m_n \beta_{\delta_1}(\varepsilon_n) + b\zeta(\varepsilon_n) + \sum_{i=1}^{b} \mathbb{P}_x\big[X_{t_n^i}^{\varepsilon_n} \notin N^{\delta_1}(K)\big]$$

for n sufficiently large. Therefore we have, using the invariance properties of μ_n,

$$\int \mu_n(dx)\mathbb{P}_x\big[X_{m_n}^{\varepsilon_n} \in V_j\big]$$

$$\le m_n \beta_{\delta_1}(\varepsilon_n) + b\zeta(\varepsilon_n) + \sum_{i=1}^{b} \int \mu_n(dx)\mathbb{P}_x\big[X_{t_n^i}^{\varepsilon_n} \notin N^{\delta_1}(K)\big]$$

$$\le m_n \beta_{\delta_1}(\varepsilon_n) + b\zeta(\varepsilon_n) + \sum_{i=1}^{b} \lambda_n^{t_n^i} \mu_n\big((N^{\delta_1}(K))^c\big)$$

$$\le m_n \beta_{\delta_1}(\varepsilon_n) + b\zeta(\varepsilon_n) + b\mu_n\big((N^{\delta_1}(K))^c\big),$$

which converges to 0 as $n \to \infty$. By our choice of the sequence m_n,

$$\liminf_{n \to \infty} \lambda_n^{m_n} \ge \liminf_{n \to \infty}(1 - \beta_{\delta_1}(\varepsilon_n))^{m_n} = 1.$$

Hence, $\lim_{n \to \infty} \lambda_n^{m_n} = 1$ and

$$\lim_{n \to \infty} \mu_n(V_j) = \lim_{n \to \infty} \frac{1}{\lambda_n^{m_n}} \int \mu_n(dx)\mathbb{P}_x\big[X_{m_n}^{\varepsilon_n} \in V_j\big] = 0.$$

The proof of the first statement is complete since $\mu(V_j) \le \liminf_n \mu_n(V_j)$.

Now, under Hypothesis 2.3, Lemma 3.9 implies that the support of μ is contained in $M \setminus V_0$. In particular, $\mu(\mathcal{R}_{ap} \cap M_1) = 1$. Hence, $\mu(\bigcup_{i=1}^{v} K_j) = 1$ and the second statement of Theorem 2.4 follows.

3.5. *A derivation of Theorem* 3 *of Klebaner, Lazar and Zeitouni* (1998). We assume Hypotheses 2.1, 2.2 and 2.3 hold. We can obtain Theorem 3 of Klebaner, Lazar and Zeitouni (1998) as a consequence of our proof of Theorem 2.4. To see why, we describe how their assumptions (A1)–(A6) imply our main assumptions. For the sake of brevity, we do not state their assumptions here. Rather we refer the interested reader to their article. Under their assumptions (A1)–(A6), Klebaner, Lazar and Zeitouni (1998) prove nonconvergence to the finite set of unstable equilibria for one-dimensional maps. Their assumptions (A1) and (A2) guarantee the following, calling $M_0 = \{0, 1\}$ to fit our settings,[2] so we have:

(a) $f(M_0) \subset M_0$ and $f(M_1) \subset M_1$;

(b) the absorbing state $\{0\}$, the unstable equilibria $(x_i^*)_{i=0,\dots,k}$, and the stable equilibria $(s_i)_{i=1,\dots,l}$ form a Morse decomposition for the dynamical system induced by f; hence there is a finite number of ap-basic classes (see Proposition 5.1) and the ap-quasiattractors in M_1 are the $(s_i)_{i=1,\dots,l}$. Our Hypothesis 2.1 is verified.

By their assumption (A4), we derive the uniform (in x) upper bound

$$\exists \lambda_0 > 0, C > 0 \qquad \text{such that } P_\varepsilon\big(x, \big(N^\delta(f(x))\big)^c\big) \leq Ce^{-\lambda_0 \delta/\varepsilon}$$

for any $\delta > 0$, $\varepsilon > 0$. Hence our Standing Hypothesis 1.1 is satisfied, with $\beta_\delta(\varepsilon) = Ce^{-\lambda_0 \delta/\varepsilon}$.

It remains to check Hypothesis 2.2. The nonattractors in this case are the unstable equilibria $(x_j^*)_{j=0,\dots,k}$. Let δ_1 be the positive parameter associated with the stable points $(s_i)_{i=1,\dots,l}$ in our Corollary 3.8. By their assumption (A5), there exists $\beta > 0$ such that we have, for $j = 1, \dots, k$, ε and δ small enough,

$$\inf_{x \in V_j} \mathbb{P}_x\big(\xi^\varepsilon(x) > \varepsilon\big) \geq \beta, \qquad \inf_{x \in V_j} \mathbb{P}_x\big(\xi^\varepsilon(x) < -\varepsilon\big) \geq \beta,$$

where $V_j = N^\delta(x_j^*)$. Call $\alpha_n = \delta/\varepsilon_n$ and assume without loss of generality that it is an integer. Using the fact that x_j^* is an unstable equilibrium, for n large enough,

$$\inf_{x \in V_j} \mathbb{P}_x\big(|X_{\alpha_n}^{\varepsilon_n} - x_j^*| \geq \delta\big) \geq \beta^{\alpha_n}.$$

Now choose $0 < a < \lambda_0 \delta_1$. We have, by Markov's property,

$$\sup_{x \in V_j} \mathbb{P}_x\big(\tau_{V_j}^n \geq e^{a/\varepsilon_n}\big) \leq \big(1 - \beta^{\alpha_n}\big)^{\nu_n},$$

where $\nu_n = \frac{\varepsilon_n e^{a/\varepsilon_n}}{\delta}$. By construction, $e^{a/\varepsilon_n}\beta_{\delta_1}(\varepsilon_n)$ goes to zero as n goes to infinity. Additionally,

$$\big(1 - \beta^{\alpha_n}\big)^{\nu_n} \sim_{n \to +\infty} \exp\left(-\frac{\varepsilon_n e^{a/\varepsilon_n} e^{\delta \log(\beta)/\varepsilon_n}}{\delta}\right).$$

---

[2]In the quoted paper, $M_1 = [0, 1]$ and M_0 is $[0, 1]^c$, but it does not change the problem as they consider continuous state space; see Ramanan and Zeitouni (1999).

This quantity vanishes as n goes to infinity if we choose δ small enough (more precisely, δ must be chosen smaller than $-a/\log\beta$). Therefore, we have verified Hypothesis 2.2, and we can apply our result to conclude that the support of weak* limit points of the QSDs do not include the unstable equilibria x_j^*.

4. Properties of ρ-basic sets and proof of Theorem 2.7. In this section, we assume that the assumptions of Theorem 2.7 hold. We prove that Hypotheses 2.5, 2.6 and finiteness of the ρ-basic classes imply Hypotheses 2.1, 2.2 and 2.3. From these implications, it follows that Theorem 2.4, Lemma 3.9 and Proposition 3.11 imply that Theorem 2.7 holds. Indeed, by Lemma 3.9, we obtain the lower bound for λ_ε and, by Proposition 3.11, that any weak*-limit point of $\{\mu_\varepsilon\}_{\varepsilon>0}$ is F-invariant. The fact that these limiting measures sit on ρ-quasiattractors follows from Theorem 2.4.

LEMMA 4.1. *Hypotheses 2.5 and 2.6 imply Hypothesis 2.3.*

PROOF. Let $\delta_0 > 0$ be given. Assertion (iii) of Hypothesis 2.5 implies that

$$c = \tfrac{1}{4}\inf\{\rho(x, y) : x, y \in M, d(F(x), y) > \delta_0\} > 0.$$

It follows from the definition of β_δ and inequality (5) that $\beta_{\delta_0}(\varepsilon) \leq \exp(-3c/\varepsilon)$ for $\varepsilon > 0$ sufficiently small. Hypothesis 2.6 implies that there exists an open neighborhood V_0 of M_0 such that $p^\varepsilon(x, M_0) \geq \exp(-2c/\varepsilon)$ for $\varepsilon > 0$ sufficiently small and $x \in V_0$. Hence,

$$\lim_{\varepsilon\to 0}\frac{\beta_{\delta_0}(\varepsilon)}{\inf_{x\in V_0} p^\varepsilon(x, M_0)} \leq \lim_{\varepsilon\to 0}\exp(-c/\varepsilon) = 0. \qquad \square$$

For the remaining implications, we need to gain some insights about the relation between ap and ρ-chain recurrence. As their ap counterparts, the ρ-basic classes whose closure is in M_1 are actually closed. This follows from the next two lemmas. We consider the quantity

$$\alpha^* = \sup_{\delta>0}\inf\{\rho(x, y) : x \in M, y \in M, d(F(x), y) > \delta\} \in (0, +\infty].$$

LEMMA 4.2. *For any $0 \leq \alpha < \alpha^*$, there exists $\delta > 0$ such that, for any $\xi = (\xi_0, \ldots, \xi_n)$ satisfying $A_n(\xi) < \alpha$, we have $\|\xi_i\| \leq \|F\| + \delta$ for $i = 1, \ldots, n$. In particular, $\mathcal{R}_\rho$ is bounded.*

PROOF. Given $0 \leq \alpha < \alpha^*$, there exists $\delta > 0$ such that

$$\rho(x, y) < \alpha \Rightarrow d(F(x), y) < \delta.$$

Therefore, if $A_n(\xi) < \alpha$, then $\rho(\xi_i, \xi_{i+1}) < \alpha$ for $i = 0, \ldots, n-1$, which implies that $d(\xi_{i+1}, F(\xi_i)) < \delta$ for $i = 0, \ldots, n-1$. $\quad\square$

LEMMA 4.3.   *We have the following*:

  (i) *The function B_ρ is upper semicontinuous on $M_1 \times M$.*

 (ii) *Let $\eta \in (0, \alpha^*)$ and $y \in M$. If the set $\{x \in M : B_\rho(x, y) \leq \eta\}$ has its closure in M_1, then it is closed.*

(iii) *Let $x \in M_1$. Assume that the set $\{y \in M : B_\rho(x, y) \leq \eta\}$ has its closure in M_1 for η small enough. Then there exists η_0 such that, for any $\eta < \eta_0$, $\{y \in M : B_\rho(x, y) \leq \eta\}$ is closed.*

PROOF.   Part (i) is proved in Kifer [(1988), Lemma 5.1, page 58]. It relies on the continuity of ρ on $M_1 \times M$.

For part (ii), let $\{x_k\}_{k \geq 1}$ be a sequence of points in M_1, such that $\lim_{k \to \infty} x_k = x \in M_1$ and $B_\rho(x_k, y) \leq \eta$. Pick $r > 0$ such that $\overline{N^r(x)} \subset M_1$. For any $\gamma > 0$ and any $k \geq 1$ there exists $\xi^{\gamma,k} = (\xi_0^{\gamma,k}, \ldots, \xi_{n_k}^{\gamma,k})$ such that $\xi_0^{\gamma,k} = x_k$, $\xi_{n_k}^{\gamma,k} = y$ and $A_{n_k}(\xi^{\gamma,k}) \leq \eta + \gamma$. By Lemma 4.2, for $\gamma < \alpha^* - \eta$, there exists $\delta > 0$ such that $\|\xi_1^{\gamma,k}\| \leq \|F\| + \delta$. Since $\overline{N^r(x)}$ is a compact set contained in M_1, and ρ is continuous on $M_1 \times M$, ρ is uniformly continuous on $\overline{N^r(x)} \times \overline{N^{\|F\|+\delta}(0)}$. Thus

$$\lim_{k \to \infty} \left| \rho(x, \xi_1^{\gamma,k}) - \rho(x_k, \xi_1^{\gamma,k}) \right| = 0.$$

Therefore,

$$B_\rho(x, y) \leq \liminf_{k \to \infty} \left(\rho(x, \xi_1^{\gamma,k}) + \rho(\xi_1^{\gamma,k}, \xi_2^{\gamma,k}) + \cdots + \rho(\xi_{n_k-1}^{\gamma,k}, y) \right)$$

$$\leq \liminf_{k \to \infty} \left| \rho(x, \xi_1^{\gamma,k}) - \rho(x_k, \xi_1^{\gamma,k}) \right| + A_{n_k}(\xi^{\gamma,k}) \leq \eta + \gamma.$$

Since this holds for any $\gamma > 0$, part (ii) follows.

Proof of part (iii) is similar. However, we have to be careful since ρ is not continuous on $M_0 \times M$. Let $x \in M_1$ be given, and assume that there exists $\overline{\eta} > 0$ such that $K = \overline{\{y \in M : B_\rho(x, y) \leq \overline{\eta}\}} \subset M_1$. Define $a = d(M_0 \cap N^{\|F\|}(0), K) > 0$. Since ρ is continuous on $M_1 \times M$ and $\rho(z, y) = 0$ if and only if $y = F(z)$, there exists $\alpha > 0$ such that

$$\rho(z, y) < \alpha \Rightarrow d(F(z), y) < \min(a/2, 1)$$

for all $z, y \in M$. Let $\eta_0 = \min(\alpha, \overline{\eta})$ and choose $0 < \eta < \eta_0$. We claim that $\{y \in M : B_\rho(x, y) \leq \eta\}$ is closed. To see why, let $\{y_k\}_{k \geq 1}$ be a sequence in M_1 such that $\lim_{k \to \infty} y_k = y \in M_1$ and $B_\rho(x, y_k) \leq \eta$ for all k. For any $\gamma > 0$, there exists a family $\xi^{\gamma,k} = (\xi_0^{\gamma,k}, \ldots, \xi_{n_k}^{\gamma,k})$ such that $\xi_0^{\gamma,k} = x$, $\xi_{n_k}^{\gamma,k} = y_k$ and $A_{n_k}(\xi^{\gamma,k}) \leq \eta + \gamma$. For $\gamma < \eta_0 - \eta$, $\rho(\xi_{n_k-1}^{\gamma,k}, y_k) < \alpha$. Therefore, $d(F(\xi_{n_k-1}^{\gamma,k}), y_k) < \min(a/2, 1)$, which implies that $d(F(\xi_{n_k-1}^{\gamma,k}), M_0) > a/2$. By continuity of F and F-invariance of M_1, the sequence $\{\xi_{n_k-1}^{\gamma,k}\}_k$ is bounded away from M_0. Since

$\eta_0 < \alpha$, Lemma 4.2 implies that there exists $\delta > 0$ such that $\{\xi_{n_k-1}^{\gamma,k}\}_k \subset N^{\|F\|+\delta}(0)$. The remainder of the proof is as for part (ii), with $\xi_{n_k-1}^{\gamma,k}$ playing the role of $\xi_1^{\gamma,k}$. $\qquad \square$

Since boundedness of ρ-basic classes follows from Lemma 4.2, Lemma 4.3 implies that given a ρ-chain recurrent point x, if $\overline{[x]_\rho} \subset M_1$, then $[x]_\rho$ is compact. Clearly, if x is ρ-chain recurrent and $[x]_\rho$ is closed, then x is ap-chain recurrent and $[x]_\rho \subset [x]_{\text{ap}}$, but the converse is not true in general. For example, consider a situation where the ap δ-chains joining x to itself have arbitrarily large length as δ goes to zero, in which case we could have $x \sim_{\text{ap}} x$ but $x \not\sim_\rho x$. While Remark 3.7 holds for ρ-chain recurrence, it is not immediate, and therefore we provide a short proof. In the sequel, we will call δ-ρ-pseudoorbit any family $\xi_0, \ldots, \xi_n$ such that $A_n(\xi) \leq \delta$.

LEMMA 4.4. *Let $[x]_\rho$ be a closed ρ-basic class in M_1. For any $\theta > 0$, there exists $\delta > 0$ such that any δ-ρ-pseudoorbit joining $[x]_\rho$ to itself is contained in $N^\theta([x]_\rho)$.*

PROOF. Pick θ small enough so that $\overline{N^\theta([x]_\rho)} \subset M_1$. Since F is M_1-invariant and $\overline{N^\theta([x]_\rho)}$ is compact and contained in M_1, so is its image by F. Hence, by closedness of M_0, there exists a compact set $K \subset M_1$, which contains the γ-neighborhood of $F(N^\theta([x]_\rho))$, for some $\gamma > 0$. Assume, by contradiction, that there exist a decreasing sequence $\delta_k \downarrow 0$ and δ_k-ρ-pseudoorbits $(\xi_0^k, \ldots, \xi_{n_k}^k)$ [i.e., $A_{n_k}(\xi^k) \leq \delta_k$] such that $\lim_{k\to\infty} \xi_0^k = u \in [x]_\rho$, $\lim_{k\to\infty} \xi_{n_k}^k = w \in [x]_\rho$ and $j_k = \min\{j \geq 1 : \xi_j^k \notin N^\theta([x]_\rho)\} < n_k$. For k large enough, we have

$$\rho(z, y) < \delta_k \quad \Rightarrow \quad d(F(z), y) < \gamma$$

for all $z, y \in M$. Since $\xi_{j_k-1}^k \in N^\theta([x]_\rho)$, we have $\xi_{j_k}^k \in K$ for k large enough. By passing to a subsequence if necessary, $\xi_{j_k}^k$ converges to some point $v \in K \setminus N^\theta([x]_\rho)$. On the other hand, consider the pseudoorbits $\tilde{\xi}^k = (\xi_0^k, \ldots, \xi_{j_k-1}^k, v)$. They satisfy

$$\lim_{k\to\infty} A_{j_k}(\tilde{\xi}^k) \leq \lim_{k\to\infty} A_{n_k}(\xi^k) + \lim_{k\to\infty} |\rho(\xi_{j_k-1}^k, v) - \rho(\xi_{j_k-1}^k, \xi_{j_k}^k)| = 0$$

due to $A_{n_k}(\xi^k) \leq \delta_k$ and by uniform continuity of ρ on $K \times K$. Hence, $x <_\rho v$. Similarly, one can show that $v <_\rho x$. Consequently, $v \in [x]_\rho$, a contradiction. $\quad \square$

LEMMA 4.5. *The ρ-basic classes $[x]_\rho$ closed in M_1 are invariant: $F([x]_\rho) = [x]_\rho$.*

PROOF. Since $F(x) >_{\mathrm{ap}} x$, $F([x]_\rho) \subset [x]_\rho$ follows if we prove that $F(x) <_\rho x$. Since $[x]_\rho$ is compact and contained in M_1 (see Lemma 4.2), we can find a compact set K bounded away from M_0, $\delta_k \downarrow 0$ and a family of δ_k-ρ-pseudoorbits $\xi^k = (\xi_0^k, \ldots, \xi_{n_k}^k)$ in K such that $\xi_0^k = \xi_{n_k}^k = x$. Consider the family $\tilde{\xi}^k = (F(x), \xi_2^k, \ldots, \xi_{n_k-1}^k, x)$. By uniform continuity of ρ on $K \times K$ and the fact that $\rho(x, y) = 0$ if and only if $y = F(x)$, we may assume by passing to a subsequence if necessary that $\lim_{k \to \infty} \xi_1^k = F(x)$. Hence

$$\lim_{k \to \infty} A_{n_k-1}(\tilde{\xi}^k) \le \lim_{k \to \infty} A_{n_k}(\xi^k) + |\rho(F(x), \xi_2^k) - \rho(\xi_1^k, \xi_2^k)| = 0$$

as $A_{n_k}(\xi^k) \le \delta_k$ and using uniform continuity of ρ on $K \times K$. Hence, $F(x) <_\rho x$.

For the inclusion $[x]_{\mathrm{ap}} \subset F([x]_{\mathrm{ap}})$, pick $y \in [x]_\rho$ such that $y \ne F(y)$ (if there is no such y, there is nothing to prove). For $\delta_k \downarrow 0$, choose a family of δ_k-ρ-pseudoorbits $\xi^k = (\xi_0^k, \ldots, \xi_{n_k}^k)$ in K such that $\xi_0^k = \xi_{n_k}^k = y$. Passing to a subsequence if necessary, we can assume that $\lim_{k \to \infty} \xi_{n_k-1}^k = z \in K$. Clearly, $F(z) = y$ and $z \sim_\rho y$. Hence, $[x]_\rho \subset F([x]_\rho)$. $\square$

The following proposition is a straightforward consequence of Proposition 5.1 in Kifer (1988).

PROPOSITION 4.6. *Let $[x]_\rho$ be an isolated ρ-quasiattractor in M_1. Then it is an attractor and $[x]_\rho = \bigcap_{\eta > 0} D_\eta$, where*

$$D_\eta = \{y \in M : B_\rho(x, y) < \eta\}.$$

Define the maximum distance on M^{n+1} by $d_n(\zeta, \xi) = \max_{j=0,\ldots,n} d(\zeta_j, \xi_j)$ for $(\zeta_0, \ldots, \zeta_n), (\xi_0, \ldots, \xi_n) \in M^{n+1}$. The following theorem and lemma are analogous to the statements of Theorem 5.2(a) and Lemma 5.3 in Kifer [(1988), pages 66 and 72, resp.]. We provide a proof of Theorem 4.7 that slightly differs from the proof of Kifer. The proof of Lemma 4.8 follows directly from Kifer's proof of his Lemma 5.3.

THEOREM 4.7. *Let $K \subset M_1$ be a compact set. Given $\eta, \delta, N > 0$, there exists $\varepsilon_0 > 0$ such that*

$$\mathbb{P}_x[d_n((X_0^\varepsilon, \ldots, X_n^\varepsilon), \xi) < \eta] \ge \exp\left(-\frac{A_n(\xi) + \delta}{\varepsilon}\right)$$

for any $x \in K$, $\varepsilon < \varepsilon_0$, $n \le N$ and $\xi = (\xi_0, \ldots, \xi_n) \in K^{n+1}$ with $\xi_0 = x$.

PROOF. Analogously to Kifer's proof, we introduce the quantity

$$n_\gamma^K = \sup\{|\rho(y, z) - \rho(y', z')| : y, y' \in K, d(y, y') \le \gamma, d(z, z') \le \gamma\}.$$

By uniform continuity of ρ on compact subsets of $M_1 \times M_1$, $\lim_{\gamma \to 0} n_\gamma^K = 0$. Let η, δ, and N be given. Choose $0 < \gamma < \eta$ such that $n_\gamma^k < \delta/2N$ and $N^\gamma(K) \subset M_1$.

Now let $\xi = (\xi_0 = x, \xi_1, \ldots, \xi_n) \in K^{n+1}$. By the uniform lower bound (4) of Hypothesis 2.5 there exists a function $g :]0, +\infty[\rightarrow]0, +\infty[$ such that $\lim_{\varepsilon \to 0} g(\varepsilon) = 0$ and

$$\varepsilon \log p^\varepsilon(x, N^\gamma(\xi_i)) \geq -\inf_{y \in N^\gamma(\xi_i)} \rho(x, y) - g(\varepsilon)$$

for any $x \in N^\gamma(K)$ and $1 \leq i \leq n$.

Hence we have

$$\mathbb{P}_x[d_n(X^\varepsilon, \xi) < \eta]$$
$$\geq \mathbb{P}_x[d_n(X^\varepsilon, \xi) < \gamma]$$
$$= \int_{x_1 \in N^\gamma(\xi_1)} p^\varepsilon(x, dx_1) \cdots \int_{x_n \in N^\gamma(\xi_n)} p^\varepsilon(x_{n-1}, dx_n)$$
$$\geq p^\varepsilon(x, N^\gamma(\xi_1)) \prod_{i=1}^{n-1} \inf_{x_i \in N^\gamma(\xi_i)} p^\varepsilon(x_i, N^\gamma(\xi_{i+1}))$$
$$\geq \exp\left[-\frac{1}{\varepsilon}\left(\inf_{y \in N^\gamma(\xi_1)} \rho(x, y) + \sum_{i=1}^{n-1} \sup_{x_i \in N^\gamma(\xi_i)} \inf_{y_i \in N^\gamma(\xi_{i+1})} \rho(x_i, y_i) + ng(\varepsilon)\right)\right]$$
$$\geq \exp\left[-\frac{1}{\varepsilon}\left(A_n(\xi) + ng(\varepsilon) + nn_\gamma^K\right)\right].$$

The result follows by choosing ε_0 small enough so that $g(\varepsilon) \leq \delta/2N$, for every $\varepsilon < \varepsilon_0$. $\square$

LEMMA 4.8. *Let K be a compact set in M which does not contain any entire semiorbits $\{F^i(x), i \in \mathbb{N}\}$. Then there exists $a > 0$ and $N \in \mathbb{N}$ (which depend on K) such that*:

(a) *for any sequence $\xi \in K^n$ with $n > N$, we have $A_n(\xi) > (n - N)a$;*
(b) *there exists $\varepsilon_0 > 0$ such that, for any $n > N$ and any $0 < \varepsilon < \varepsilon_0$,*

$$\sup_{x \in K} \mathbb{P}_x[\tau_K^\varepsilon > n] \leq e^{-((n-N)a)/\varepsilon},$$

where $\tau_K^\varepsilon = \inf\{j \geq 0 : X_j^\varepsilon \notin K\}$.

Recall that $\omega(x) = \bigcap_{n \geq 1} \overline{\bigcup_{m \geq n} F^m(x)}$ and that a point $x \in M$ is called *non-wandering* if for all open neighborhoods U of x and any $N \in \mathbb{N}$, there exists $n \geq N$ such that $F^n(U) \cap U \neq \varnothing$. We denote by $\mathrm{NW}(F)$ the set of nonwandering points of F. Note that ω-limit points are always nonwandering: $\{y \in M : y \in \omega(x), \text{ for some } x\} \subset \mathrm{NW}(F)$.

LEMMA 4.9. *The set $\mathrm{NW}(F) \cap M_1$ is contained in $\mathcal{R}_\rho$. In particular, any ω-limit point in M_1 is also in $\mathcal{R}_\rho$.*

PROOF.   Let $x \in M_1 \cap \mathrm{NW}(F)$ and $\delta > 0$ be given. By continuity of ρ in $M_1 \times M$ and $\rho(x, F(x)) = 0$ for all x, there exists $\gamma > 0$ such that

$$\rho(x, F(y)) < \delta/2 \quad \text{and} \quad \rho(z, x) < \delta/2$$

for $y \in N^\gamma(x)$ and $F(z) \in N^\gamma(x)$. Since x is nonwandering, there exists $n \geq 1$ such that

$$F^n(N^\gamma(x)) \cap N^\gamma(x) \neq \varnothing.$$

Pick $y, z \in N^\gamma(x)$ such that $F^n(y) = z$. Now consider the chain $\xi = (x, F(y), \ldots, F^{n-1}(y), x)$. Since $F(F^{n-1}(y)) = z \in N^\gamma(x)$, we have $A(\xi) = \rho(x, F(y)) + \rho(F^{n-1}(y), x) < \delta$. Taking $\delta \downarrow 0$ yields $x \sim_\rho x$ as claimed.   □

COROLLARY 4.10.   *Assume that μ is an F-invariant probability measure whose support S lies in M_1. Then $S \subset \mathcal{R}_\rho$.*

PROOF.   By the Poincaré recurrence theorem, S is included in the set

$$\overline{\{x \in M_1 : x \in \omega(x)\}} \subset \mathrm{NW}(F).$$

Applying Lemma 4.9 completes the result.   □

COROLLARY 4.11.   *Assume that $\mathcal{R}_\rho \cap M_1$ admits a neighborhood U, whose closure lies in some compact set $K \subset M_1$. Then there exists $N \in \mathbb{N}$ such that any partial solution $\zeta = (x, F(x), \ldots, F^n(x)) \in K^{n+1}$ with $n \geq N$ must pass through U.*

PROOF.   The set $K \setminus U$ does not contain any entire semiorbit of F, by Lemma 4.9. Since $A_n(\zeta) = 0$, applying Lemma 4.8(a) completes the proof.   □

We already stated that if $[x]_\rho$ is a closed ρ-basic class, then $[x]_\rho \subset [x]_{\mathrm{ap}}$. Under the finiteness assumption of Theorem 2.7, we derive the equality between ap and ρ-basic classes. Call $K_1, \ldots, K_v$ the ρ-basic classes in M_1 (recall that they are supposed to be closed), and label $K_1, \ldots, K_\ell$ the quasi-attractors among them. Proposition 4.6 implies that $K_1, \ldots, K_\ell$ are attractors. The following lemma implies that finiteness of the ρ-basic classes in M_1 implies finiteness of the *ap*-basic classes in M_1. In particular, Hypothesis 2.1 holds under the assumptions of Theorem 2.7.

THEOREM 4.12.   *Assume that there is a finite number of ρ-basic classes in M_1. Then $\mathcal{R}_\rho \cap M_1 = \mathcal{R}_{\mathrm{ap}} \cap M_1$ and $[x]_\rho = [x]_{\mathrm{ap}}$ for any $x \in \mathcal{R}_{\mathrm{ap}} \cap M_1$.*

PROOF.   Let $x \in \mathcal{R}_{\mathrm{ap}} \cap M_1$. We prove that $x \in \mathcal{R}_\rho \cap M_1$ and $[x]_{\mathrm{ap}} \subset [x]_\rho$. If $[x]_{\mathrm{ap}} = \{x\}$, then x is a fixed point and there is nothing to left to prove. Let $y \in [x]_{\mathrm{ap}}$, $y \neq x$ and $\alpha > 0$.

Remark 3.7 implies there exists a compact set $K \subset M_1$ with $\bigcup_i K_i \subset K$, a sequence $\delta_k \downarrow_k 0$ and a family $\xi^k = (\xi_0^k = x, \ldots, \xi_{n_k}^k = y) \in K^{n_k+1}$ of ap δ_k-pseudoorbits joining x to y.

Let $\gamma > 0$ be chosen so that $B_\rho(a, b) < \alpha$ for all $i = 1, \ldots, v$ and $a, b \in N^\gamma(K_i)$, the closure of $U = \bigcup_i N^\gamma(K_i)$ is contained in K, and $N^\gamma(K_i)$ is an isolating neighborhood for K_i for all i. Corollary 4.11 implies that there exists a positive integer N such that every partial solution $\{a, F(a), \ldots, F^n(a)\}$ in K of length $n \geq N$ must pass through U. Consequently, by compactness of K^N and continuity of F, we can find k_0 such that ξ^k cannot have more than N consecutive terms in $K \setminus U$ for $k \geq k_0$.

Now, given $k \in \mathbb{N}$, define $\sigma_0(k) = 0$ and $\tau_0(k) = \min\{j > 0 : \xi_j^k \notin U\}$. For $i \geq 1$, define inductively the terms $\sigma_i(k) = \min\{j > \tau_{i-1}(k) : \xi_j^k \in U\}$ and $\tau_i(k) = \min\{j > \sigma_i(k) : \xi_j^k \notin U\}$. This defines two sequences $\{\tau_i(k)\}_{i=0,\ldots,p_k}$ and $\{\sigma_i(k)\}_{i=0,\ldots,q_k}$. Notice that $q_k = p_k$ if $y \notin \bigcup_i K_i$ and $q_k = p_k + 1$ otherwise. By truncating multiple entries of ap pseudoorbits into each set $N^\gamma(K_i)$, we can assume that $q_k \leq v - 1$. After truncation, these pseudoorbits may only satisfy $d(F(\xi_j^k), \xi_{j+1}^k) \leq \delta_k$ for $\tau_i(k) - 1 \leq j \leq \sigma_{i+1}(k)$. Therefore

$$
\begin{aligned}
B_\rho(x, y) \leq \sum_{i=0}^{p_k} \big(& B_\rho\big(\xi_{\sigma_i(k)}^k, \xi_{\tau_i(k)-1}^k\big) + B_\rho\big(\xi_{\tau_i(k)-1}^k, \xi_{\tau_i(k)}^k\big) \\
& + B_\rho\big(\xi_{\tau_i(k)}^k, \xi_{\sigma_{i+1}(k)-1}^k\big) + B_\rho\big(\xi_{\sigma_{i+1}(k)-1}^k, \xi_{\sigma_{i+1}(k)}^k\big)\big) \\
& + B_\rho\big(\xi_{\tau_{q_k}(k)}^k, y\big)
\end{aligned}
$$

in the case where $y \in \bigcup_i K_i$, and

$$
\begin{aligned}
B_\rho(x, y) \leq \sum_{i=0}^{p_k} \big(& B_\rho\big(\xi_{\sigma_i(k)}^k, \xi_{\tau_i(k)-1}^k\big) + B_\rho\big(\xi_{\tau_i(k)-1}^k, \xi_{\tau_i(k)}^k\big)\big) \\
& + \sum_{i=0}^{p_k-1} \big(B_\rho\big(\xi_{\tau_i(k)}^k, \xi_{\sigma_{i+1}(k)-1}^k\big) + B_\rho\big(\xi_{\sigma_{i+1}(k)-1}^k, \xi_{\sigma_{i+1}(k)}^k\big)\big) \\
& + B_\rho\big(\xi_{\sigma_{p_k}(k)}^k, y\big)
\end{aligned}
$$

otherwise. In either case, our choice of γ implies

$$
B_\rho(x, y) \leq v\alpha + \big(v(N + 2) + 1\big) \sup\{\rho(a, b) : d\big(F(a), b\big) \leq \delta_k, a, b \in K\}
$$

for k sufficiently large. Uniform continuity of ρ on $K \times K$ implies that $\lim_{k \to \infty} \sup\{\rho(a, b) : d(F(a), b) \leq \delta_k, a, b \in K\} = 0$, and we obtain that $B_\rho(x, y) \leq v\alpha$. Since this holds for any $\alpha > 0$ we get that $x <_\rho y$. Similarly, $y <_\rho x$, which yields $x \sim_\rho y$. Therefore, $x \in \mathcal{R}_\rho$ and $[x]_\rho = [x]_{\mathrm{ap}}$. □

The next proposition shows that Hypothesis 2.2 holds.

PROPOSITION 4.13. *Let $j \in \{\ell + 1, \ldots, v\}$. We can find $\eta > 0$ such that, for any $\gamma > 0$, there exists $\varepsilon_0 > 0$ (which depends on η and γ) and a function ζ on $(0, \varepsilon_0)$ such that $\lim_{\varepsilon \to 0} \zeta(\varepsilon) = 0$ and*

$$\sup_{x \in N^{\eta}(K_j)} \mathbb{P}_x\big[\tau^{\varepsilon}_{N^{\eta}(K_j)} > e^{\gamma/\varepsilon}\big] \leq \zeta(\varepsilon)$$

for any $\varepsilon < \varepsilon_0$.

PROOF. First of all, by definition of a non ρ-quasiattractor, there exists $\eta > 0$ such that the closure of $N^{2\eta}(K_j)$ belongs to M_1, and for any $\gamma > 0$ and any $x \in N^{\eta}(K_j)$, there exists a sequence $\xi^{\gamma} = (\xi^{\gamma}_0, \ldots, \xi^{\gamma}_{n(\gamma)})$ such that

$$\xi^{\gamma} = x, \qquad \xi^{\gamma}_{n(\gamma)} \notin N^{2\eta}(K_j) \quad \text{and} \quad A_{n(\gamma)}(\xi^{\gamma}) < \gamma.$$

Call $U = N^{2\eta}(K_j)$. Since M_1 is invariant by F, $F(\overline{U})$ is compact and contained in M_1. Hence there exists $r > 0$ and a compact set $K \subset M_1$ such that

$$N^r(F(\overline{U})) \subset K.$$

By continuity of ρ on $\overline{U} \times M$ and since ρ is strictly positive on the compact set $\overline{U} \times (N^r(F(\overline{U})))^c$, there exists $\gamma_0 > 0$ such that

$$\rho(x, y) > \gamma_0 \qquad \text{for all } x \in \overline{U}, y \in K^c.$$

In particular, this means that, for $\gamma < \gamma_0$, the sequence ξ^{γ} must pass through $K \setminus U$, and we can therefore assume without loss of generality that $\xi^{\gamma}_{n(\gamma)} \in K \setminus U$ and ξ^{γ} lives in K.

Pick $\delta > 0$. We now apply Theorem 4.7 in the compact set K, with δ, η and $N = n(\gamma)$: there exists $\varepsilon_0 > 0$ [which depends on η, δ and $n(\gamma)$] such that, for any $\varepsilon < \varepsilon_0$,

$$\mathbb{P}_x\big[d_{n(\gamma)}(X^{\varepsilon}, \xi^{\gamma}) < \eta\big] \geq \exp\Big(-\frac{\gamma + \delta}{\varepsilon}\Big).$$

Consequently there exists $\varepsilon'_0 > 0$ such that, for any $0 < \varepsilon < \varepsilon'_0$ [up to changing slightly $n(\gamma)$], we have

$$\mathbb{P}_x\big[\tau^{\varepsilon}_{N^{\eta}(K_j)} \leq n(\gamma)\big] > e^{-\gamma/\varepsilon}.$$

Consequently,

$$\mathbb{P}_x\big[\tau^{\varepsilon}_{N^{\eta}(K_j)} \geq e^{2\gamma/\varepsilon}\big] \leq \big(1 - e^{-\gamma/\varepsilon}\big)^{[e^{2\gamma/\varepsilon}/n(\gamma)]}.$$

The last quantity is of order $\exp(-e^{\gamma/\varepsilon}/(2n(\gamma)))$ and therefore goes to zero. $\square$

By assumption of Theorem 2.7, there is at least one ρ-quasiattractor (which turns out to be an attractor by Proposition 4.6) among the ρ-basic classes included

in M_1, such that $\mu_n(U) > 0$ for all n, for all open neighborhoods U. Hence, by Lemma 3.9 there exists $\delta_0 > 0$ such that $\lambda_n > 1 - \beta_{\delta_0}(\varepsilon_n) > 1 - e^{-c_0/\varepsilon_n}$, where

$$c_0 = \tfrac{1}{2}\inf\{\rho(x, y): (x, y) \in M_1 \times M, d(F(x), y) \geq \delta_0\} > 0.$$

Let $\delta_1 < \delta_0$ and $\{V_i\}_{i=1,\dots,v}$ be chosen so that Corollary 3.8 holds.

We are now ready to prove Theorem 2.7. Since Hypotheses 2.1 and 2.3 are satisfied, it remains to verify Hypothesis 2.2. Choose the neighborhoods $\{V_i\}_{i=\ell+1,\dots,v}$ such that $V_i \subset N^\eta(K_i)$, where η is given by Proposition 4.13. Choose $\gamma < \frac{c_1}{2}$ where

$$0 < c_1 = \inf\{\rho(x, y): (x, y) \in M_1 \times M, d(F(x), y) \geq \delta_1\} \leq 2c_0.$$

Define $h(\varepsilon) = e^{\gamma/\varepsilon}$. Proposition 4.13 implies that there exists $\varepsilon_0 > 0$ and a function ζ such that $\lim_{\varepsilon \to 0} \zeta(\varepsilon) = 0$ and

$$\sup_{x \in N^\eta(K_j)} \mathbb{P}_x\big[\tau^{\varepsilon_n}_{N^\eta(K_j)} > h(\varepsilon_n)\big] \leq \zeta(\varepsilon_n) \to 0$$

for any $\varepsilon < \varepsilon_0$. Since $\lim_{n \to \infty} h(\varepsilon_n)\beta_{\delta_1}(\varepsilon_n) = 0$, Hypothesis 2.2 holds. This completes the proof of Theorem 2.7.

5. Applications. Our results are broadly applicable to many Markov chain models in population biology. To give some flavor of this applicability, we introduce two classes of Markov chains satisfying our probabilistic assumptions and some illustrative applications to metapopulation dynamics, competing species, host-parasitoid interactions and evolutionary games. For each application there are two ingredients for verifying the conditions of Theorem 2.7. The probabilistic ingredient involves verifying that there exist quasi-stationary distributions and verifying the large deviation assumptions. We defer verifying these conditions until Section 6. The dynamical ingredient involves verifying there is a finite number of ρ-basic classes and identifying the attractors. For the second ingredient, we introduce a proposition, that is, applicable to most of our examples.

5.1. *A dynamical proposition.* To state the proposition, we need a few definitions from dynamical systems. For $x \in M$, let $\omega(x) = \{y$: there exists $n_k \to \infty$ such that $\lim_{k \to \infty} F^{n_k}(x) = y\}$ be the *ω-limit set for x* and $\alpha(x) = \{y$: there exist $n_k \to \infty$ and $y_k \in M$ such that $F^{n_k}(y_k) = x$ and $\lim_{k \to \infty} y_k = y\}$ be the *α-limit set for x*. Our assumption that F is bounded implies that there exists a global attractor given by the compact, F-invariant set $\Lambda = \bigcap_{n \geq 0} F^n(M)$. For all $x \in \Lambda$, $\omega(x)$ and $\alpha(x)$ are compact, nonempty, F-invariant sets.

A *Morse decomposition* of the dynamics of F is a collection of F-invariant, compact sets $K_1, \dots, K_k$ such that:

- K_i is isolated, that is, there exists a neighborhood of K_i such that it is the maximal F-invariant set in the neighborhood, and

- for every $x \in \Lambda \setminus \bigcup K_i$, there exist $i > j$ such that $\omega(x) \subset K_i$ and $\alpha(x) \subset K_j$.

Modulo replacing the invariant sets K_i by points, one can think of F being gradient-like as all orbits move from lower indexed invariant sets to higher indexed invariant sets.

PROPOSITION 5.1. *Assume Hypothesis 2.5 holds. If F admits a Morse decomposition $K_1, \ldots, K_k$ such that*:

- $K_i \subset M_1$ or $K_i \subset M_0$ for each i, and
- K_i is transitive whenever $K_i \subset M_1$, that is, there exists $x \in K_i$ such that $\{x, F(x), F^2(x), \ldots\}$ is dense in K_i,

then ρ-basic classes in M_1 are given by the $K_i \subset M_1$. In particular, there is a finite number of ρ-basic classes in M_1, and each of them is closed.

PROOF. Let $K_1, \ldots, K_k$ be a Morse decomposition for F. Let $I \subset \{1, \ldots, k\}$ be such that $K_i \subset M_1$ if and only $i \in I$. By assumption, $K_i \subset M_0$ for $i \notin I$, and K_i is transitive for $i \in I$. Transitivity of K_i for $i \in I$ and continuity of ρ restricted to $M_1 \times M$ implies that K_i is contained in a ρ-basic class for $i \in I$, that is, $x \sim_\rho y$ for all $x, y \in K_i$. As shown in Section 4, assertion (iii) of Hypothesis 2.5 implies that the ρ-basic classes are contained in the ap-chain recurrent set of F which is contained in $\bigcup_i K_i$. Hence, the ρ-basic classes in M_1 are given by $\{K_i\}_{i \in I}$. □

5.2. *Nonlinear Poisson branching processes.* To describe structured populations with k types of individuals (e.g., different genotypes or species, individuals living in different spatial locations), let x represent the vector of population densities which lies in the nonnegative cone $\mathbb{R}_+^k$ of $\mathbb{R}^k$. A widely used class of models in population biology [Caswell (2001)] is the nonlinear matrix model of the form $F(x) = A(x)x$ where $A(x)$ are nonnegative matrices representing transitions births, deaths and transitions between types of individuals (e.g., due to mutation or dispersal).

Since real populations involve finite numbers of individuals, these deterministic models can be viewed as approximations of more realistic, stochastic representations of the population dynamics. More specifically, let $1/\varepsilon$ that represents the "size" (e.g., area, volume) of the habitat. Let $N_t^\varepsilon \in \mathbb{Z}_+^k$ denote the vector of population abundances where $\mathbb{Z}_+^k$ is the nonnegative cone of the k-dimensional integer lattice. Then $X_t^\varepsilon = \varepsilon N_t^\varepsilon \in \varepsilon \mathbb{Z}_+^k$ is the vector of population densities. For every $x \in \mathbb{R}_+^k$, let $Z_1(x), Z_2(x), \ldots$ be a sequence of i.i.d. random vectors with independent components, and whose i component has a Poisson distribution with mean x_i. Given $N_0^\varepsilon \in \mathbb{Z}_+^k$, we can define the Markov chains $\{X_t^\varepsilon\}$ iteratively by

$$N_{t+1}^\varepsilon = Z_{t+1}(A(X_t^\varepsilon)N_t^\varepsilon) \quad \text{and} \quad X_{t+1}^\varepsilon = \varepsilon N_{t+1}^\varepsilon.$$

Equivalently, we can write $X_{t+1}^\varepsilon = \varepsilon Z_{t+1}(F(X_t^\varepsilon)/\varepsilon)$.

A useful observation about these Poisson processes, from the modeling stand-point, is that multinomial sampling of a Poisson process still corresponds to a Poisson process. More specifically, consider a multinomial random vector $(X_1, \ldots, X_k)$ where the number of samples N is Poisson distributed with mean $\lambda > 0$ and the sampling probabilities are $(p_1, \ldots, p_k)$. Then

$$
\begin{aligned}
\mathbb{P}[X_1 &= x_1, \ldots, X_k = x_k] \\
&= \mathbb{P}[X_1 = x_1, \ldots, X_k = x_k \mid N = x_1 + \cdots + x_k]P[N = x_1 + \cdots + x_k] \\
&= \frac{(x_1 + \cdots + x_k)!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k} \exp(-\lambda) \frac{\lambda^{x_1 + \cdots + x_k}}{(x_1 + \cdots + x_k)!} \\
&= \prod_{i=1}^{k} \frac{(p_i \lambda)^{x_i}}{x_i!} \exp(-p_i \lambda).
\end{aligned}
$$

Hence, $X_1, \ldots, X_k$ are independent Poisson random variables with rate parameters $p_1\lambda, \ldots, p_k\lambda$. We make repeated use of this observation in the examples provided below.

Since F is bounded, one can show quite generally that these Markov chains support quasi-stationary distributions whenever there are absorbing sets. For all of our examples, these absorbing sets are $\{0\}$ or $\partial\mathbb{R}_+^k = \{x \in \mathbb{R}_+^k : \prod x_i = 0\}$. A proof of this assertion is given in Proposition 6.1 of Section 6. Under slightly stronger assumptions (namely A is continuous and F_i is strictly positive), we show in Proposition 6.3 of Section 6 that these Poisson processes also satisfy our large deviation Hypotheses 2.5 and 2.6.

To provide a taste of the possible applications, we apply our results to three particular classes of nonlinear Poisson branching processes.

Metapopulation dynamics.   A fundamental question in population biology is how do local demographic processes, such as reproduction and survivorship, inter-act with dispersal (a regional demographic process) to determine spatial-temporal patterns of abundance [Earn, Levin and Rohani (2000), Earn and Levin (2006), Hastings and Botsford (2006), Schreiber (2010)]. This issue has been studied extensively with discrete-time deterministic models representing space as a finite collection of patches connected by dispersal. To illustrate how our results apply to these metapopulation models, we introduce a stochastic version of the spatial Ricker map, which was originally studied by Hastings (1993) for 2 patches, and for which we allow an arbitrary number, k, of patches.

Let $1/\varepsilon > 0$ be the area or volume of a single patch, $N_t^{\varepsilon,i}$ denote the number of individuals in patch i, $N_t^\varepsilon = (N_t^{\varepsilon,1}, \ldots, N_t^{\varepsilon,k})$ the vector of population abundances across space and $X_t^\varepsilon = \varepsilon N_t^\varepsilon$ the vector of population densities. To describe reproduction within a patch, let $f(x) = f_0 \exp(-x)$ be the mean fecundity of an individual when the local population density is x and the "intrinsic" fitness is $f_0 > 0$.

The map $x \mapsto x f(x)$ is known as the Ricker map in theoretical ecology and is commonly used to describe the population dynamics of a single species [Hastings (1997), Ricker (1954), Wysham and Hastings (2008)]. Let $D = (d_{ij})$ be an irreducible, row-stochastic matrix where d_{ij} corresponds to the probability of an individual dispersing from patch i to patch j. Given N_t^ε, we define the *spatial Ricker process* as follows:

- Each individual in patch i independently produces a Poisson-distributed number of offspring with mean $f(X_t^{\varepsilon,i})$ to replace themselves. Let $Z_{t+1}^{i,\varepsilon}$ be the total number of offspring produced in patch i, which is Poisson distributed with mean $N_t^{\varepsilon,i} f(X_t^{\varepsilon,i})$. We assume that the $Z_{t+1}^{1,\varepsilon}, \ldots, Z_{t+1}^{k,\varepsilon}$ are independent; there are no correlations in the reproductive output between distinct patches.
- Independent of one another, offspring in patch i move to patch j with probability d_{ij}. To represent this movement, let $W_{t+1}^\varepsilon(i) = (W_{t+1}^{\varepsilon,1}(i), \ldots, W_{t+1}^{\varepsilon,k}(i))$ be a multinomial random vector with sampling probabilities $d_{i1}, \ldots, d_{ik}$ and $Z_{t+1}^{\varepsilon,i}$ trials.
- Define
$$N_{t+1}^\varepsilon = \sum_i W_{t+1}^\varepsilon(i) \quad \text{and} \quad X_{t+1}^\varepsilon = \varepsilon N_{t+1}^\varepsilon.$$

By our earlier observation about multinomial sampling of a Poisson random variable, $\{X_t^\varepsilon\}$ is a nonlinear Poisson process with
$$F_i(x) = \sum_j d_{ji} x_j f(x_j).$$

Since $\|F\| \le f_0 \|D\|$ and D is irreducible, Proposition 6.1 implies the process $\{X_t^\varepsilon\}$ has a quasi-stationary distribution μ_ε with respect to the absorbing state $M_0 = \{0\}$ for all $\varepsilon > 0$.

Let μ be a weak* limit point of μ_ε as $\varepsilon \to 0$. To say something about the support of μ, we need to understand the dynamics of the map $F(x)$. The simplest applicable result is a persistence and extinction dichotomy. Since $F_i(x) \le f_0 x_i$, it follows that 0 is a global attractor for $F(x)$ whenever $f_0 < 1$. Alternatively, when $f_0 > 1$, a result of Kon, Saito and Takeuchi [(2004), Theorem 3] implies that $F(x)$ has a positive attractor. Lemma 3.9(b$'$) and Theorem 3.12 imply the following result.

PROPOSITION 5.2. *Let μ be a weak* limit point for quasi-stationary distributions μ_ε of the spatial Ricker process $\{X_t^\varepsilon\}$. Then*:

Extinction. *If $f_0 < 1$, then $\mu(\{0\}) = 1$.*
Metastability. *If $f_0 > 1$, then there exists a $\delta > 0$ such that $\mu(N^\delta(\{0\})) = 0$.*

In the limiting cases where the population is either weakly mixed or well mixed, we can say more about the support of the limiting measure μ. These stronger

assertions rely on the one-dimensional map $x \mapsto f_0 x \exp(-x)$ having a linearly stable periodic orbit, call it $\mathcal{S} = \{p, F(p), \ldots, F^{n-1}(p)\}$ where n is the period. Kozlovski [(2003), Theorem C] proved that, for an open and dense set of f_0 values, such a stable periodic orbit exists. Hence, this assumption is not very restrictive.

THEOREM 5.3. *Assume the one-dimensional map* $x \mapsto f_0 x \exp(-x)$ *has a linearly stable periodic orbit, call it* $\mathcal{S} = \{p, F(p), \ldots, F^{n-1}(p)\}$, *and D is an irreducible, nonnegative matrix whose row sums equal one (i.e., a row stochastic matrix).*

Weakly mixing. *If D is sufficiently close to the identity matrix, then there exists* n_k *linearly stable periodic orbits for* $F(x)$, *and* μ *is supported by the union of these stable periodic orbits.*

Strongly mixing. *If all the entries of D are sufficiently close to* $1/k$ *and the column sums of D equal one (i.e., D is doubly stochastic), then there exists a unique globally stable periodic orbit for* $F(x)$, *and the support of* μ *is given by this periodic orbit.*

We remark that in the special case of a single patch, $k = 1$, we recover results of Högnäs (1997), Klebaner, Lazar and Zeitouni (1998) and Ramanan and Zeitouni (1999) for one-dimensional maps on a compact interval. See Section 3.5 for further discussion about this point.

PROOF OF THEOREM 5.3. To prove the first assertion, consider the uncoupled map

$$\widetilde{F}(x) = \big(x_1 f(x_1), x_2 f(x_2), \ldots, x_k f(x_k) \big).$$

Each of the components of this limiting map are given by the one-dimensional map $g(x_i) = x_i f(x_i)$ which by assumption has a linearly stable periodic orbit $\mathcal{S} = \{p, g(p), \ldots, g^{n-1}(p)\}$. This linearly stable periodic orbit gives rise to n^k periodic orbits of the form $(g^{n_1}(p), \ldots, g^{n_k}(p))$ with $0 \le n_j < n$ for $\widetilde{F}$. Since g has a negative Schwartzian derivative and a single critical point, van Strien [(1981), Theorem A] proved that the complement of the basin of attraction of $\mathcal{S}$ for g can be decomposed into a finite number of compact, g-invariant sets which have a dense orbit and are hyperbolic repellers: there exists $C > 0$ and $\lambda > 1$ such that $|(g^n)'(x)| \ge C\lambda^n$ for all points x in the set and $n \ge 1$. Consequently, the k-dimensional mapping $\widetilde{F}$ is an Axiom A endomorphismn [Przytycki (1976), page 271]: the derivative of $\widetilde{F}$ is nonsingular for all points in the nonwandering set $\Omega(\widetilde{F}) = \{x \in \mathbb{R}_+^k : \text{for every neighborhood } U \text{ of } x, \widetilde{F}^n(U) \cap U \neq \varnothing$ for some $n\}$, $\Omega(\widetilde{F})$ is a hyperbolic set and the periodic points are dense in $\Omega(\widetilde{F})$. Results of Przytycki [(1976), 3.11–3.14 and 3.17] imply that key attributes of Axiom A endomorphism are: (i) $\Omega(\widetilde{F})$ decomposes in a finite number of invariant sets $\Omega^1(\widetilde{F}), \ldots, \Omega^m(\widetilde{F})$, (ii) for each orbit $\{x_n\} \subset \Omega^i(\widetilde{F})$ of $\widetilde{F}$, the unstable manifold

at x_0 intersects $\Omega^i(\widetilde{F})$ in a dense set and (iii) maps F sufficiently C^1 close to $\widetilde{F}$ are Axiom A endomorphisms. Property (iii) implies that $F(x) = D\widetilde{F}(x)$ is an Axiom A endomorphism provided that D is sufficiently close to the ident ity matrix. Property (ii) implies that each of the invariant sets $\Omega^i(F)$ is a ρ-equivalence class. Linear stability of the n^k periodic orbits $(g^{n_1}(p), \ldots, g^{n_k}(p))$ with $0 \le n_j < n$ for $\widetilde{F}$ implies that, for sufficiently small perturbations F of $\widetilde{F}$, n^k of the invariant sets $\Omega^i(F)$ correspond to linearly stable periodic points, while the remaining invariant sets are either hyperbolic repellers or saddles. Since the stable periodic orbits are the only ρ-quasi-attractors, Theorem 2.7 implies the first assertion of the proof.

We prove the second assertion. Since D is doubly stochastic, the nonnegative half-line $\mathcal{L} = \{x : x_1 = \cdots = x_k \ge 0\}$ is F-invariant, that is, $F(x_1 \mathbf{1}) = g(x_1)\mathbf{1}$ where $\mathbf{1}$ is the vector of ones. As in the case of the proof of the first assertion, the dynamics of F of restricted $\mathcal{L}$ has the stable periodic point $(p\mathbf{1}, g(p)\mathbf{1}, \ldots, g^{n-1}(p)\mathbf{1})$, and the complement of its basin of attraction can be decomposed into a finite number, say m, of compact, g-invariant sets which have a dense orbit and are hyperbolic repellers. Define $\widetilde{D}$ by $\widetilde{d}_{ij} = d_{ij} - d_{ik}$ for all i, j. By choosing d_{ij} sufficiently close to $1/k$ for all i, j, we can make the matrix $\widetilde{D}$ as close to zero as we want. Hence, Earn and Levin [(2006), Theorem 1] implies that $\mathcal{L}$ is a global attractor for the dynamics of F. Moreover, the stable periodic orbit for F restricted to $\mathcal{L}$ is stable for F. Proposition 5.1 implies that each of these invariant sets is a ρ-equivalence class. Since the stable periodic orbit is the only ρ-quasi-attractor, Theorem 2.7 implies the second assertion. $\square$

Competing species. During the mid twentieth century, laboratory experiments played a key role in establishing the competitive exclusion principle in ecology. One classic set of competition experiments was conducted by Park (1948, 1954) with flour beetles. To model the dynamics of these competing beetles, collaborators of Park [Leslie and Gower (1958)] used difference equations, rather than the classical Lotka–Volterra differential equation model of competition. Cushing et al. (2004) showed that these difference equations exhibit the same dynamical outcomes as the Lotka–Volterra models. Namely, one or both species may go extinct for all initial conditions, may coexist about a globally stable equilibrium or may exhibit contingent exclusion where the initially "more abundant" species excludes the other species. Here, we consider a stochastic counterpart of the Leslie–Gower model.

Let $N_t^\varepsilon = (N_t^{\varepsilon,1}, N_t^{\varepsilon,2})$ and $X_t^\varepsilon = \varepsilon N_t^\varepsilon$ denote the abundances and densities of the competing species at time t. Once again, $1/\varepsilon$ corresponds to the volume of their habitat. The per-capita mean fecundity $f_i(x)$ for species i is given by $f_i(x) = \frac{b_i}{1+c_{ii}x_i+c_{ij}x_j}$ where $j \ne i$, $b_i > 0$ is the "intrinsic" birth rate, $c_{ii} > 0$ is the strength of intraspecific competition and $c_{ij} > 0$ is the strength of interspecific competition. If individual births are independent given the current density of

individuals and Poisson distributed with means $f_i(X_t^\varepsilon)$ $i = 1, 2$, then $N_t^{\varepsilon,i}$ is a non-linear Poisson process associated with the map $F(x) = (f_1(x)x_1, f_2(x)x_2)$. Proposition 6.1 implies the Leslie–Gower process has quasi-stationary distributions μ_ε for $\varepsilon > 0$ with $M = \mathbb{R}_+^2 = \{x \in \mathbb{R}^2 : x_i \geq 0\}$ and $M_0 = \partial\mathbb{R}_+^2 = \{x \in \mathbb{R}_+^2 : x_1 x_2 = 0\}$. Results of Cushing et al. [(2004), Theorem 4], our Theorems 3.12 and 2.7 imply the following result.

THEOREM 5.4. *Let μ_ε be a quasi-stationary distribution for the Leslie–Gower process $\{X_t^\varepsilon\}$. Let μ a weak* limit point of these quasi-stationary distributions.*

Coexistence. *If $b_i > 1$ for $i = 1, 2$ and $c_i(b_j - 1) < b_i - 1$ for $i = 1, 2$ and $i \neq j$, then μ is a Dirac measure supported by the point*

$$\left(\frac{b_2 - 1}{c_1 c_2 - 1}\left(c_1 - \frac{b_1 - 1}{b_2 - 1}\right), \frac{b_1 - 1}{c_1 c_2 - 1}\left(c_2 - \frac{b_2 - 1}{b_1 - 1}\right) \right).$$

Extinction or exclusion. *If $b_i < 1$ for some i, or $b_i > 1$ for $i = 1, 2$, $b_2 - 1 > (b_1 - 1)/c_1$ and $b_1 - 1 > (b_2 - 1)/c_2$, or $b_i > 1$ for $i = 1, 2$, $b_2 - 1 < (b_1 - 1)/c_1$ and $b_1 - 1 < (b_2 - 1)/c_2$, then μ is supported by $\partial\mathbb{R}_+^2$.*

The case for which our results are not conclusive is when the dynamics of the Leslie–Gower model are bistable [i.e., $b_i > 1$ for $i = 1, 2$ and $c_i(b_j - 1) > b_i - 1$ for $i = 1, 2$ and $i \neq j$] in which case there is a positive unstable equilibrium and all initial conditions not lying on its stable manifold (which has dimension one) go to $\partial\mathbb{R}_+^2$. However, we conjecture that μ is supported on the boundary of the positive quadrant in this case.

Host-parasitoid interactions. Predator-prey interactions involve one species benefiting by harming another species. These interactions are the fundamental building blocks for all food webs. An important class of predators is parasitoids such as wasps or flies whose young develop in and ultimately kill their host [Godfray (1994)]. Mathematical models of these interactions have been studied for almost a century [Thompson (1924), Nicholson and Bailey (1935), Hassell (1978, 2000), May (1995), Schreiber (2006a, 2007), Gidea et al. (2011)]. As predator-prey interactions are inherently oscillatory, these studies often focused on identifying mechanisms that stabilize predator-prey interactions.

Here, we introduce a stochastic analog of these deterministic models. Let H_t^ε and P_t^ε denote the abundances of host and the parasitoid in generation t, respectively. Let $X_t^\varepsilon = \varepsilon N_t^\varepsilon = \varepsilon(H_t^\varepsilon, P_t^\varepsilon)$ be their densities where $1/\varepsilon$ is the size of the environment. Let $f(X_t^{\varepsilon,1})$ be the mean number of offspring produced by an individual host. Let $g(X_t^\varepsilon)$ be the probability that an offspring escapes parasitism from the parasitoids. We update the population state X_t^ε according to the following rules:

- Each adult host independently produces a Poisson distributed number of offspring with mean $f(X_t^{\varepsilon,1})$. Let M_{t+1} be the total number of offspring which is Poisson distributed with mean $H_t^{\varepsilon} f(X_t^{\varepsilon,1})$.
- Each offspring survives parasitism independently with probability $g(X_t^{\varepsilon})$. Let H_{t+1}^{ε} equal the number of surviving offspring and $P_{t+1}^{\varepsilon} = M_{t+1} - H_{t+1}^{\varepsilon}$ be the number of parasitized offspring which all emerge as parasitoids in the next generation.

Since $(H_{t+1}^{\varepsilon}, P_{t+1}^{\varepsilon})$ is binomial distributed with M_{t+1} trials, $(H_{t+1}^{\varepsilon}, P_{t+1}^{\varepsilon})$ are independent Poisson random variables. Hence, X_t^{ε} is a nonlinear Poisson process associated with the map

$$F(x) = \left(f(x_1)x_1 g(x), f(x_1)x_1(1 - g(x)) \right)$$

on $\mathbb{R}_+^2 = \{(x_1, x_2) \in \mathbb{R}^2 : x_i \geq 0\}$. Provided that F is continuous, and f is a compact map, Proposition 6.1 implies that there is a quasi-stationary distribution μ_{ε} for X_t^{ε} with $\varepsilon > 0$.

To understand the support of the weak* limit points μ of μ_{ε}, we focus on a generalized Thompson model [Thompson (1924), Getz and Mills (1996), Schreiber (2006a, 2007)]. For this model, $f(x_1) = \exp(r(1 - x_1/K))$ is given by the Ricker equation where $r > 0$ is the intrinsic rate of growth of the host, and $K > 0$ is the host's carrying capacity. The escape function $g(x) = (1 + x_2/(bx_1k))^{-k}$ corresponds to a negative binomial escape function with egg-limited encounter rates. Here, $b > 0$ is the attack rate of the parasitoid and $1/k > 0$ represents how "clumped" or "aggregated" parasitoid attack are; that is, smaller k correspond to greater aggregation of parasitoid attacks. Notice that while g is not defined at $x_1 = 0$, the map F extends continuously to $x_1 = 0$ if we set $F(x) = 0$ whenever $x_1 = 0$. Combining results from Schreiber [(2007), Theorem 3.1, 3.2] and Theorem 2.7 yields the following results for $k < 1$, that is, parasitoid attacks are sufficiently aggregated (Hassell et al. (1991)).

THEOREM 5.5. *Let μ be a weak* limit point of the quasi-stationary distributions for the Thompson host-parasitoid process X_t^{ε}. Assume $k < 1$, and define*

$$y^* = \max\left\{y \geq 0 : \exp(-r)\left((1 + y/(bk))^k - 1\right) = y\right\}.$$

Then:

Extinction. *If $\exp(r)(1 + y^*/(bk))^{-k} < 1$, then μ is supported by the $\partial\mathbb{R}_+^2 = \{x \in \mathbb{R}_+^2 : x_1 x_2 = 0\}$.*
Coexistence. *If $\exp(r)(1 + y^*/(bk))^{-k} > 1$, then μ is supported by $\mathbb{R}_+^2 \setminus \partial\mathbb{R}_+^2$. Moreover, for an open and dense set of parameter values (r, b) satisfying $\exp(r)(1 + y^*/(bk))^{-k} > 1$, μ is supported by a periodic orbit.*

When $k \geq 1$ (i.e., parasitoid attacks are not sufficiently aggregated), Schreiber [(2007), Theorem 3.1] implies coexistence does not occur for the deterministic model. However, this extinction often involves unstable sets in the interior of $\mathbb{R}_+^2$. Consequently, our results are not applicable. Nonetheless, we conjecture that μ is supported by $\partial \mathbb{R}_+^2$.

5.3. *Multinomial processes.* Consider a landscape with N sites that can be in one of k states. These states may correspond to occupation by individuals playing different strategies in the context of evolutionary games, or different genotypes in the context of population genetics. Let $M = \{x \in \mathbb{R}^k : x_i \geq 0, \sum_{i=1}^k x_i = 1\}$ be the k-simplex and $F : M \to M$ be a continuous map.

For each $x \in M$, let $Z_1(x), Z_2(x), \ldots$ be a sequence of independent random vectors with a multinomial distribution with N trials, k possible outcomes and probability x_i of producing type i in a single trial. If $\varepsilon = 1/N$ and $X_0^\varepsilon \in M \cap \varepsilon \mathbb{Z}^k$ is given, then we can define a Markov chain $\{X_t^\varepsilon\}_{t=0}^\infty$ on $M \cap \varepsilon \mathbb{Z}^k$ iteratively by

$$X_{t+1}^\varepsilon = \varepsilon Z_{t+1}\big(F(X_t^\varepsilon)\big).$$

Since $\{X_t^\varepsilon\}$ is a finite-state Markov chain, quasi-stationary distributions exist uniquely whenever the transition matrix restricted to the transient states is aperiodic and irreducible [Darroch and Seneta (1965)]. When M_0 is the boundary of the simplex and $F_i(x) = x_i f_i(x)$, with f_i continuous and positive, we therefore always have a unique quasi-stationary distribution, and we prove in Proposition 6.5 of Section 6 that the large deviation hypotheses are satisfied. As a particular application of these multinomial processes, we consider evolutionary games.

Evolutionary game dynamics. Evolutionary game theory studies the dynamics of populations of players, each programmed to play a fixed strategy throughout their life time [Hofbauer and Sigmund (1998, 2003), Cressman (2003)]. These populations often exhibit frequency dependent selection; the reproductive success of a player changes in time due to the composition of strategies in the population. The study of evolutionary games has led to fundamental insights into the evolution of animal conflicts [Maynard Smith (1974)], cooperation [Imhof and Nowak (2010), Nowak et al. (2004)], habitat selection [Cressman, Krivan and Garay (2004), Schreiber, Fox and Getz (2000)] and mating systems [Sinervo and Lively (1996)].

A basic deterministic model for evolutionary games is the discrete-time replicator equation [Hofbauer and Sigmund (2003)],

$$F_i(x) = x_i \frac{\sum_j a_{ij} x_j + c}{\sum_{jl} a_{jl} x_j x_l + c},$$

where x_i is the frequency of strategy i in the population, the entries a_{ij} of the "pay-off" matrix A describe the fitness gain to strategy i when interacting with

strategy j, and c is the "basal" fitness of an individual. The dynamics of this discrete system have been studied extensively [Hofbauer and Sigmund (1998, 2003)]. Here, we describe a stochastic analog of these games that account for finite population sizes. In the case of two-strategies, this stochastic analog corresponds to the frequency dependent Wright–Fisher processes studied by Imhof and Nowak (2006).

Let $N_t^{\varepsilon,i}$ and $X_t^{\varepsilon,i}$ denote the abundance and frequency of the ith strategy at time t. Here, $1/\varepsilon$ is assumed to be an integer that corresponds to the total population size which does not change over time. If we update N_t^ε by taking a multinomial random variable with $1/\varepsilon$ trials and probabilities $F(X_t^\varepsilon)$, then we get a multinomial process. One can interpret this process as individuals producing many offspring proportional to their fitness $\sum_j a_{ij}x_j + c$, and randomly selecting individuals from the offspring "pool" to replace their parents. For this stochastic process $\{X_t^\varepsilon\}$, $M_0 = \{x \in M : \prod x_i = 0\}$ is an absorbing set that corresponds to the loss of one or more strategies.

We can leverage two results from the theory of replicator dynamics to describe the support of the quasi-stationary distributions μ_ε when $\varepsilon > 0$ is sufficiently small, and the basal payoff c is sufficiently large. To first order in $1/c$,

$$F_i(x) \approx x_i + x_i \frac{1}{c}\left(\sum_j a_{ij}x_j - \sum_{jl} a_{jl}x_jx_l\right)$$

and the dynamics of the map $x \mapsto F(x)$ can be viewed as a Cauchy–Euler approximation to the classical continuous time replicator equations

$$(11) \qquad \frac{dx_i}{dt} = x_i\left(\sum_j a_{ij}x_j - \sum_{jl} a_{jl}x_jx_l\right).$$

Using this observation and work of Garay and Hofbauer (2003), Hofbauer and Sigmund (1998), we can prove a sufficient condition for the stochastic replicator processes exhibiting metastable persistence for $\varepsilon > 0$ small and $c > 0$ large.

THEOREM 5.6. *Let μ be a weak* limit point of the quasi-stationary distributions μ^ε for the replicator process with $c > 0$ sufficiently large. If there exists $p_i > 0$ such that*

$$\sum_i p_i\left(\sum_j a_{ij}x_j - \sum_{jl} a_{jl}x_jx_l\right) > 0$$

at every equilibrium $x \in M_0$ for F, then μ is supported by M_1.

PROOF. Hofbauer and Sigmund [(1998), Theorem 13.6.1] implies that the continuous-time replicator equations (11) have a global attractor $A \subset M_1$ whenever there positive weights $p_1, \ldots, p_k$ such that

$$\sum_i p_i\left(\sum_j a_{ij}x_j - \sum_{jl} a_{jl}x_jx_l\right) > 0$$

at every equilibrium $x \in M_0$ for F. In particular, in the terminology of Garay and Hofbauer [(2003), Definition 2.1], (11) admits a good average Lyapunov function whenever these positive weights exist. Consequently, Garay and Hofbauer [(2003), Theorem 8.3] implies that the discrete-time replicator equation F admits a global attractor $A \subset M_1$ whenever c is sufficiently large. This global attractor, however, need not be a unique ρ-quasi attractor. However, as suggested by Comment 3.8, we can apply Lemma 3.9(b′) to conclude that $\mu(V_0) = 0$ for some neighborhood V_0 of M_0. $\square$

As an interesting special case consider the rock-paper-scissor game where the payoff-matrix is of the form

$$A = \begin{pmatrix} 0 & -a_2 & b_3 \\ b_1 & 0 & -a_3 \\ -a_1 & b_2 & 0 \end{pmatrix}$$

with a_i and b_i positive. Zeeman (1980) proved that if $\det(A) > 0$, then the persistence condition of Theorem 5.6 is satisfied. Moreover, for the continuous-time replicator equations, there is a globally stable internal equilibrium. For $c > 0$ sufficiently large, this equilibrium is also globally stable for the map F and, consequently μ is a Dirac measure supported by this equilibrium. When $\det(A) < 0$, Zeeman proved that the internal equilibrium is unstable and all other orbits of the continuous-time deterministic system approach the boundary and one can show that the same conclusion holds for the discrete-time system when $c > 0$ is sufficiently large. Since the boundary is not a global attractor in this case, we cannot apply Theorem 3.12. None the less, we conjecture that μ is supported on the boundary M_0 in this case.

6. Large deviation results for Poisson and multinomial models. In this section we prove the existence of quasi-stationary distributions (as needed) and verify our large deviation hypotheses for the Poisson and multinomial models introduced in Section 5.

6.1. *Nonlinear Poisson branching model.* We first prove the existence of quasi-stationary distributions for the nonlinear Poisson processes introduced in Section 5.2.

PROPOSITION 6.1. *If* $\sup_{x \in \mathbb{R}_+^k} \|F(x)\| < \infty$ *and* $F_i(x) > 0$ *for all* $x \in M \setminus M_0$ *and* i, *then the nonlinear Poisson process* $\{X_t^\varepsilon\}$ *associated with* F *has at least one quasi-stationary distribution supported on* $M \setminus M_0$.

PROOF. For notational convenience, we prove this result for the Markov chain $\{N_t^\varepsilon\}$ and call $p^\varepsilon(x, y)$ its transition kernel. Let $X = \mathbb{Z}_+^k \setminus M_0$ and $q^\varepsilon(x, y)$ denote

the restriction of p^ε to X. Let $l^1(X)$ denote absolutely summable functions from X to $\mathbb{R}$. For $u \in l^1(X)$, define $\|u\|_1 = \sum_x |u(x)|$.

We can define a linear operator Q^ε from $l^1(X) \to l^1(X)$ by $(uQ^\varepsilon)(x) = \sum_{y \in X} q^\varepsilon(y, x)u(y)$. Recall that μ_ε is a QSD for the Markov chain N^ε if and only if it is a nonnegative eigenvector of the operator Q^ε. Since $F_i(x) > 0$ for all i and $x \in M \setminus M_0$, $q^\varepsilon(x, y) > 0$ for all $x, y \in X$.

Next, we show Q^ε is a compact operator, that is, the image of the unit ball under Q^ε is precompact. Recall that a closed set C in $l^1(X)$ is compact if and only if it is bounded and equisummable: for all $\delta > 0$, there exists N such that $\sup_{u \in C} \sum_{\|x\| \geq N} |u(x)| \leq \delta$. Defining $G(x) = F(\varepsilon x)/\varepsilon$, we get

$$q^\varepsilon(x, y) = \prod_{i=1}^{k} \frac{G_i(x)^{y_i}}{y_i!} \exp(-G_i(x)) \leq \prod_{i=1}^{k} \frac{m^{y_i}}{y_i!} \qquad \text{where } m = \sup_{x \in X} \|G(x)\|.$$

Hence, for u with $\|u\|_1 \leq 1$,

$$\|uQ^\varepsilon\|_1 = \sum_x \left| \sum_y q^\varepsilon(y, x)u(y) \right|$$

$$\leq \sum_x \sum_y \prod_{i=1}^{k} \frac{m^{x_i}}{x_i!} |u(y)|$$

$$\leq \sum_x \prod_{i=1}^{k} \frac{m^{x_i}}{x_i!} = e^{km}.$$

Moreover, given $\delta > 0$, choose $N > 0$ such that $\sum_{\|x\| \geq N} \prod_{i=1}^{k} m^{x_i}/x_i! < \delta$. Then $\sum_{|x| \geq N} |(uQ^\varepsilon)(x)| \leq \delta$ for all u such that $\|u\|_1 \leq 1$. Hence, Q^ε is a compact operator.

On the other hand, given that q^ε is strictly positive on $X \times X$, we have for any $Y \subset X$,

$$\sum_{x \notin Y} \sum_{y \in Y} q^\varepsilon(x, y) > 0.$$

Applying the following result of Jentzsch on Kernel positive operators completes the proof of this proposition. $\square$

THEOREM 6.2 [Schaefer (1974), Theorem V.6.6]. *Let $E = L^p(\mu)$, where $1 \leq p \leq +\infty$ and (X, Σ, μ) is a σ-finite measure space. Suppose $Q \in \mathcal{L}(E)$ is an operator given by a $(\Sigma \times \Sigma)$-measurable kernel $q \geq 0$, satisfying the following two assumptions*:

   (i) *some power of Q is compact*;

(ii) $Y \subset \Sigma$, $\mu(Y) > 0$ and $\mu(\Sigma \setminus Y) > 0$ implies

$$\int_{X \setminus Y} \int_Y q(x, y) \mu(dx) \mu(dy) > 0.$$

Then the spectral radius $r(Q)$ is positive, is a simple eigenvalue, its unique renormalized eigenvector v satisfies $v(x) > 0$, μ almost surely and $r(Q)$ is the only eigenvalue of Q with a positive eigenvector. Moreover, if $q(x, y) > 0$ ($\mu \otimes \mu$) almost surely, then every other eigenvalue of Q has modulus strictly smaller than $r(Q)$.

The following proposition verifies the large deviation hypotheses of Section 2 for the nonlinear Poisson processes.

PROPOSITION 6.3.   Assume that $x \mapsto F(x)$ is continuous, $F_i(x) > 0$ for all i and $x \in M \setminus M_0$, and $\sup_{x \in \mathbb{R}_+^k} \|F(x)\| < \infty$. Then the nonlinear Poisson process $\{X_t^\varepsilon\}$ associated with F satisfies Hypotheses 2.5 and 2.6.

PROOF.   Let μ_ε^x denote the distribution of the Poisson random vector X_{t+1}^ε conditional to $X_t^\varepsilon = x$. The logarithmic moment generating function relative to μ_ε^x is given by

$$\Lambda_{\varepsilon,x}(\lambda) = \log \mathbb{E}\big(e^{\langle \lambda, \varepsilon Z_1(F(x)/\varepsilon)\rangle}\big) = \sum_{i=1}^k \frac{F_i(x)}{\varepsilon}\big(e^{\varepsilon \lambda_i} - 1\big).$$

Hence, the family $\varepsilon \Lambda_{\varepsilon,x}(\cdot/\varepsilon)$ is identically equal on $\mathbb{R}^k$ to the function $\Lambda_x(\lambda) = \sum_{i=1}^k F_i(x)(e^{\lambda_i} - 1)$. Thus, by the Gärtner–Ellis theorem [see, e.g., Dembo and Zeitouni (1993), Theorem 2.3.6], the family μ_ε^x satisfies a large deviation principle with convex rate function $\Lambda_x^*(y) = \sum_{i=1}^k y_i \log \frac{y_i}{F_i(x)} + F_i(x) - y_i$; that is, for any closed set $F \subset \mathbb{R}_+^k$ and $x \in \mathbb{R}_+^k$,

$$\limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon^x(F) \leq - \inf_{y \in F} \Lambda_x^*(y)$$

and for any open set $G \subset \mathbb{R}_+^k$,

$$\liminf_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon^x(G) \geq - \inf_{y \in G} \Lambda_x^*(y).$$

Hence, if we define

$$\rho(x, y) = \Lambda_x^*(y),$$

then ρ immediately satisfies (i), (ii) and the upper bound of (iv) of Hypothesis 2.5.

Now let us derive the uniform lower bound of Hypothesis 2.5. Pick a compact set $K \subset M_1$ and an open ball $B \subset M$. Let $B^\varepsilon = B \cap (\varepsilon \mathbb{N}^k)$ be the ε-lattice

on B. Let $\alpha > 0$. For every $x \in K$, there exists $y(x) \in B$ such that $\rho(x, y(x)) \leq \inf_{y \in B} \rho(x, y) + \alpha$. Choose $\varepsilon_0 > 0$ small enough such that

$$d(y, y') < \varepsilon_0 \quad \Rightarrow \quad |\rho(x, y) - \rho(x, y')| < \alpha$$

for all $y, y' \in B$ and $x \in K$. For each $x \in K$ and each $0 < \varepsilon < \varepsilon_0$, we can choose a point $y_\varepsilon(x) = \varepsilon(n_1^\varepsilon, \ldots, n_k^\varepsilon)$ such that $d(y_\varepsilon(x), y(x)) < \varepsilon$ and, consequently, $\rho(x, y_\varepsilon(x)) \leq \inf_{y \in B} \rho(x, y) + 2\alpha$. For all $0 < \varepsilon < \varepsilon_0$ and all $x \in K$, we have

$$\mu_\varepsilon^x(B) \geq \mu_\varepsilon^x(\{y_\varepsilon(x)\}) = \prod_{i=1}^k e^{-F_i(x)/\varepsilon} \frac{(F_i(x)/\varepsilon)^{n_i^\varepsilon}}{n_i^\varepsilon!}.$$

Recall that, for any $p \in \mathbb{N}$, $-p - \log p! + p \log p \geq -(1 + \log p)$ and define $I_+ = \{i \in \{1, \ldots, k\} : n_i^\varepsilon > 0\}$. A straightforward computation gives

$$\varepsilon \log \mu_\varepsilon^x(B) \geq -\sum_{i=1}^k F_i(x) + \sum_{i \in I_+} \left(\varepsilon n_i^\varepsilon \log \frac{F_i(x)}{\varepsilon} - \varepsilon \log(n_i^\varepsilon!) \right)$$

$$= -\rho(x, y_\varepsilon(x)) + \sum_{i \in I_+} \varepsilon n_i^\varepsilon \left(-1 - \frac{1}{n_i^\varepsilon} \log(n_i^\varepsilon!) + \log n_i^\varepsilon \right)$$

$$\geq -2\alpha - \inf_{y \in B} \rho(x, y) - \varepsilon \sum_{i \in I_+} (1 + \log n_i^\varepsilon).$$

The last quantity goes to zero as ε goes to zero, independently of x since n_i^ε is of order $(y_\varepsilon(x))_i/\varepsilon$ and the quantities $(y_\varepsilon(x))_i$ are bounded. Hence, we have shown that the lower bound (4) holds uniformly for any compact set $K \subset M_1$ and open ball B.

To verify that ρ satisfies (iii) of Hypothesis 2.5, we first prove the following lemma.

LEMMA 6.4.   *Define* $g : (0, \infty) \times [0, \infty) \to [0, \infty)$ *by* $g(x, y) = y \log \frac{y}{x} + x - y$. *Then for all* $\delta > 0$ *and* $m > 0$,

$$\inf\{g(x, y) : |x - y| \geq \delta, x \leq m\} \geq a > 0.$$

PROOF.   Let $m > 0$ and $\delta > 0$. We have that $C := \{(x, y) : |x - y| \geq \delta, x \leq m\} = A \cup B$ with $A = \{(x, y) : m \geq x \geq \delta, 0 \leq y \leq x - \delta\}$ and $B = \{(x, y) : y \geq \delta, 0 < x \leq y - \delta\}$. Since A is compact and g restricted to A is positive and continuous, $\inf_{(x,y) \in A} g(x, y) > 0$. Restricted to B, $g(x, y)$ is positive and increasing in y. Hence, $\inf_{(x,y) \in B} g(x, y) = \inf_{0 < x \leq m} g(x, x + \delta)$. Since

$$\frac{d}{dx} g(x, x + \delta) = \log(1 + \delta/x) - \delta/x < 0,$$

we get $\inf_{0 < x \leq m} g(x, x + \delta) = g(m, m + \delta) > 0$. Thus, $\inf_{(x,y) \in C} g(x, y) > 0$.   □

Let now $\delta > 0$ be given, $d(x, y) = \max_i |x_i - y_i|$ and g be as defined in the lemma. If $|y_i - F_i(x)| \geq \delta$, then Lemma 6.4, with $m = \sup_{x \in \mathbb{R}_+^k} \|F(x)\|$, implies that $\rho(x, y) = \sum_j g(F_j(x), y_j) \geq g(F_i(x), y_i) \geq a$.

To check that the uniform upper bound (5) holds, notice that it is sufficient to prove that quantities $\mu_{\varepsilon,i}^x([F_i(x) + \delta, +\infty[)$ (where $\mu_{\varepsilon,i}^x(\cdot)$ is the i-marginal of μ_ε^x, namely the distribution of the i component of X_{t+1}^ε, conditional to $X_t^\varepsilon = x$) are bounded above by some expression which goes to zero as ε goes to zero, uniformly in $x \in K$. This is an easy consequence of Chernov's upper bound,

$$\mu_{\varepsilon,i}^x([F_i(x) + \delta, +\infty[) \leq e^{-(1/\varepsilon)g(F_i(x), F_i(x)+\delta)} \leq e^{-1/\varepsilon\beta},$$

where $\beta = \inf\{g(x, y) : |x - y| \geq \delta, x \leq m\} > 0$, and the quantity on the right-hand side goes to zero uniformly in x by (iii).

Finally to verify Hypothesis 2.6, notice that $p^\varepsilon(x, 0) = \exp(-\sum_i F_i(x)/\varepsilon)$, and recall that $F(0) = 0$. Hence, given $c > 0$, choose a neighborhood V_0 of $\{0\}$ such that $\sum_i F_i(x) \leq c$ whenever $x \in V_0$. Then $\varepsilon \log p^\varepsilon(x, 0) \geq -c$ whenever $x \in V_0$. $\square$

6.2. *Multinomial model.* Here we verify the large deviation assumptions for the multinomial processes introduced in Section 5.3.

PROPOSITION 6.5. *Assume $F_i(x) = x_i f_i(x)$, with f_i continuous and positive. Then the multinomial process $\{X_t^\varepsilon\}$ associated with F satisfies Hypotheses 2.5 and 2.6 with respect to the absorbing set $M_0 = \{x \in M : \prod x_i = 0\}$.*

PROOF. Let μ_N^x be the law of the multinomial random vector $\frac{1}{N}Z_1(F(x))$, which can be written as $\frac{1}{N}\sum_{i=1}^N Y_i(F(x))$, where $(Y_i(F(x)))_i$ is an i.i.d. sequence with distribution $\mathbb{P}[Y_1(F(x)) = e_j] = F_j(x)$ (e_j is the unitary vector in $\mathbb{R}^k$ whose jth component equals one). By Cramér's theorem [see, e.g., Dembo and Zeitouni (1993), Theorem 2.2.30], the sequence μ_N^x satisfies a large deviation principle with convex rate function $\Lambda_x^*(y) = \sum_{i=1}^k y_i \log \frac{y_i}{F_i(x)}$. Hence, if we define $\rho(x, y) = \Lambda_x^*(y)$, then ρ immediately satisfies (i), (ii), (iv) of Hypothesis 2.5.

The proof of the uniform lower bound is similar to the Poisson branching process case. Let $K \subset M_1$ be a compact set and $B \subset M$ an open ball. Let $B^N = B \cap \frac{1}{N}\mathbb{N}^k$. Let $\alpha > 0$ be given. For every $x \in K$, there exists $y(x) \in B$ such that $\rho(x, y(x)) \leq \inf_{y \in B} \rho(x, y) + \alpha$. Choose $N_0 \geq 1$ sufficiently large such that

$$d(y, y') < 1/N_0 \quad \Rightarrow \quad |\rho(x, y) - \rho(x, y')| < \alpha$$

for all $x \in K$ and $y, y' \in B$. For each $x \in K$ and $N \geq N_0$, we choose $y_N(x) = \frac{1}{N}(n_1^N, \ldots, n_k^N)$ such that $d(y_N(x), y(x)) < 1/N$. Let $I_+ = \{i : n_i^N > 0\}$. For N

large enough,

$$
\frac{1}{N}\log\mu_N^x(B) \geq \frac{1}{N}\log\mu_N^x(\{y_N(x)\})
$$

$$
= \frac{1}{N}\log\left(N!\prod_{i=1}^{k}\frac{(F_i(x))^{n_i^N}}{n_i^N!}\right)
$$

$$
\geq -\rho(x, y_N(x)) + \left(1 + \frac{1}{N}\log N! - \log N\right)
$$

$$
+ \frac{1}{N}\sum_{i\in I_+}n_i^N\left(-1 - \frac{1}{n_i^N}\log n_i^N! + \log n_i^N\right)
$$

$$
\geq -2\alpha - \inf_{y\in B}\rho(x, y) - \frac{1}{N}\sum_{i\in I_+}(1 + \log n_i^N).
$$

The last quantity goes to zero as $N \to \infty$, independently of x since n_i^N is of order $N(y_N(x))_i$ and the quantities $(y_N(x))_i$ are bounded. Hence, we have shown that the lower bound (4) holds uniformly for any compact set $K \subset M_1$ and open ball B.

To verify that ρ satisfies (iii) of Hypothesis 2.5, assume by contradiction that there exist $\beta > 0$ and two sequences $(x_n)_n$ and $(y_n)_n$ in the k-simplex M, converging, respectively, to x and y, and such that

$$
\lim_n \rho(x_n, y_n) = 0 \quad \text{and} \quad d(x_n, y_n) \geq \beta.
$$

Define $I_0 = \{i \in \{1, \ldots, k\} : x^i = 0\}$. Notice that, if $i \in I_0$, then $\lim_n y_n^i \log\frac{y_n^i}{x_n^i} = 0$ if $y^i = 0$ and $y_n^i \log\frac{y_n^i}{x_n^i} \to +\infty$ otherwise. As a consequence, $y^i = 0 \ \forall i \in I_0$. We consider two cases separately.

Assume first that all the components of x are zero except the first one, $I_0 = \{2, \ldots, k\}$. Then $y^i = 0$ for $i = 2, \ldots, k$, which implies that $x = y$, a contradiction.

Assume now that I_0 contains at most $k - 2$ terms. Call I_1 its complementary and assume without loss of generality that $I_1 = \{1, \ldots, n_1\}$. Define $\tilde{x} = \{x^1, \ldots, x^{n_1}\}$, $\tilde{y} = \{y^1, \ldots, y^{n_1}\}$ and notice that $\tilde{x} \in \text{Int}(\Delta^{n_1})$ and $\tilde{y} \in \Delta^{n_1}$, where Δ^{n_1} is the n_1-simplex. Define analogously the sequences $(\tilde{x}_n)_n$ and $(\tilde{y}_n)_n$, which belong to the set $\{u \in \mathbb{R}^{n_1} : u^i \geq 0 \sum_i u^i \leq 1\}$. Let now $\tilde{\rho}$ be the application given by

$$
(u, v) \in \mathbb{R}_*^{n_1} \times \mathbb{R}^{n_1} \mapsto \sum_{i=1}^{n_1} v_i \log\frac{v_1}{u_i}.
$$

This map is continuous and strictly positive in $(\tilde{x}, \tilde{y})$, since $|\tilde{x} - \tilde{y}| > \beta$. Therefore, there exists $\delta > 0$ and $r > 0$ such that $\tilde{\rho} > \delta$ on $B_{\mathbb{R}^{n_1}}(\tilde{x}, r) \times B_{\mathbb{R}^{n_1}}(\tilde{y}, r)$. This

concludes the proof since

$$\liminf_{n \to +\infty} \rho(x_n, y_n) = \liminf_{n \to +\infty} \left(\tilde{\rho}(\tilde{x}_n, \tilde{y}_n) + \sum_{i \in I_0} y_n^i \log \frac{y_n^i}{x_n^i} \right)$$

$$\geq \beta + \lim_{n \to +\infty} \sum_{i \in I_0} y_n^i \log \frac{y_n^i}{x_n^i} = \beta.$$

The uniform upper bound (5) holds by an application of Chernov upper bound. Let $\delta > 0$. Then

$$\mu_{N,i}^x \left([F_i(x) + \delta, +\infty[\right) \leq e^{-N\beta},$$

where $\beta = \inf\{y \log \frac{y}{x}, (x, y) \in (0, 1]^2, |x - y| \geq \delta\} > 0$.

Finally to verify Hypothesis 2.6, we have $p^\varepsilon(x, M_0) \geq \max_i (1 - F_i(x))^{1/\varepsilon}$. Hence, $\varepsilon \log p^\varepsilon(x, M_0) \geq \max_i \log(1 - x_i f_i(x))$. Given $c > 0$, choose a neighborhood V_0 of M_0 such that $\min_i x_i f_i(x) \leq 1 - e^{-c}$ whenever $x \in V_0$. Then $\varepsilon \log p^\varepsilon(x, 0) \geq -c$ whenever $x \in V_0$. $\square$

Acknowledgments. The authors would also like to thank the participants of the Probability workshop of the University of Neuchâtel for useful remarks and comments.

REFERENCES

ANDRONOV, A., PONTRYAGIN, L. and WITT, A. (1933). On the statistical investigation of dynamical systems. *J. Exp. Theor. Phys.* **3** 165–180.

ARJAS, E., NUMMELIN, E. and TWEEDIE, R. L. (1980). Semi-Markov processes on a general state space: α-theory and quasistationarity. *J. Austral. Math. Soc. Ser. A* **30** 187–200. MR0607930

BARBOUR, A. D. (1976). Quasi-stationary distributions in Markov population processes. *Adv. in Appl. Probab.* **8** 296–314. MR0410960

BUCKLEY, F. M. and POLLETT, P. K. (2010). Limit theorems for discrete-time metapopulation models. *Probab. Surv.* **7** 53–83. MR2645217

CASWELL, H. (2001). *Matrix Population Models*. Sinauer, Sunderland, MA.

CHAN, T. (1998). Large deviations and quasi-stationarity for density-dependent birth-death processes. *J. Austral. Math. Soc. Ser. B* **40** 238–256. MR1668135

CONLEY, C. (1978). *Isolated Invariant Sets and the Morse Index. CBMS Regional Conference Series in Mathematics* **38**. Amer. Math. Soc., Providence, RI. MR0511133

COOLEN-SCHRIJNER, P. and VAN DOORN, E. A. (2006). Quasi-stationary distributions for a class of discrete-time Markov chains. *Methodol. Comput. Appl. Probab.* **8** 449–465. MR2329282

CRESSMAN, R. (2003). *Evolutionary Dynamics and Extensive Form Games. MIT Press Series on Economic Learning and Social Evolution* **5**. MIT Press, Cambridge, MA. MR2004666

CRESSMAN, R., KRIVAN, V. and GARAY, J. (2004). Ideal free distributions, evolutionary games, and population dynamics in multiple-species environments. *Am. Nat.* **164** 473–489.

CUSHING, J. M., LEVARGE, S., CHITNIS, N. and HENSON, S. M. (2004). Some discrete competition models and the competitive exclusion principle. *J. Difference Equ. Appl.* **10** 1139–1151. MR2100718

DARROCH, J. N. and SENETA, E. (1965). On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *J. Appl. Probab.* **2** 88–100. MR0179842

DEMBO, A. and ZEITOUNI, O. (1993). *Large Deviation Techniques and Applications*. Springer, New York.

EARN, D. J., LEVIN, S. A. and ROHANI, P. (2000). Coherence and conservation. *Science* **290** 1360–1364.

EARN, D. J. D. and LEVIN, S. A. (2006). Global asymptotic coherence in discrete dynamical systems. *Proc. Natl. Acad. Sci. USA* **103** 3968–3971. MR2215201

FERRARI, P. A., KESTEN, H., MARTINEZ, S. and PICCO, P. (1995). Existence of quasi-stationary distributions. A renewal dynamical approach. *Ann. Probab.* **23** 501–521. MR1334159

FREIDLIN, M. I. and WENTZELL, A. D. (1984). *Random Perturbations of Dynamical Systems*. Springer, New York. MR0722136

GARAY, B. M. and HOFBAUER, J. (2003). Robust permanence for ecological differential equations, minimax, and discretizations. *SIAM J. Math. Anal.* **34** 1007–1039. MR2001657

GETZ, W. M. and MILLS, N. J. (1996). Host-parasitoid coexistence and egg-limited encounter rates. *Am. Nat.* **148** 333–347.

GIDEA, M., MEISS, J. D., UGARCOVICI, I. and WEISS, H. (2011). Applications of KAM theory to population dynamics. *J. Biol. Dyn.* **5** 44–63. MR2787172

GODFRAY, H. C. J. (1994). *Parasitoids*. Princeton Univ. Press, Princeton, NJ.

GOSSELIN, F. (2001). Asymptotic behavior of absorbing Markov chains conditional on nonabsorption for applications in conservation biology. *Ann. Appl. Probab.* **11** 261–284. MR1825466

HASSELL, M. P. (1978). *The Dynamics of Arthropod Predator–Prey Systems. Monographs in Population Biology* **13**. Princeton Univ. Press, Princeton, NJ. MR0508052

HASSELL, M. P. (2000). *The Spatial and Temporal Dynamics of Host-Parasitoid Interactions*. Oxford Univ. Press, Oxford.

HASSELL, M. P., MAY, R. M., PACALA, S. W. and CHESSON, P. L. (1991). The persistence of host-parasitoid associations in patchy environments. I. A general criterion. *Am. Nat.* **138** 586–583.

HASTINGS, A. (1993). Complex interactions between dispersal and dynamics: Lessons from coupled logistic equations. *Ecology* **74** 1362–1372.

HASTINGS, A. (1997). *Population Biology*: *Concepts and Models*. Springer, New York.

HASTINGS, A. and BOTSFORD, L. W. (2006). Persistence of spatial populations depends on returning home. *Proc. Natl. Acad. Sci. USA* **103** 6067–6072.

HOFBAUER, J. and SIGMUND, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge Univ. Press, Cambridge. MR1635735

HOFBAUER, J. and SIGMUND, K. (2003). Evolutionary game dynamics. *Bull. Amer. Math. Soc.* (*N.S.*) **40** 479–519. MR1997349

HÖGNÄS, G. (1997). On the quasi-stationary distribution of a stochastic Ricker model. *Stochastic Process. Appl.* **70** 243–263. MR1475665

IMHOF, L. A. and NOWAK, M. A. (2006). Evolutionary game dynamics in a Wright–Fisher process. *J. Math. Biol.* **52** 667–681. MR2235522

IMHOF, L. A. and NOWAK, M. A. (2010). Stochastic evolutionary dynamics of direct reciprocity. *Proceedings of the Royal Society B*: *Biological Sciences* **277** 463–468.

JACOBS, F. and SCHREIBER, S. J. (2006). Random perturbations of dynamical systems with absorbing states. *SIAM J. Appl. Dyn. Syst.* **5** 293–312. MR2237149

KERMACK, W. O. and MCKENDRICK, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **115** 700–721.

KIFER, Y. (1988). *Random Perturbations of Dynamical Systems*. Birkhäuser, Boston, MA. MR1015933

KIFER, Y. (1989). Attractors via random perturbations. *Comm. Math. Phys.* **121** 445–455. MR0990775

KIFER, Y. (1990). A discrete-time version of the Wentzell–Freidlin theory. *Ann. Probab.* **18** 1676–1692. MR1071818

KIJIMA, M. (1992). On the existence of quasi-stationary distributions in denumerable *R*-transient Markov chains. *J. Appl. Probab.* **29** 21–36. MR1147764

KLEBANER, F. C., LAZAR, J. and ZEITOUNI, O. (1998). On the quasi-stationary distribution for some randomly perturbed transformations of an interval. *Ann. Appl. Probab.* **8** 300–315. MR1620378

KON, R., SAITO, Y. and TAKEUCHI, Y. (2004). Permanence of single-species stage-structured models. *J. Math. Biol.* **48** 515–528. MR2067114

KOZLOVSKI, O. S. (2003). Axiom A maps are dense in the space of unimodal maps in the C^k topology. *Ann. of Math.* (2) **157** 1–43. MR1954263

LASSERRE, J. B. and PEARCE, C. E. M. (2001). On the existence of a quasistationary measure for a Markov chain. *Ann. Probab.* **29** 437–446. MR1825158

LESLIE, P. H. and GOWER, J. C. (1958). The properties of a stochastic model for two competing species. *Biometrika* **45** 316–330. MR0102448

LOTKA, A. J. (1925). *Elements of Physical Biology*. Williams and Witkins, Baltimore, MD.

MAY, R. M. (1995). Necessity and chance: Deterministic chaos in ecology and evolution. *Bull. Amer. Math. Soc.* (*N.S.*) **32** 291–308. MR1307905

MAYNARD SMITH, J. (1974). The theory of games and the evolution of animal conflicts. *J. Theoret. Biol.* **47** 209–221. MR0444115

NICHOLSON, A. J. and BAILEY, V. A. (1935). The balance of animal populations. *Proc. Zool. Soc. London* 551–598.

NOWAK, M. A., SASAKI, A., TAYLOR, C. and FUDENBERG, D. (2004). Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428** 646–650.

NUMMELIN, E. and ARJAS, E. (1976). A direct construction of the *R*-invariant measure for a Markov chain on a general state space. *Ann. Probab.* **4** 674–679. MR0407983

PARK, T. (1948). Experimental studies of interspecies competition. *Ecol. Monogr.* **18** 265–308.

PARK, T. (1954). Experimental studies of the interspecies competition. II. Temperature, humidity and competition in two species of *Tribolium*. *Physiol. Zool.* **27** 177–238.

PRZYTYCKI, F. (1976). Anosov endomorphisms. *Studia Math.* **58** 249–285. MR0445555

RAMANAN, K. and ZEITOUNI, O. (1999). The quasi-stationary distribution for small random perturbations of certain one-dimensional maps. *Stochastic Process. Appl.* **84** 25–51. MR1720096

RICKER, W. E. (1954). Stock and recruitment. *J. Fish. Bd. Can.* **11** 559–623.

RUELLE, D. (1981). Small random perturbations of dynamical systems and the definition of attractors. *Comm. Math. Phys.* **82** 137–151. MR0638517

SCHAEFER, H. H. (1974). *Banach Lattices and Positive Operators*. Springer, New York. MR0423039

SCHREIBER, S. J. (2006a). Host-parasitoid dynamics of a generalized Thompson model. *J. Math. Biol.* **52** 719–732. MR2235524

SCHREIBER, S. J. (2006b). Persistence despite perturbations for interacting populations. *J. Theoret. Biol.* **242** 844–852. MR2279750

SCHREIBER, S. J. (2007). Periodicity, persistence, and collapse in host-parasitoid systems with egg limitation. *J. Biol. Dyn.* **1** 273–288. MR2350952

SCHREIBER, S. J. (2010). Interactive effects of temporal correlations, spatial heterogeneity and dispersal on population persistence. *Proceedings of the Royal Society B: Biological Sciences* **277** 1907–1914.

SCHREIBER, S., FOX, L. and GETZ, W. (2000). Coevolution of contrary choices in host-parasitoid systems. *Am. Nat.* **155** 637–648.

SENETA, E. and VERE-JONES, D. (1966). On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probab.* **3** 403–434. MR0207047

SINAI, J. G. (1972). Gibbs measures in ergodic theory. *Russian Math. Surveys* **27** 21–69. MR0399421

SINERVO, B. and LIVELY, C. (1996). The rock-paper-scissors game and the evolution of alternative male strategies. *Nature* **380** 240–243.

THOMPSON, W. R. (1924). La théorie mathématique de l'action des parasites entomophages et le facteur du hassard. *Ann. Fac. Sci. Marseille* **2** 69–89.

TWEEDIE, R. L. (1974). Quasi-stationary distributions for Markov chains on a general state space. *J. Appl. Probab.* **11** 726–741. MR0443090

VAN STRIEN, S. J. (1981). On the bifurcations creating horseshoes. In *Dynamical Systems and Turbulence*, *Warwick* 1980 (*Coventry*, 1979/1980). *Lecture Notes in Math.* **898** 316–351. Springer, Berlin. MR0654898

VOLTERRA, V. (1926). Fluctuations in the abundance of a species considered mathematically. *Nature* **118** 558–560.

WYSHAM, D. B. and HASTINGS, A. (2008). Sudden shifts in ecological systems: Intermittency and transients in the coupled Ricker population model. *Bull. Math. Biol.* **70** 1013–1031. MR2391177

ZEEMAN, E. C. (1980). Population dynamics from game theory. In *Global Theory of Dynamical Systems* (*Proc. Internat. Conf.*, *Northwestern Univ.*, *Evanston*, *Ill.*, 1979). *Lecture Notes in Math.* **819** 471–497. Springer, Berlin. MR0591205

GREQAM, CENTRE DE LA VIEILLE CHARITÉ
2 RUE DE LA VIEILLE CHARITÉ
13236 MARSEILLE CEDEX 02
FRANCE
E-MAIL: mathieu.faure@univ-amu.fr

DEPARTMENT OF EVOLUTION AND ECOLOGY
UNIVERSITY OF CALIFORNIA
DAVIS, CALIFORNIA 95616
USA
E-MAIL: sschreiber@ucdavis.edu