The Quasi-Stationary Distribution for Small Random Perturbations of Certain One-Dimensional Maps

Kavita Ramanan*
Bell Laboratories
Murray Hill, NJ 07974
U.S.A.

Ofer Zeitouni †
Department of Electrical Engineering
Technion—Israel Institute of Technology
Haifa 32000, Israel

April 3, 1998. Revised April 4, 1999

Abstract

We analyze the quasi-stationary distributions of the family of Markov chains $\{X_n^\varepsilon\}, \varepsilon > 0$, obtained from small non-local random perturbations of iterates of a map $f : I \to I$ on a compact interval. The class of maps considered is slightly more general than the class of one-dimensional Axiom A maps. Under certain conditions on the dynamics, we show that as $\varepsilon \to 0$ the limit quasi-stationary distribution of the family of Markov chains is supported on the union of the periodic attractors of the map f. Moreover, we show that these conditions are satisfied by Markov chains obtained as perturbations of the logistic map $f(x) = \mu x(1-x)$ by additive Gaussian noise and also by Markov chains that model density-dependent branching processes.

Key words and phrases. Quasi-stationary distribution, one-dimensional dynamics, Axiom A maps, periodic attractors, logistic map, density-dependent branching processes.

1 Introduction

A transformation $f : I \to I$ of a compact interval $I \subset I\!\!R$ defines a one-dimensional discrete dynamical system $\{x_n\}$ given by $x_0 = x \in I$ and for $n = 0, 1, \ldots,$

$$x_{n+1} = f(x_n). \tag{1.1}$$

Small perturbations of the deterministic system (1.1) by state-dependent noise $\xi^\varepsilon(\cdot)$ give rise to the family of Markov chains $\{X_n^\varepsilon\}, \varepsilon > 0$, defined iteratively by $X_0^\varepsilon = x$ and for $n = 0, 1, \ldots,$

$$X_{n+1}^\varepsilon = f(X_n^\varepsilon) + \xi^\varepsilon(X_n^\varepsilon). \tag{1.2}$$

Note that the recursion (1.2) is well defined only as long as $\{X_n^\varepsilon\}$ remains in the interval I. The main objective of this paper is to characterize the long term behaviour of the chains $\{X_n^\varepsilon\}$, $\varepsilon > 0$, conditioned on staying within the interval I, in the limit as noise tends to zero. In applications one often encounters the situation where for each $\varepsilon > 0$, $\{X_n^\varepsilon\}$ is a Markov chain

*The work of this author was done at the Department of Electrical Engineering, Technion. Partially supported by a US-Israel BSF grant.

†Partially supported by a US-Israel BSF grant.

on a countable lattice $\mathcal{L}^\varepsilon$ which is eventually absorbed (with probability one) into the boundary ∂I of I. From general Markov chain theory, it then follows that the stationary distribution of the Markov chain $\{X_n^\varepsilon\}$ is concentrated on ∂I. In contrast, the deterministic system could have invariant measures that are supported in the interior I° of I. Thus the stationary distributions of the Markov chains do not give much insight into the behaviour of the deterministic system under small perturbations. In such situations it is more pertinent to study the behaviour of the Markov chain $\{X_n^\varepsilon\}$ conditioned on not being absorbed or, alternatively, conditioned on staying within the interior I° of the interval I. Thus we introduce the *quasi-stationary distribution* $\rho^\varepsilon(\cdot)$ of the Markov chain $\{X_n^\varepsilon\}$, which is defined to be a probability measure that satisfies

$$\rho^\varepsilon(A) = \lim_{n\to\infty} P(X_n^\varepsilon \in A | X_n^\varepsilon \in I(\varepsilon)) \tag{1.3}$$

for every Borel set $A \subset I\!R$, where $I(\varepsilon) = I$ if $\{X_n^\varepsilon\}$ is a Markov chain on a continuous state space and $I(\varepsilon) = I^\circ \cap \mathcal{L}^\varepsilon$ if $\{X_n^\varepsilon\}$ is a Markov chain taking values on a countable lattice $\mathcal{L}^\varepsilon$. We then define a *limit quasi-stationary distribution* ρ of the family of Markov chains $\{X_n^\varepsilon\}$, $\varepsilon > 0$, to be a probability measure that satisfies

$$\rho^{\varepsilon_k} \Rightarrow \rho \tag{1.4}$$

along some subsequence $\varepsilon_k \to 0$, where $\Rightarrow$ denotes weak convergence and ρ^ε is the quasi-stationary distribution of the chain $\{X_n^\varepsilon\}$ as defined in (1.3). Quasi-stationary distributions were first studied by Yaglom for Markov chains on countable state spaces [23]. Consequently the quantity $\rho^\varepsilon(\cdot)$ defined in (1.3) is also sometimes referred to as a Yaglom limit. For subsequent work on quasi-stationary distributions for Markov chains, see [7, 8, 21, 22] and the references therein. Using the theory of Krein-Rutman, under suitable conditions standard arguments guarantee the existence and uniqueness of a quasi-stationary distribution ρ^ε for every $\varepsilon > 0$, and show that any limit quasi-stationary distribution must be supported on f-invariant subsets of I (see Theorems 3.1 and 3.2). Our goal in this paper is to show that for a certain class of once continuously differentiable endomorphisms f of a compact interval, under suitable conditions on the noise, the support of the limit quasi-stationary distribution of the family of Markov chains $\{X_n^\varepsilon\}, \varepsilon > 0$, is contained in the union of the periodic attractors of f.

The motivation for our work stems from the fact that dynamical systems like (1.1) are often used to model physical phenomena. For example $\{x_n\}$ may represent the population density of the nth generation in some region, or may represent the proportion of predators in a predator-prey population at the nth time step [19, 11]. Markov chains that satisfy (1.2) represent the natural stochastic analogues of these deterministic models. For some models, the Markov chains have been shown to approach the corresponding deterministic system (in the sense of the strong law of large numbers and the central limit theorem) as the parameter $\varepsilon \to 0$ [15, 16]. The well-known logistic map $f(x) = \mu x(1-x)$ with $\mu \in (0, 4]$ is one of the most commonly used maps to model the evolution of population density [19]. Moreover the dynamics of the logistic map, as the parameter μ increases from 0 to 4, exhibits most of the features present in one-dimensional maps of a compact interval. Thus in Section 5, we apply our main result to Markov chains obtained as small perturbations of the logistic map. In particular, we verify our assumptions for the case of additive Gaussian noise and also for a model of noise arising from density-dependent branching processes like that considered in [12, 17, 18]. We show that for a certain set of parameters μ that is dense in $(1, 4)$, any limit quasi-stationary distribution of the family of Markov chains $\{X_n^\varepsilon\}$, $\varepsilon > 0$, obtained by each type of perturbation described above is concentrated on the union of the periodic attractors of f. This set of parameters includes certain values μ for which the logistic map has infinitely many repelling periodic orbits.

Our results generalize those of [18] and [12], both of which consider non-local random perturbations of a map on an interval. For a class of piecewise $\mathcal{C}^2$ maps $f : [0,1] \to [0,1]$ that satisfy $f(0) = f(1) = 0$, have one attracting periodic orbit of period $p = 1$ or $p = 2$, and have at most one repelling fixed point of $f^{(p)}$ in $(0,1)$, it was shown in [18] that under certain conditions on the noise the limit quasi-stationary distribution is uniform on the periodic attractor. The logistic map satisfies these properties only for parameter values $\mu \in (1, 1+\sqrt{6})$. In [12] a class of discrete branching processes whose mean behaviour is described by the Ricker model, $f(x) = xe^{\mu-x}$, was considered. It was shown there that for a range of parameters of μ for which f has one periodic attractor and a finite number of repelling periodic orbits, the limit quasi-stationary distribution is uniform on the periodic attractor. This corresponds in the logistic map case to parameters $\mu \in (1, r^*)$, where $r^* < 4$ is the value that denotes the end of the period-doubling regime [3, Chapter 1.12]. Analogous results for (small noise) diffusion processes have been derived in [13], and for local perturbations in [20]. For stationary limiting distributions (when no extinction is present, and the noise enters via random compositions of maps), see [9].

The outline of the paper is as follows. In Sections 2 and 3 we state our assumptions on the deterministic dynamics and the noise respectively. In Section 3 we also show that the assumptions on the noise imply the existence and uniqueness of a quasi-stationary distribution. The statement, outline of proof, and proof of the main result, Theorem 4.1, is given in Section 4. Section 5 contains applications to perturbations of the logistic map. We state some open problems and make some concluding remarks in Section 6.

2 Description of the Deterministic Dynamics

Let I be a compact interval. We consider deterministic dynamical systems of the form (1.1), where $f : I \to I$ belongs to the class of generalized Axiom A maps described in Definition 2.2, and satisfies the additional condition stated in (2.5). We provide some concrete examples of generalized Axiom A maps following Definition 2.2. We then show in Lemma 2.4 that maps in this class possess a desirable expansive property that allows a useful decomposition of the interval I. This decomposition greatly facilitates the analysis of the behaviour of these maps when subjected to certain non-local random perturbations (whose properties are specified precisely in the next section).

In order to define the class of generalized Axiom A maps we first introduce some terminology from the theory of dynamical systems, taken mainly from [4]. We alert the reader that contrary to normal convention, when we refer to a neighbourhood we do not necessarily imply that it is open unless explicitly stated. Let $\mathbb{N} \doteq \{0, 1, \ldots\}$ denote the set of natural numbers and for $i \in \mathbb{N}$, let f^i represent the ith iterate of f and let $f^{-i}(A) = \{x : f^i x \in A\}$. For $i \in \mathbb{N}$, $\mathcal{C}^i$ represents the space of i times continuously differentiable functions, and given any compact interval I, $\mathcal{C}^i[I, I]$ is defined to be the space of $\mathcal{C}^i$ endomorphisms of I such that $f(\partial I) \subset \partial I$, where ∂I is the boundary of I. Suppose $f : I \to I$ is $\mathcal{C}^1$. Then either f' or Df, as is convenient, will be used to denote the derivative of f, and the set $C(f) = \{x : f'(x) = 0\}$ is defined to be the set of critical points of f. A point $s \in I$ is said to be *periodic* for f with period p if $f^p(s) = s$ and $f^i(s) \neq s$ for $i = 1, \ldots, p-1$. A fixed point is a point that is periodic with period 1. The (forward) *orbit* of a point $x \in I$ is the set $\mathcal{O}(x) \doteq \{f^i(x), i \in \mathbb{N}\}$. The orbit $\mathcal{O}(s)$ of a periodic point s is called a *periodic orbit.* A periodic orbit $\mathcal{O}(s)$ is said to be an *attractor* if $B(s) \doteq \{x \in I : f^i(x) \to \mathcal{O}(s) \text{ as } i \to \infty\}$ contains an open set. The set $B(s)$ is called the *basin* of $\mathcal{O}(s)$. For a $\mathcal{C}^1$ map, any attracting periodic point s which has period p must clearly satisfy

$|Df^p(s)| \leq 1$. In this case, there exists a neighbourhood $W(s) \subset B(s)$ of the periodic orbit which possesses the following contracting property. Every interval $G \subset W(s)$ that contains a point of the orbit $\mathcal{O}(s)$ satisfies $f(G) \subset W(s)$. We call $W(s)$ the *contracting basin* of the periodic attractor. If $W(s)$ is an open neighbourhood of the periodic orbit, then the periodic attractor is said to be *two-sided*, and otherwise it is said to be *one-sided*. We define the basin $B(f)$ of f to be the union of the basins of all the periodic attractors of f. Likewise $W(f)$ is the union of all the contracting basins of f. Finally, a set D is said to be *forward invariant* if $f(D) \subset D$, and *fully invariant* if in addition $f^{-1}(D) \subset D$. We now introduce the definition of a hyperbolic set [4, p. 220].

Definition 2.1 (Hyperbolic Set) *Let $f : I \to I$ be a $\mathcal{C}^1$ map. A subset $K \subset I$ is a hyperbolic set if K is forward invariant and there exist constants $H > 0$ and $\lambda > 1$ such that for all $x \in K$ and $n \in \mathbb{N}$,*

$$|Df^n(x)| > H\lambda^n. \tag{2.1}$$

We can now define the class of generalized Axiom A maps.

Definition 2.2 (Generalized Axiom A Maps) *A map $f \in \mathcal{C}^1(I, I)$ is said to be generalized Axiom A if*

1. *f has a finite number of two-sided periodic attractors and no one-sided periodic attractors.*
2. *The set $K = I \setminus B(f)$ is a hyperbolic set, where $B(f)$ is the union of the basins of the periodic attractors of f.*

Remark 2.3 *In the literature Axiom A maps are defined to be those that satisfy both properties stated in Definition 2.2, and in addition satisfy the condition that all periodic attractors be hyperbolic (in the sense that $|Df^p(s)| < 1$ for every point s on a periodic attractor with period p) [4, p. 221]. Since we do not require this additional hyperbolicity condition, which in particular allows us to consider points at which period-doubling bifurcations occur [3, 4], we refer to our class of maps as generalized Axiom A. Observe that for a map to satisfy Definition 2.2 it must have at least one two-sided periodic attractor since the fact that I is a compact interval and $f : I \to I \in \mathcal{C}^1$ implies that the whole interval cannot be a hyperbolic set.*

Generalized Axiom A maps occur quite commonly in one-dimensional dynamical systems. They are dense in the space of $\mathcal{C}^1$ maps endowed with the metric

$$d(f, g) \doteq \sup_{x \in I} \left(|f(x) - g(x)|, |Df(x) - Dg(x)| \right).$$

Furthermore, Mane's theorem shows that any $\mathcal{C}^2$ map that possesses a two-sided periodic attractor and has all its critical points in the basin $B(f)$ is generalized Axiom A [4, Theorems III.2.1 and III.2.2]. When the map is $\mathcal{C}^3$, and each critical point lies in the basin of a two-sided attractor, the negativity of the Schwarzian derivative

$$Sf(x) \doteq \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2$$

implies that f is a generalized Axiom A map [4, Theorem III.3.2]. In particular this property is satisfied by unimodal $\mathcal{C}^3$ maps with negative Schwarzian derivative that possess a two-sided periodic attractor, and for which the boundary points of I are not in the basin of the periodic attractor [2, Theorem II.4.1]. The much studied logistic map $f : [0,1] \to [0,1]$, $f(x) = \mu x(1-x)$, satisfies $f(0) = f(1) = 0$, has one critical point, a negative Schwarzian derivative and a repelling fixed point at 0 for all $\mu \in (1,4)$. Consequently the logistic map is generalized Axiom A for all parameter values $\mu \in (1,4)$ for which it has a two-sided periodic attractor. This set of parameter values is in fact dense in the interval $(1,4)$ [4, p. 223].

Certain maps on non-compact intervals can also be transformed into maps of the generalized Axiom A class. More precisely, if $f : J \to J \in \mathcal{C}^1$ is a map on a non-compact interval for which there exist a, b such that $-\infty < a = \inf_{x\in J} f(x) < \sup_{x\in J} f(x) = b < \infty$, then one can equivalently study $f : [a,b] \to [a,b]$ by considering the first iterates of points in J as the initial points in $[a,b]$. It is easy to see that the case when the set of points $U = \{x : f(x) \to \infty\}$ is such that $J \setminus U$ is an interval can also be reduced to the compact case in a similar fashion. This shows us that certain other families of maps commonly used to model population growth like the Ricker map, which is defined on $[0,\infty)$ by $f(x) = xe^{\mu - x}$, also satisfy Definition 2.2 for a subset of parameter values μ.

We now show that for generalized Axiom A maps it is possible to decompose the interval I into two disjoint "stable" and "unstable" regions such that there exists a finite iterate of f which maps all points in the stable region into the contracting basin, and which is expansive on the complement. This property will be used in Lemma 4.7 to estimate the exit time of the perturbed dynamical system from a certain subset of I. We will henceforth drop the dependence on f in the notation for the basin $B(f)$, contracting basin $W(f)$, and critical points $C(f)$, and instead simply refer to them as B, W and C, respectively. As usual for any set $A \subset I$, the interior of A is denoted by A°, the closure by $\bar{A}$ and the complement by A^c. The α-fattening W^α of a set $W \subset I$ is defined to be $\{x : |x-y| < \alpha \text{ for some } y \in W\}$, and correspondingly $W^{-\alpha}$ is given by $\{x : |x-y| > \alpha \text{ for all } y \in W^c\}$. The $\alpha-$fattening of a point x is denoted by $U_\alpha(x)$. For a set $S \subset I\!R$ and a point $x \in I\!R$ we let $d(x,S) = \inf_{y\in S} |x-y|$ denote the distance of x from the set S.

Lemma 2.4 *Suppose $f : I \to I$ is a generalized Axiom A map and let $K = I \setminus B$ be the associated hyperbolic invariant set. Let $\delta > 0$ be such that $W^{-\delta}$ contains the union of the periodic attractors in its interior. Then there exist constants $m, j < \infty$, $\eta > 1$ and $\gamma, L > 0$ such that*

1. *For every $x \in K^\gamma \cap I$*
$$|Df^m(x)| > \eta. \tag{2.2}$$

2.
$$\inf_{i=0,\dots,m-1} \inf_{\{z : d(z, f^i(K^\gamma \cap I)) < \gamma/2\}} |f'(z)| \geq L. \tag{2.3}$$

3.
$$Z \doteq \{x \in I : f^i(x) \notin W^{-\delta} \text{ for } i = 0,1,\dots,j-1\} \subset K^{\gamma/2}. \tag{2.4}$$

Proof. Since f is generalized Axiom A, K is a hyperbolic set and so by Definition 2.1 there exist $H > 0$, $\lambda > 1$ such that (2.1) holds. Since the periodic attractors are two-sided, W is open and hence $B = \cup_{i\in I\!N} f^{-i}(W)$ is also open, and consequently K is compact. Fix $\eta > 1$ and choose $m < \infty$ such that $H\lambda^m > \eta > 1$. Then by (2.1) for $n \geq m$ and $x \in K$ we have $|Df^n(x)| > \eta > 1$. The continuity of Df^m guarantees the existence of $\gamma > 0$ such that $|Df^m(x)| > \eta > 1$ for all $x \in K^\gamma \cap I$, which proves (2.2) above.

The fact that K satisfies (2.1) implies that it contains no critical points of f and so $K \cap C = \emptyset$. Note that because C and K are closed and $C \subset B$, one can choose $\gamma > 0$ smaller if necessary to ensure that $\overline{C^\gamma} \subset B$ (and still satisfy (2.2)). Since $f \in \mathcal{C}^1$ and B is fully invariant, $\cup_{i=0}^{m-1} f^{-i}(\overline{C^\gamma})$ is closed and is contained in B. Hence by choosing $\gamma > 0$ yet smaller if necessary one can guarantee that

$$K^\gamma \cap \left[\cup_{i=0}^{m-1} f^{-i}(\overline{C^\gamma})\right] = \emptyset,$$

which implies that $\left[\cup_{i=0}^{m-1} f^{i}(K^\gamma \cap I)\right] \cap \overline{C^\gamma} = \emptyset$. Thus (2.3) is satisfied with $L \doteq \inf_{z \in I \setminus C^{\gamma/2}} |f'(z)| > 0$.

Since B is the basin of attraction it follows that for every $x \in \overline{B^{-\gamma/2}}$, there exists an open interval $U(x) \subset B$ containing x and a constant $m(x) \in I\!N$ such that $f^{m(x)}(U(x)) \subset W^{-\delta}$. Since $\{U(x), x \in \overline{B^{-\gamma/2}}\}$ is an open covering of the compact set $\overline{B^{-\gamma/2}}$, there exists a finite subcover $\{U(x_i), x_i \in \overline{B^{-\gamma/2}}, i = 1, \ldots, F\}$. Hence if $M_1 \doteq \max_{i=1,\ldots,F} m(x_i)$ then $f^n(\overline{B^{-\gamma/2}}) \subset W^{-\delta}$ for all $n \geq M_1$ because $f(W^{-\delta}) \subset W^{-\delta}$. Let $j = 2M_1$ and define Z as in (2.4). Then clearly $Z \cap \overline{B^{-\gamma/2}} = \emptyset$ since f maps $\overline{B^{-\gamma/2}}$ into $W^{-\delta}$ in less than j steps, which implies that $Z \subset I \setminus \overline{B^{-\gamma/2}} = K^{\gamma/2} \cap I$. ■

We will require that the mapping f has an additional contraction property stated below in (2.5), which is needed to control the time of exit of the perturbed dynamical system from the interval I (see Lemma 4.7). For notational convenience, throughout this paper we denote $\overline{I^{-\theta}}$ by I_θ for any $\theta > 0$.

Assumption 2.1 *The map f is a generalized Axiom A map. Moreover, there exists $\theta_0 > 0$ such that for every $\theta \in (0, \theta_0)$ there exists $\theta' > \theta$ such that*

$$f(I_\theta) \subset I_{\theta'}. \tag{2.5}$$

For the rest of the paper we will always assume that the deterministic dynamical system satisfies Assumption 2.1. Since $f \in \mathcal{C}^1[I, I]$ maps ∂I to ∂I it follows that either both end points of I are fixed, or both end points are mapped to a fixed end point, or the end points form a periodic orbit of period two. It is easy to see that a consequence of Assumption 2.1 is that all the periodic points in ∂I are hyperbolic repelling, i.e. there exists $r > 1$ such that $|Df^2(s)| > r^2$ for any periodic point $s \in \partial I$.

3 Assumptions on the Noise

In this section we state our assumptions on the family of Markov chains $\{X_n^\varepsilon\}, \varepsilon > 0$, satisfying (1.2), i.e. $X_0^\varepsilon = x$ and

$$X_{n+1}^\varepsilon = f(X_n^\varepsilon) + \xi^\varepsilon(X_n^\varepsilon).$$

Here, the function $f(\cdot)$ extends to all of $I\!R$ simply by taking $f(x) = 0$ for $x \in I^c$. The distribution Q_x^ε of the noise $\xi^\varepsilon(x)$ is assumed to depend only on x and ε. We denote the set of Borel sets in $I\!R$ by $\mathcal{B}(I\!R)$. Let $\Pi^\varepsilon : (I\!R, \mathcal{B}(I\!R)) \to [0,1]$ be the transition kernel of the time homogeneous Markov chain $\{X_n^\varepsilon\}$ so that for every set $A \subset \mathcal{B}(I\!R)$ and $x \in I\!R$,

$$\Pi^\varepsilon(A|x) \doteq P(X_1^\varepsilon \in A | X_0^\varepsilon = x). \tag{3.1}$$

The time homogeneity of the Markov chain implies that $\Pi^\varepsilon(A|x) = P(X_{n+1}^\varepsilon \in A | X_n^\varepsilon = x)$. We denote by $\Pi_n^\varepsilon(A|x) : (I\!R, \mathcal{B}(I\!R)) \to [0,1]$ the n-step transition kernel of the Markov chain $\{X_n^\varepsilon\}$. We use dx or $\lambda(dx)$, as is convenient, to denote Lebesgue measure. For two functions on $I\!R$, we use the obvious notation $f > g$ to mean that $f(y) > g(y)$ for all $y \in I\!R$.

Assumption 3.1 *The family of Markov chains* $\{X_n^\varepsilon\}$, $\varepsilon > 0$, *satisfies the conditions stated below.*

1. *For each* $\varepsilon > 0$, *there exists a probability measure* V^ε *on* $(I, \mathcal{B}(I))$ *and a non-negative function* $\pi_x^\varepsilon(y)$ *on* $I \times I$ *such that for every* $x \in I$,

$$\Pi^\varepsilon(A|x) = \int_A \pi_x^\varepsilon(y) V^\varepsilon(dy),$$

for all $A \in \mathcal{B}(I)$. *(Recall that* $\Pi^\varepsilon(I\!R|x) = 1$, *and note that* $\pi_x^\varepsilon(\cdot)$ *is a sub-probability kernel.) Moreover, for every* $x \in I$, *let* $\pi_{x,n}^\varepsilon$ *be the density of* $\Pi_n^\varepsilon(\cdot|x)$ *with respect to* $V^\varepsilon(\cdot)$. *Then one of the following holds.*
a) For every $\varepsilon > 0$, $V^\varepsilon(dy) = dy$ *is Lebesgue measure and there exist integers* $M_0(\varepsilon)$ *and real numbers* $a(\varepsilon)$ *and* $b(\varepsilon)$ *such that for all* $x, y \in I$,

$$0 < a(\varepsilon) \le \pi_{x,M_0(\varepsilon)}^\varepsilon(y) \le b(\varepsilon) < \infty.$$

b) For every $\varepsilon > 0$, V^ε *is proportional to the counting measure on a countable lattice* $\mathcal{L}^\varepsilon$, $V^\varepsilon(\mathcal{L}^\varepsilon \cap I^\circ) = 1$, *and the matrix* $[\pi_x^\varepsilon(y) V^\varepsilon(y)]$ *restricted to* $\mathcal{L}^\varepsilon \cap I^\circ$ *is strictly substochastic and irreducible. Moreover, for every* $\varepsilon > 0$ *and* $x \in \partial I$, $\Sigma_{y \in \partial I} \pi_x^\varepsilon(y) = 1$. *Finally,* $\Sigma_{x,y \in I^\circ \cap \mathcal{L}^\varepsilon} \pi_x^\varepsilon(y) < \infty$ *(which ensures that* $\pi^\varepsilon(x|y)$ *is a positive compact operator).*

2. *Let* Q_x^ε *denote the distribution of* $\xi^\varepsilon(x)$. *One of the following holds.*
a) $\{X_n^\varepsilon\}$, $\varepsilon > 0$, *satisfies (1a) and there exists* $\beta \in (0, 1/2)$ *such that for every* $x \in I$,

$$\pi_x^\varepsilon(\cdot) \ge \frac{\beta}{\varepsilon} 1_{[f(x)-\varepsilon, f(x)+\varepsilon]}(\cdot). \tag{3.2}$$

b) $\{X_n^\varepsilon\}$, $\varepsilon > 0$, *satisfies (1b) and there exists* $a_1 \in (0, \infty)$ *such that* $w_\varepsilon \le \varepsilon^{a_1}$, *where* $w_\varepsilon \doteq \inf\{d(x, \partial I) : x \in I^\circ \cap \mathcal{L}^\varepsilon\}$. *Also for every* $\theta > 0$ *there exists* $\beta(\theta) \in (0, 1/2)$ *such that, for every* ϵ *small enough and for every* $x \in I_\theta$,

$$\pi_x^\varepsilon(\cdot) \ge \frac{\beta(\theta)}{\varepsilon} 1_{[f(x)-\varepsilon, f(x)+\varepsilon]}(\cdot). \tag{3.3}$$

Moreover, there exists $\kappa \in (0, 1/2)$ *such that for any fixed point* $s \in \partial I$ *of* f^2, *and any* $\varepsilon > 0$ *small enough,*

$$\inf_{x \in I^\circ \cap U_\kappa(s)} Q_x^\varepsilon(0, \infty) > \kappa \quad \text{and} \quad \inf_{x \in I^\circ \cap U_\kappa(s)} Q_x^\varepsilon(-\infty, 0) > \kappa. \tag{3.4}$$

3. *There exists* $\lambda_0 > 0$ *such that for all* $|\lambda| \le \lambda_0$,

$$\sup_{x \in I, \varepsilon} \Lambda_\varepsilon^\xi(x, \lambda/\varepsilon) < \infty, \tag{3.5}$$

where for $x \in I$ *and* $\varepsilon > 0$,

$$\Lambda_\varepsilon^\xi(x, \lambda) \doteq \log\left(\int_{I\!R} e^{\lambda y} Q_x^\varepsilon(dy)\right).$$

The assumptions on the noise imposed above are quite natural. As shown in Theorem 3.1, condition (1) guarantees the existence and uniqueness of a quasi-stationary distribution ρ^ε for the Markov chain $\{X_n^\varepsilon\}$ for every $\varepsilon > 0$. Condition (2) ensures that the noise does not become too small (i.e. less than order ε) too fast. The necessity for this assumption is best seen by considering the extreme case when there is no noise at all - in which case any invariant measure of the deterministic system is a limit quasi-stationary distribution and thus its support need not be contained in the periodic attractors of the map. When the Markov chain has a countable state space with almost sure absorbing states on the boundary of I, the noise $\xi^\varepsilon(x)$ is zero for $x \in \partial I$. For such Markov chains condition (3.2) cannot hold throughout the interval I, but it is reasonable to expect the chain to satisfy (3.3). In that case one also imposes the additional condition (3.4), which ensures that the noise near the boundary of the interval is sufficiently large so that the chain exits a neighbourhood of the repelling periodic points on the boundary ∂I in polynomial time as long as w_ε decays to zero at most polynomially in ε. Finally, condition (3) implies that the random variables $\xi^\varepsilon(x)$ converge in probability to zero at an exponential rate as $\varepsilon \to 0$, uniformly in $x \in I$.

We now quote theorems that prove existence of quasi-stationary distributions for Markov chains that satisfy Assumption 3.1. We will use $\Rightarrow$ to denote weak convergence. We define $I(\varepsilon) \doteq I$ or $I(\varepsilon) \doteq I^\circ \cap \mathcal{L}^\varepsilon$ for every $\varepsilon > 0$ depending on whether the family of Markov chains satisfies (1a) or (1b) respectively of Assumption 3.1.

Theorem 3.1 *Consider the family of Markov chains $\{X_n^\varepsilon\}$, $\varepsilon > 0$, defined in (1.2). Suppose f satisfies Assumption 2.1 and the Markov chain satisfies (1) and (3) of Assumption 3.1. Then the following hold.*

1. *For every $\varepsilon > 0$ the quasi-stationary distribution ρ^ε defined by (1.3) exists, it is a probability measure on $(I(\varepsilon), \mathcal{B}(I(\varepsilon)))$, and there exists a number $R^\varepsilon > 1$ such that for all $A \in \mathcal{B}(I(\varepsilon))$,*

$$\rho^\varepsilon(A) = R^\varepsilon \int_{I(\varepsilon)} \pi_x^\varepsilon(A)\rho^\varepsilon(dx). \tag{3.6}$$

2. *There exists $c > 0$ such that $R^\varepsilon \leq 1 + e^{-c/\varepsilon}$, and so $\lim_{\varepsilon\to 0} R^\varepsilon = 1$.*

Proof. When (1a) of Assumption 3.1 is satisfied, the existence of $R^\varepsilon > 1$ and a quasi-stationary distribution ρ^ε satisfying (3.6) follows from [18, Theorem 1]. If (1b) of Assumption 3.1 is satisfied then for each $\varepsilon > 0$ the measure $V^\varepsilon(\cdot)$ is the counting measure on a countable lattice $\mathcal{L}^\varepsilon$ and the matrix $\pi_x^\varepsilon(y)$ is a compact positive operator on $\ell^1(\mathcal{L}^\varepsilon \cap I^\circ)$. By the Krein-Rutman Theorem [5, p.2130] this guarantees that the spectral radius $1/R^\varepsilon$ of the operator $\pi_x^\varepsilon(y)$ is a simple eigenvalue with positive right and left eigenvectors. The strict substochasticity and irreducibility of the matrix $\pi_x^\varepsilon(y)V^\varepsilon(y)$ restricted to $\ell^1(\mathcal{L}^\varepsilon \cap I^\circ)$ show that $R^\varepsilon > 1$ and that the eigenvectors are strictly positive. In both cases, the exponential rate of convergence of R^ε to 1 follows as a consequence of (3) of Assumption 3.1, as shown in the proof of (b) of Theorem 1 in [18]. ■

Theorem 3.2 *Consider the family of Markov chains $\{X_n^\varepsilon\}$, $\varepsilon > 0$, defined in (1.2). Suppose f satisfies Assumption 2.1 and the Markov chain satisfies (1) of Assumption 3.1. Then any weak limit ρ of ρ^ε is an invariant measure for f. In other words, for any $A \in \mathcal{B}(I)$,*

$$\rho(A) = \rho(f^{-1}(A)). \tag{3.7}$$

Proof. Let ρ be any weak limit of $\{\rho^\varepsilon\}$. When Assumptions 2.1 and 3.1 are satisfied, the fact that $\rho(A) = \rho(f^{-1}(A))$ for all $A \in \mathcal{B}(I^\circ)$ follows from the proof of Theorem 2 in [18]. Note that the condition (2.5) in Assumption 2.1 guarantees that $f^{-1}(\partial I) = \partial I$, and so the dynamics on ∂I can be of three types. Let s_1, s_2 be the left and right end points of I respectively. If $f(s_1) = s_1$ and $f(s_2) = s_2$, then clearly $f^{-1}(s_1) = s_1$ and $f^{-1}(s_2) = s_2$ and thus (3.7) holds for all $A \in \mathcal{B}(I)$. The case when s_1, s_2 form a periodic orbit of period two is dealt with in a similar manner since then s_1 and s_2 are both fixed points of f^2. Finally, consider the case when $s_1 = f(s_1) = f(s_2)$. By (2.5) of Assumption 2.1 and the continuity of f, it follows that there exists $\delta' > 0$ such that $f(I) \subset [s_1, s_2 - \delta')$. Thus

$$\begin{aligned}\rho^\varepsilon((s_2 - \delta'/2, s_2]) &= R^\varepsilon \int_I P_x(X_1^\varepsilon \in (s_2 - \delta', s_2])\rho^\varepsilon(dx) \\ &\le R^\varepsilon \max\nolimits_{x \in I} P(\xi^\varepsilon(x) > \delta'/2),\end{aligned}$$

which tends to zero as $\varepsilon \to 0$ due to (3) of Assumption 3.1 and the fact that $\lim_{\varepsilon \to 0} R^\varepsilon = 1$. This leads to the conclusion that $\rho(\{s_2\}) = 0$. Since ∂I is f-invariant, this implies that $\rho(\{s_1\}) = \rho(\{f^{-1}(s_1)\})$. Thus in this case too (3.7) is valid for all $A \in \mathcal{B}(I)$. ■

4 The support of the quasi-stationary distribution

Consider a family of Markov chains $\{X_n^\varepsilon\}, \varepsilon > 0$, defined by (1.2), that satisfies Assumptions 2.1 and 3.1. Theorems 3.1 and 3.2 showed that for every $\varepsilon > 0$ there exists a quasi-stationary distribution ρ^ε for the Markov chain $\{X_n^\varepsilon\}$, and moreover that there exists an f-invariant measure ρ such that $\rho^\varepsilon \Rightarrow \rho$ as $\varepsilon \to 0$. Our main result is the following theorem.

Theorem 4.1 *Suppose the family of Markov chains $\{X_n^\varepsilon\}$, $\varepsilon > 0$, satisfies Assumptions 2.1 and 3.1. Let ρ be a limit quasi-stationary distribution defined by (1.4). If $\mathcal{A}$ is the union of the periodic attractors of f, then*

$$\rho(I) = \rho(\mathcal{A}).$$

4.1 Outline of the proof of Theorem 4.1

We first provide a heuristic argument as to why we expect the result to hold. The next section contains a sequence of lemmas which culminate in the precise proof. For simplicity, in this description we consider only the case where the noise possesses a density with respect to Lebesgue measure, i.e. satisfies parts (1a), (2a) and (3) of Assumption 3.1.

Let $\delta, \gamma, L > 0$, $\eta > 1$ and m, $j < \infty$ be as defined in Lemma 2.4, and recall the definition

$$Z \doteq \{x : f^i(x) \notin W^{-\delta}, i = 0, 1, \ldots, j-1\}.$$

Since $W^{-\delta}$ is a finite union of open intervals and $Z = I \setminus \cup_{i=0}^{j-1} f^{-i}(W^{-\delta})$, Z is a finite union of closed intervals. We will refer to each of these closed intervals as a component of Z. Define $Y = I \setminus Z$. Then since $f(W^{-\delta}) \subset W^{-\delta}$, it is clear that for all $n \ge j$,

$$f^n(Y) \subset W^{-\delta}. \tag{4.1}$$

Let $\mathcal{A}$ be the finite union of the two-sided periodic attractors of f. The f-invariance property of ρ that was proved in Theorem 3.2, along with (4.1) and the fact that all f-invariant sets in Y are subsets of $\mathcal{A}$, leads to the conclusion that $\rho(Y) = \rho(\mathcal{A})$. Thus to prove Theorem 4.1 it suffices to show that $\rho(Y^c) = 0$. In fact, with $F \doteq W^{-\delta/2} \subset Y$, we will show that $\rho(\bar{F}^c) = 0$ by exploiting the properties of the pre-limit quasi-stationary distributions derived in Theorem 3.1.

We know from (3.7) that $\rho^\varepsilon \Rightarrow \rho$ as $\varepsilon \to 0$, and hence $\rho(\bar{F}^c) \leq \liminf_{\varepsilon\to 0} \rho^\varepsilon(\bar{F}^c)$ by the Portmanteau theorem. Furthermore, from (3.6) we know that for every $N \in I\!N$,

$$\rho^\varepsilon(\bar{F}^c) = (R^\varepsilon)^N \int_I P_x(X_N^\varepsilon \in \bar{F}^c)\rho^\varepsilon(dx).$$

Thus it is enough to show that for some $N = N(\varepsilon) \in I\!N$, the right hand side of the above equation goes to zero as $\varepsilon \to 0$. Using the decomposition $I = Y \cup Z$, from the above display we conclude that

$$\rho(\bar{F}^c) \leq \liminf_{\varepsilon\to 0} \rho^\varepsilon(\bar{F}^c) \leq \liminf_{\varepsilon\to 0}(R^\varepsilon)^N \left[\sup_{x\in Y} P_x(X_N^\varepsilon \in \bar{F}^c) + \sup_{x\in Z} P_x(X_N^\varepsilon \in \bar{F}^c).\right] \tag{4.2}$$

Since $R^\varepsilon \leq 1+e^{-c\varepsilon}$ by Theorem 3.1, it follows that for all $N(\varepsilon)$ of less than exponential order $\lim_{\varepsilon\to 0}(R^\varepsilon)^{N(\varepsilon)} = 1$. The first probability on the right hand side of (4.2) is shown in Theorem 4.4 to decay to zero as $\varepsilon \to 0$, using Lemmas 4.2 and 4.3. Lemma 4.2 uses the boundedness of the exponential moments of the noise to show that as $\varepsilon \to 0$, the Markov chain lies with probability approaching 1 in an arbitrarily small neighbourhood of the deterministic trajectory for any fixed number (that is independent of ε) of steps. By (4.1), there exists j such that $f^j(Y) \subset W^{-\delta}$. Hence if $X_0^\varepsilon \in Y$ then X_j^ε lies in $F = W^{-\delta/2}$ with probability approaching 1 as $\varepsilon \to 0$. In Lemma 4.3 we use the large deviation principle for the Markov chain $\{X_n^\varepsilon\}$ established in [17] to infer that there exists a $T < \infty$ such that when starting inside $W^{-\delta}$, the exit time from the region $F \subset W$ is greater than $e^{T/\varepsilon}$ with probability approaching 1. The Markov property is then used to show that if the chain starts in Y it is highly unlikely to be outside F after $N(\varepsilon)$ steps if $j < N(\varepsilon) < e^{T/\varepsilon}$. Therefore for $N(\varepsilon)$ in that range, the first term in (4.2) decays to zero as $\varepsilon \to 0$. Bounds on the second term in (4.2) are obtained in Theorem 4.8, which uses estimates for the time of exit from the "unstable" region Z obtained in Lemma 4.7. Lemma 4.7 in turn uses estimates on the rate of growth of the support of the noise as long as the process remains within the region $K^\gamma \cap Z$, that are derived in Lemmas 4.5 and 4.6. These two lemmas, which are at the heart of the proof of Theorem 4.1, exploit the expansive property (2.2) of the deterministic system derived in Lemma 2.4 as well as the lower bound (3.2) on the noise, and are related to an argument of G. Zohar [24].

The case of discrete noise taking values in the lattice $\mathcal{L}^\varepsilon$ (when (1b), (2b) and (3) of Assumption 3.1 hold) requires some technical modifications, most notably near the endpoints of I. This is taken care of in Theorem 4.11, which relies on bounds on the exit time from a neighbourhood of a hyperbolic repelling fixed point derived in Lemma 4.10.

4.2 Statement of Theorems and Proofs

We now present the lemmas and their rigorous proofs. The sets W, Z and $Y = I \setminus Z$ and the constants $\delta, \gamma, L > 0$, $\eta > 1$ and $j, m < \infty$ are chosen as in Lemma 2.4. Recall that $F = W^{-\delta/2}$.

For every $\varepsilon > 0$, $\{\mathcal{F}_n^\varepsilon\}$ denotes the filtration associated with the Markov chain $\{X_n^\varepsilon\}$. As usual, E_x and P_x denote the expectation and probability respectively, conditioned on starting at x.

Lemma 4.2 *Suppose (3) of Assumption 3.1 is satisfied. Then*

$$\lim_{\varepsilon\to 0} \sup_{x\in Y} P_x(X_j^\varepsilon \in F^c) = 0. \tag{4.3}$$

Proof. We first establish that for any $i \in I\!N$ and $\delta > 0$, there exists $\nu > 0$ such that

$$\lim_{\varepsilon \to 0} \sup_{x \in I} \sup_{y \in U_\nu(x)} P_y(|X_i^\varepsilon - f^i(x)| > \delta) = 0. \tag{4.4}$$

As shown below, this is a simple consequence of the Markov property of $\{X_n^\varepsilon\}$ and the condition (3.5) on the noise which guarantees that for any $c > 0$

$$\lim_{\varepsilon \to 0} \sup_{x \in I} P(|\xi^\varepsilon(x)| > c) = 0. \tag{4.5}$$

Since f is $\mathcal{C}^1$ it is Lipschitz continuous on I with some Lipschitz constant $b \in (0, \infty)$. Now choose $\nu = \delta/(2b)$. Then the dynamics (1.2) and the fact that $|f(y) - f(x)| < \delta/2$ for all $y \in U_\nu(x)$ imply that for such y,

$$\begin{aligned} P_y(|X_1^\varepsilon - f(x)| > \delta) &= P_y(|f(y) + \xi^\varepsilon(y) - f(x)| > \delta) \\ &\leq P_y(|f(y) - f(x)| + |\xi^\varepsilon(y)| > \delta) \\ &\leq P_y(|\xi^\varepsilon(y)| > \delta/2). \end{aligned} \tag{4.6}$$

Taking the supremum in (4.6) over $y \in U_\nu(x)$ and $x \in I$ and then taking limits as $\varepsilon \to 0$, the right hand side goes to zero due to (4.5). Thus (4.4) holds when $i = 1$.

Suppose (4.4) is true for $i = 1, 2, \ldots, k-1$. Then we show below that it is also true for $i = k$. The dynamics (1.2) and the Markov property show that for any $\tilde{\delta} > 0$

$$\begin{aligned} P_y(|X_k^\varepsilon - f^k(x)| > \delta) &= E_y[P_y(|X_k^\varepsilon - f^k(x)| > \delta | \mathcal{F}_{k-1}^\varepsilon)] \\ &= E_y[P_{X_{k-1}^\varepsilon}(|f(X_{k-1}^\varepsilon) + \xi^\varepsilon(X_{k-1}^\varepsilon) - f^k(x)| > \delta)] \\ &\leq P_y(|X_{k-1}^\varepsilon - f^{k-1}(x)| > \tilde{\delta}) \\ &\quad + E_y\left[P_{X_{k-1}^\varepsilon}(|f(X_{k-1}^\varepsilon) + \xi^\varepsilon(X_{k-1}^\varepsilon) - f^k(x)| > \delta) 1_{\{|X_{k-1}^\varepsilon - f^{k-1}(x)| \leq \tilde{\delta}\}}\right], \end{aligned}$$

which implies that

$$\begin{aligned} \sup_{y \in U_\nu(x)} P_y(|X_k^\varepsilon - f^k(x)| > \delta) \leq\; & \sup_{y \in U_\nu(x)} P_y(|X_{k-1}^\varepsilon - f^{k-1}(x)| > \tilde{\delta}) \\ & + \sup_{y \in U_{\tilde{\delta}}(f^{k-1}(x))} P_y(|f(y) + \xi^\varepsilon(y) - f(f^{k-1}(x))| > \delta). \end{aligned}$$

Choose $\tilde{\delta} \doteq \delta/(2b)$ and take the supremum over $x \in I$ and limits as $\varepsilon \to 0$ in the last display. By assumption, (4.4) holds for $i = k - 1$ and thus there exists $\nu > 0$ for which the first term goes to zero. The second term on the right hand side goes to zero by (4.6) for the case $i = 1$. This establishes (4.4) for $i = k$ and therefore, by induction, for all $i \in I\!N$. Recall (c.f. (4.1)) that $f^j(Y) \subset W^{-\delta}$, replace δ by $\delta/2$, i by j and y by x in (4.4) to get the desired result (4.3). ∎

Lemma 4.3 *Suppose (3) of Assumption (3.1) is satisfied. Define*

$$\sigma^\varepsilon \doteq \inf\{n > 0 : X_n^\varepsilon \in F^c\}. \tag{4.7}$$

Then there exists $T < \infty$ such that

$$\lim_{\varepsilon \to 0} \sup_{x \in F} P_x(\sigma^\varepsilon \leq e^{T/\varepsilon}) = 0. \tag{4.8}$$

Proof. The map f is generalized Axiom A, and therefore has only finitely many periodic attractors. Hence there exists $l < \infty$ (for example one can choose l to be the lowest common multiple of the periods of all the periodic attractors of f) such that the set $F = W^{-\delta/2}$ can be expressed as the union of intervals in the contracting basins of attraction of fixed points of f^l. Since the time of exit of the Markov chain $\{X_n^\varepsilon\}$ from the set F is larger than the time of exit of the chain from any subset of F, it suffices to derive an estimate of the form (4.8) for the exit time from any one of the intervals comprising F that lies in the contracting basin of a fixed point of f^l. Such estimates, for exit times from a single basin of attraction of a fixed point of a map g, were derived in [17, Lemma 2.2] using the large deviation principle for the chain $\{X_k^\varepsilon\}$ that was proved in [17, Lemma 2.1] and [14, Theorem 5.2 and Corollary 5.2]. The large deviation lemma in [17, Lemma 2.1] requires that g be Lipschitz continuous, which is certainly satisfied by the map f^l considered here. The lemma also requires that certain assumptions (A1), (A2) and (A5) on the Markov chain stated in [17] hold. By letting $Y_n^\varepsilon = X_{nl}^\varepsilon$ and $\phi^\varepsilon(\cdot) = X_l^\varepsilon - f^l(x)$, we see that $\{Y_n^\varepsilon\}$ is a Markov chain that satisfies (1.2) with X replaced by Y, ξ by ϕ and f by f^l. It is not hard to verify from standard large deviation arguments that the Markov chain $\{Y_n^\varepsilon\}$ defined above satisfies these assumptions because $\{X_n^\varepsilon\}$ satisfies (3) of Assumption 3.1. Consequently $\{Y_n^\varepsilon\}$ satisfies the large deviation principle. An estimate of the form (4.8) for the exit time of $\{Y_n^\varepsilon\}$ from the interval in the contracting basin of attraction of a fixed point of f^l can then be obtained from the large deviation principle for $\{Y_n^\varepsilon\}$ in the same way as in [17, Lemma 2.2]. This automatically yields the required estimate for the exit time of $\{X_n^\varepsilon\}$. ■

Theorem 4.4 *Suppose (3) of Assumption 3.1 is satisfied. Let T be as chosen in Lemma 4.3. Then*

$$\lim_{\varepsilon\to 0} \sup_{x\in Y} \sup_{j\le n\le e^{T/\varepsilon}} P_x(X_n^\varepsilon \in F^c) = 0.$$

Proof. Recall the definition of the stopping time $\sigma^\varepsilon \doteq \inf\{n \ge 0 : X_n^\varepsilon \in F^c\}$. Let $N = e^{T/\varepsilon}$. For any $x \in Y$, using the Markov property we see that

$$\begin{aligned}
\sup_{j\le n\le N} P_x(X_n^\varepsilon \in F^c) &\le \sup_{1\le n\le N} P_x(X_{n+j}^\varepsilon \in F^c) \\
&= \sup_{1\le n\le N} E_x[P_x(X_{n+j}^\varepsilon \in F^c|\mathcal{F}_j^\varepsilon)] \\
&= \sup_{1\le n\le N} E_x[P_{X_j^\varepsilon}(X_n^\varepsilon \in F^c)] \\
&= \sup_{1\le n\le N}\{E_x[P_{X_j^\varepsilon}(X_n^\varepsilon \in F^c)1_{F^c}(X_j^\varepsilon)] \\
&\quad +E_x[P_{X_j^\varepsilon}(X_n^\varepsilon \in F^c)1_F(X_j^\varepsilon)]\},
\end{aligned}$$

which implies that

$$\sup_{j\le n\le N} P_x(X_n^\varepsilon \in F^c) \le P_x(X_j^\varepsilon \in F^c) + \sup_{y\in F} P_y(\sigma^\varepsilon \le N). \tag{4.9}$$

Taking the supremum over $x \in Y$ and then the limit as $\varepsilon \to 0$, the first and second terms on the right hand side go to zero due to Lemmas 4.2 and 4.3 respectively. ■

In Theorem 4.8 we derive an upper bound for the time taken for the Markov chain $\{X_n^\varepsilon\}$ to exit an α-neighbourhood of the expansive region $Z \cap I_\theta$, where $\theta > 0$ satisfies certain conditions stated below. The theorem uses some estimates which we derive in Lemmas 4.5 and 4.6. In order to handle simultaneously both discrete and continuous noise, it is advantageous to introduce some additional notations. Let $\theta_0 > 0$ be chosen to satisfy (2.5) of Assumption 2.1. Recall that a consequence of Assumption 2.1 is that there exists $r > 1$ with the property that $|Df^2(s)| > r^2 > 1$ for all periodic points $s \in \partial I$. For the rest of this section we fix $\theta \in (0, \theta_0)$ to satisfy the following properties. For every periodic point $s \in \partial I$, $|Df^2(x)| > r^2 > 1$ for all $x \in U_{3\theta}(s)$. Moreover $\lambda(I) > (3G+3)\theta$, where $G = \sup_{x\in I}|f'(x)|$ and $\lambda(I)$ is the Lebesgue measure of I. Finally we choose $\theta < \delta/6$ so that $\bar{F} = \overline{W^{-\delta/2}} \subset I_{3\theta}$, and, when relevant, we further assume that $\theta < \kappa$ (see (3.4)). We let $Z_\theta \doteq Z \cap I_\theta$. In the lemmas and theorems that follow, the case of continuous density satisfying (1a), (2a) and (3) of Assumption 3.1 can be handled with Z_θ replaced throughout by Z, c.f. Remark 4.9 below.

Recall that the unstable region Z can be decomposed into a finite disjoint union $Z = \cup_{i=1}^S J_i$, where each component J_i is a closed interval. Also recall from (2.4) in Lemma 2.4 that $Z \subset K^{\gamma/2}$. Thus we can choose $\alpha \in (0, \theta/2)$ small enough so that

1. $Z^{2\alpha} \subset K^\gamma$.
2. $\overline{J_i^{2\alpha}} \cap \overline{J_j^{2\alpha}} = \emptyset$ for every $i \neq j$, $i,j \in \{1,\dots,S\}$.
3. $$d(\overline{f(Z_\theta^{2\alpha})}, I \setminus I_\theta) > (\theta' - \theta)/2, \tag{4.10}$$

   where $\theta' > \theta$ is such that $f(I_\theta) \subset I_{\theta'}$ as in (2.5) of Assumption 2.1.

By (2.2), the first condition on α ensures that f^m is uniformly expansive on $Z^{2\alpha}$. The second condition is imposed for convenience to ensure that each point in $Z^{2\alpha}$ belongs to the 2α-neighbourhood of a unique component of Z. Finally the last condition is used in Theorem 4.8, in conjunction with the exponential bound (3.5) on the noise, to guarantee that with exponentially high probability the Markov chain lies in I_θ at the time of exit from Z_θ^α.

For $k \in \mathbb{N}$, let $Y_k^\varepsilon \doteq X_{km}^\varepsilon(x)$ be the (km)th iterate of the Markov chain $\{X_n^\varepsilon\}$ defined in (1.2) and let the measure V^ε and transition kernel $\pi_{x,m}$ be as specified in (1) of Assumption 3.1. Then for every $x \in \mathbb{R}$ and $A \in \mathcal{B}(\mathbb{R})$,

$$P(Y_1^\varepsilon \in A | Y_0^\varepsilon = x) = \int_A \pi_{x,m}^\varepsilon(y) V^\varepsilon(dy).$$

Fix $\beta \in (0, 1/2)$ such that it satisfies (3.2) if (2a) of Assumption 3.1 holds, or is equal to $\beta(\theta - 2\alpha)$ which satisfies (3.3) (where $\theta, \alpha > 0$ are as chosen above) if (2b) of Assumption 3.1 is satisfied. The following lemma shows that $\pi_{x,m}^\varepsilon$ satisfies an estimate analogous to that satisfied by π_x^ε in (3.2) and (3.3). Let $g = f^m$ and define $G = \max_{x\in I}(|f'(x)|, |g'(x)|)$. Since $f \in \mathcal{C}_1$, $G < \infty$.

Lemma 4.5 *Suppose (1) and (2) of Assumption 3.1 are satisfied. Let g, Y_k^ε and π_x^ε be defined as above. Then there exists $\chi \in (0,1)$ such that for any $x \in Z_\theta^{2\alpha}$,*

$$\pi_{x,m}^\varepsilon(\cdot) \geq \frac{\chi}{\varepsilon} 1_{[g(x)-\frac{\varepsilon}{2}, g(x)+\frac{\varepsilon}{2}]}. \tag{4.11}$$

Proof. For notational convenience, we assume in the proof that $V^\varepsilon(dz) = dz$. The discrete case is handled in exactly the same way.

Fix $x \in Z_\theta^{2\alpha}$ and $X_0^\varepsilon = x$ and for $k \in \mathbb{N}$, let $h_k^\varepsilon = \pi_{x,k}^\varepsilon$ be the density of X_k^ε with respect to V^ε. Clearly $Y_1^\varepsilon(x)$ has density $\pi_{x,m}^\varepsilon(\cdot) = h_m^\varepsilon(\cdot)$. Let L be as defined in Lemma 2.4. We now show that if $x \in Z_\theta^{2\alpha}$ then for $k = 1, \dots, m$,

$$h_k^\varepsilon(\cdot) \geq \frac{1}{\varepsilon}\left(\frac{\beta L}{G}\right)^k 1_{[f^k(x)-\frac{\varepsilon}{2}, f^k(x)+\frac{\varepsilon}{2}]}. \tag{4.12}$$

By (2) of Assumption 3.1, the choice of β, the fact that $L < G$ and $h_1^\varepsilon(\cdot) = \pi_x^\varepsilon(\cdot)$, (4.12) holds for $k = 1$. Now suppose (4.12) holds for some $k < m$. Using the definition of the transition kernel and the estimates in (4.12) and (3.3) we obtain

$$\begin{aligned} h_{k+1}^\varepsilon(y) &= \int_{\mathbb{R}} \pi_z^\varepsilon(y) h_k^\varepsilon(z) dz \\ &\geq \left(\frac{\beta L}{G}\right)^k \frac{1}{\varepsilon} \int_{[f^k(x)-\frac{\varepsilon}{2}, f^k(x)+\frac{\varepsilon}{2}]} \pi_z^\varepsilon(y) dz. \\ &\geq \left(\frac{\beta L}{G}\right)^k \frac{\beta}{\varepsilon^2} \int_{[f^k(x)-\frac{\varepsilon}{2}, f^k(x)+\frac{\varepsilon}{2}]} 1_{[-\varepsilon,\varepsilon]}(y - f(z)) dz. \end{aligned}$$

Assume without loss of generality that $\varepsilon < \gamma/2$ and recall that by the choice of α, $Z_\theta^{2\alpha} \subset K^\gamma$. Then by the definition of G and the choice (2.3) of L, for all $x \in Z_\theta^{2\alpha}$ and $0 \leq k \leq m-1$, $L \leq |f'(z)| \leq G$ for $z \in [f^k(x) - \frac{\varepsilon}{2}, f^k(x) + \frac{\varepsilon}{2}]$ since $[f^k(x) - \frac{\varepsilon}{2}, f^k(x) + \frac{\varepsilon}{2}]$ is contained in a $\gamma/2$-fattening of $f^k(K^\gamma)$. The continuity of f' requires that f is in fact strictly monotone on the interval. Assume that f is increasing so that $f'(z) \geq L$ for all $z \in [f^k(x) - \frac{\varepsilon}{2}, f^k(x) + \frac{\varepsilon}{2}]$. Thus $f(f^k(x) + \frac{\varepsilon}{2}) - f(f^k(x)) > L\varepsilon/2$ and similarly $f(f^k(x) - \frac{\varepsilon}{2}) - f(f^k(x)) < -L\varepsilon/2$. Using this last property and the upper bound G on f', we substitute $w = y - f(z)$ in the last display to obtain

$$\begin{aligned} h_{k+1}^\varepsilon(y) &= \left(\frac{\beta L}{G}\right)^k \frac{\beta}{G\varepsilon^2} \int_{[y-f(f^k(x)+\frac{\varepsilon}{2}), y-f(f^k(x)-\frac{\varepsilon}{2})]} 1_{[-\varepsilon,\varepsilon]}(w) dw \\ &\geq \left(\frac{\beta L}{G}\right)^k \frac{\beta}{G\varepsilon^2} \int_{[y-\frac{L\varepsilon}{2}-f^{k+1}(x), y-f^{k+1}(x)+\frac{L\varepsilon}{2}]\cap[-\varepsilon.\varepsilon]} dw. \end{aligned}$$

Observing that for all $y \in [f^{k+1}(x) - \frac{\varepsilon}{2}, f^{k+1}(x) + \frac{\varepsilon}{2}]$, the length of the interval $[y - \frac{L\varepsilon}{2} - f^{k+1}(x), y - f^{k+1}(x) + \frac{L\varepsilon}{2}] \cap [-\varepsilon, \varepsilon]$ is greater than $L\varepsilon$ the last display shows that

$$h_{k+1}^\varepsilon(\cdot) \geq \left(\frac{\beta L}{G}\right)^k \frac{\beta L}{G\varepsilon} 1_{[f^{k+1}(x)-\frac{\varepsilon}{2}, f^{k+1}(x)+\frac{\varepsilon}{2}]}(\cdot) = \left(\frac{\beta L}{G}\right)^{k+1} \frac{1}{\varepsilon} 1_{[f^{k+1}(x)-\frac{\varepsilon}{2}, f^{k+1}(x)+\frac{\varepsilon}{2}]}(\cdot).$$

It is easy to verify that the same estimate would hold if f were strictly monotone decreasing with $-G \leq f'(z) \leq -L$ on the interval. The last inequality shows that (4.12) holds with k replaced by $k+1$. Thus by induction it is true for all $k \leq m$, and the lemma is established setting $\chi = (\beta L/G)^m$. ■

Let $\chi = (\beta L/G)^m$ be as in Lemma 4.5 and define $\tilde{\beta} \doteq \chi/4G$. For $k \in \mathbb{N}$ define the set $A_k^\varepsilon = A_k^\varepsilon(x)$ to be the maximal interval containing $g^k(x)$ such that

$$A_k^\varepsilon \subset \left\{ y : h_{km}^\varepsilon(y) > \frac{\chi \tilde{\beta}^{k-1}}{\varepsilon} \right\}, \tag{4.13}$$

and let $\nu_k^\varepsilon \doteq \lambda(A_k^\varepsilon)$ denote its length. Notice that A_k^ε and ν_k^ε are purely deterministic quantities that depend on the starting point x of the Markov chain, although this dependence is not denoted explicitly. The uniform expansiveness of the map $g = f^m$ on $Z^{2\alpha}$ and the estimate (3.3) on the noise lead one to expect that as long as A_k^ε is contained in $Z_\theta^{2\alpha}$ for $x \in Z_\theta^{2\alpha}$, its support must grow. The following lemma provides a lower bound for this growth. Recall that $\eta > 1$ was chosen to satisfy (2.2).

Lemma 4.6 *Suppose (1) and (2) of Assumption 3.1 are satisfied. Let $g = f^m$ and for $k \in \mathbb{N}$, let A_k^ε and ν_k^ε be defined as above. If for any $x \in Z_\theta^{2\alpha}$, $A_k^\varepsilon \subset Z_\theta^{2\alpha}$ for $k = 1, \ldots, n-1$, then*

$$\nu_k^\varepsilon \geq \eta^{k-1}\varepsilon + (k-1)\frac{\varepsilon}{2} \tag{4.14}$$

for $k = 1, \ldots, n$.

Proof. As in the previous lemma, we assume that $V^\varepsilon(dz) = dz$, the discrete case being handled similarly. Fix $x \in Z_\theta^{2\alpha}$. Recall from Lemma 4.5 that $\pi_{x,m}(\cdot) = h_m^\varepsilon(\cdot)$ is the density of the transition kernel for the chain $\{Y_n^\varepsilon(\cdot)\}$. Since (2) of Assumption 3.1 is satisfied, from the estimate (4.11) in Lemma 4.5 it follows that $\nu_1^\varepsilon \geq \varepsilon$. Hence (4.14) is satisfied for $k = 1$. Suppose (4.14) is satisfied for some $k < n$. Since A_k^ε is a closed interval contained in $Z_\theta^{2\alpha}$, by the choice of α there exists a unique component J of Z such that $A_k^\varepsilon \subset J^{2\alpha}$. From the definition of $\pi_{z,m}^\varepsilon(\cdot)$ the evolution of the densities of $Y_k^\varepsilon = X_{km}^\varepsilon$ is described by the equation

$$h_{(k+1)m}^\varepsilon(y) = \int_{\mathbb{R}} \pi_{z,m}^\varepsilon(y) h_{km}^\varepsilon(z)dz$$

which, by the definition of A_k^ε, satisfies

$$h_{(k+1)m}^\varepsilon(y) \geq \frac{\chi\tilde{\beta}^{k-1}}{\varepsilon}\int_{A_k^\varepsilon} \pi_{z,m}^\varepsilon(y)dz.$$

Since $A_k^\varepsilon \subset Z_\theta^{2\alpha}$, from (4.11) we infer that

$$h_{(k+1)m}^\varepsilon(y) \geq \frac{\chi^2\tilde{\beta}^{k-1}}{\varepsilon^2}\int_{A_k^\varepsilon} 1_{[-\frac{\varepsilon}{2},\frac{\varepsilon}{2}]}(y - g(z))dz. \tag{4.15}$$

Moreover, from the estimate (2.2) and the fact that $A_k^\varepsilon \subset J^{2\alpha} \subset K^{2\gamma}$, we see that $|Dg(x)| = |Df^m(x)| \geq \eta > 1$. Since Dg is continuous this implies that g must be monotone on the interval $J^{2\alpha}$. Assume that g is increasing so that it satisfies $Dg(x) \geq \eta > 1$ for every $x \in J^{2\alpha}$. Since A_k^ε is an interval containing $g^k(x)$, there exist $a, b \in (0,\infty)$ such that $A_k^\varepsilon = [g^k(x) - a, g^k(x) + b]$. Then $\nu_k^\varepsilon = a + b$ and the strict monotonicity of g shows that

$$g(A_k^\varepsilon) = g([g^k(x) - a, g^k(x) + b]) = [g(g^k(x) - a), g(g^k(x) + b)] \supset [g^{k+1}(x) - \eta a, g^{k+1}(x) + \eta b],$$

where the last inclusion follows from the Mean Value Theorem. Recall that $G \geq \sup_{x\in I} g'(x)$. Proceeding in a manner analogous to the proof of Lemma 4.5, substitute $w = y - g(z)$ in (4.15) and use the last display to obtain

$$h_{(k+1)m}^\varepsilon(y) \geq \frac{\chi^2\tilde{\beta}^{k-1}}{G\varepsilon^2}\int_{[y-g^{k+1}(x)-\eta b, y-g^{k+1}(x)+\eta a]\cap[-\frac{\varepsilon}{2},\frac{\varepsilon}{2}]} dw.$$

Note that for every $y \in [g^{k+1}(x) - \eta a - \varepsilon/4, g^{k+1}(x) + \eta b + \varepsilon/4]$, the length of the interval $[y - g^{k+1}(x) - \eta b, y - g^{k+1}(x) + \eta a] \cap [-\varepsilon/2, \varepsilon/2]$ is greater than $\varepsilon/4$. Thus substituting $\tilde{\beta} = \chi/(4G)$, we see that for all $y \in [g^{k+1}(x) - \eta a - \frac{\varepsilon}{4}, g^{k+1}(x) + \eta b + \frac{\varepsilon}{4}]$,

$$h_{(k+1)m}^\varepsilon(y) \geq \left(\frac{\chi\tilde{\beta}^{k-1}}{\varepsilon}\right)\left(\frac{\chi}{4G}\right) = \frac{\chi\tilde{\beta}^k}{\varepsilon}.$$

It is easy to check that parallel calculations yield the same result for the case when $Dg(x) \leq -\eta < -1$ for all $x \in A_k^\varepsilon$. Thus the above discussion shows us that

$$\begin{aligned} \nu_{k+1}^\varepsilon = \lambda(A_{k+1}^\varepsilon) &\geq \lambda([g^{k+1}(x) - \eta a - \tfrac{\varepsilon}{4}, g^{k+1}(x) + \eta b + \tfrac{\varepsilon}{4}]) \\ &\geq \eta(a+b) + \tfrac{\varepsilon}{2} = \eta\nu_k^\varepsilon + \tfrac{\varepsilon}{2} \geq \eta^k\varepsilon + k\tfrac{\varepsilon}{2}, \end{aligned}$$

where the last inequality follows from the initial assumption that ν_k^ε satisfies (4.14). By induction (4.14) holds now for all $k \leq n$ and the lemma is proved. ■

We now use the last two lemmas to estimate the time of exit of the Markov chain $\{X_n^\varepsilon\}$ from a neighbourhood Z_θ^α of the unstable region Z. Define the time of exit from this neighbourhood to be

$$\tau^\varepsilon \doteq \inf\{n > 0 : X_n^\varepsilon \notin Z_\theta^\alpha\}. \tag{4.16}$$

Since Z_θ^α is open, τ^ε is clearly a stopping time. Let $\mathcal{F}_{\tau^\varepsilon}$ be the associated stopped σ-field.

Lemma 4.7 *Suppose (2) and (3) of Assumption 3.1 are satisfied, and let* τ^ε *be defined by (4.16). Then there exist* $c_1, c_2' \in (0,\infty)$ *such that*

$$\lim_{\varepsilon\to 0} \sup_{x\in Z_\theta} P_x(\tau^\varepsilon > c_2'/\varepsilon^{c_1}) = 0. \tag{4.17}$$

Proof. Define $V = \max_{i=1,\ldots,S} \lambda(J_i^{2\alpha})$ to be the maximum length of any component of $Z^{2\alpha}$. Choose $M(\varepsilon)$ such that

$$M(\varepsilon) > \log\left(\frac{V}{\varepsilon}\right) / \log\eta.$$

Claim 1 : For any $\varepsilon > 0$, let $M \doteq M(\varepsilon)$ satisfy the inequality given above and define $\tilde{\chi} \doteq \chi\tilde{\beta}^M$, where χ and $\tilde{\beta}$ are as in (4.13). Then we have

$$\inf_{x\in Z_\theta^\alpha} P_x(\tau^\varepsilon < m(M(\varepsilon)+1)) \geq \tilde{\chi}.$$

Proof of Claim 1 : Fix $\varepsilon > 0$ and let M be chosen as above. We show below that there exists $k \leq M+1$ such that $A_k^\varepsilon \not\subset Z_\theta^{2\alpha}$, where A_k^ε is defined in (4.13). Notice that if $A_k^\varepsilon \subset Z_\theta^{2\alpha}$ for any k, then since A_k^ε is an interval, we must have $\nu_k^\varepsilon \leq V$. Suppose $A_k^\varepsilon \subset Z_\theta^{2\alpha}$ for every $k = 1,\ldots,M$. Then Lemma 4.6 implies that $\nu_{M+1}^\varepsilon \geq \eta^M\varepsilon + M\eta/2$, which by the choice of M implies that $\nu_{M+1}^\varepsilon > V$. Thus $A_{M+1}^\varepsilon \not\subset Z_\theta^{2\alpha}$. This shows us that there must be some $k \leq M+1$ such that $A_k^\varepsilon \not\subset Z_\theta^{2\alpha}$. Since $\nu_i^\varepsilon \geq \varepsilon$ for all $i \in \mathbb{N}$, in particular $\nu_k^\varepsilon \geq \varepsilon$ and since A_k^ε is not contained in $Z_\theta^{2\alpha}$ its intersection with $(Z_\theta^\alpha)^c$ must have length greater than ε (since we can always assume that $\varepsilon < \alpha$). In other words, $\lambda(A_k^\varepsilon \cap (Z_\theta^\alpha)^c) > \varepsilon$, and by the definition of A_k^ε and the fact that $\tilde{\beta} < 1$ we can conclude that $P_x(X_{km}^\varepsilon \notin Z_\theta^\alpha) \geq \tilde{\chi}$. Hence for any $x \in Z_\theta^\alpha$, and k chosen as above

$$P_x(\tau^\varepsilon \leq m(M+1)) = \Sigma_{i=1}^{m(M+1)} P(X_i^\varepsilon \notin Z_\theta^\alpha) \geq P(X_{km}^\varepsilon \notin Z_\theta^\alpha) \geq \tilde{\chi},$$

and the claim is proved.

Claim 2 : Let $c_1 = 4\log(\frac{1}{\beta})/\log(\eta) + 1$, and let $c_2' = 2mV^{c_1}/\chi^2\log\eta$. Then

$$\lim_{\varepsilon\to 0} \sup_{x\in Z_\theta} P_x(\tau^\varepsilon \geq c_2'/\varepsilon^{c_1}) = 0.$$

Proof of Claim 2 : Note that $c_1, c_2' \in (0,\infty)$ and choose $M \doteq M(\varepsilon) \doteq \log(\frac{V}{\varepsilon})/\log\eta + 1$. Then elementary algebraic manipulations show that for all sufficiently small ε,

$$c_2'/\varepsilon^{c_1} \geq m(M+1)(\chi\tilde{\beta}^M)^{-2}.$$

Let $s \doteq m(M+1)$ and $\tilde{\chi} \doteq \chi\tilde{\beta}^M$. Then, using the Markov property and the last display, we obtain that for any $x \in Z_\theta$

$$\begin{aligned}
P_x(\tau^\varepsilon \geq c_2'/\varepsilon^{c_1}) &\leq P_x(\tau^\varepsilon > c_2'/\varepsilon^{c_1}) \\
&\leq P_x(\tau^\varepsilon > s\tilde{\chi}^{-2}) \\
&= E_x[P_x(\tau^\varepsilon > s\tilde{\chi}^{-2}|\mathcal{F}_s)1_{\{\tau^\varepsilon \geq s\}}] \\
&= E_x[P_{X_s}(\tau^\varepsilon > s[\tilde{\chi}^{-2}-1])1_{\{\tau^\varepsilon \geq s\}}] \\
&\leq \sup\nolimits_{x\in Z_\theta^\alpha} P_x(\tau^\varepsilon > s[\tilde{\chi}^{-2}-1])(1-\inf\nolimits_{x\in Z_\theta^\alpha} P_x(\tau^\varepsilon < s)).
\end{aligned}$$

Taking the supremum over $x \in Z_\theta$, and using claim 1 this implies that

$$\sup_{x\in Z_\theta} P_x(\tau^\varepsilon \geq c_2'/\varepsilon^{c_1}) \leq (1-\tilde{\chi}) \sup_{x\in Z_\theta^\alpha} P_x(\tau^\varepsilon \geq s(\tilde{\chi}^{-2}-1)).$$

Iterating this procedure $\tilde{\chi}^{-2}$ times (where we assume without loss of generality that $\tilde{\chi}^{-2} \in I\!N$), we see that

$$\sup_{x\in Z_\theta} P_x(\tau^\varepsilon \geq c_2'/\varepsilon^{c_1}) \leq (1-\tilde{\chi})^{\tilde{\chi}^{-2}},$$

which goes to zero as $\varepsilon \to 0$ since $\tilde{\chi} = \chi\tilde{\beta}^M \to 0$ as $\varepsilon \to 0$. ■

Theorem 4.8 *Suppose that (2) and (3) of Assumption 3.1 are satisfied. Choose $T < \infty$ as in Lemma 4.3, and let c_1, $c_2' \in (0,\infty)$ be as in Lemma 4.7. Moreover let $c_2 = c_2' + j$. Then for all $N(\varepsilon) \in (c_2/\varepsilon^{c_1}, e^{T/\varepsilon})$*

$$\lim_{\varepsilon\to 0} \sup_{x\in Z_\theta} P_x(X^\varepsilon_{N(\varepsilon)} \in F^c) = 0.$$

Proof. Fix j as chosen in Lemma 4.2 and recall the definition (4.16) of τ^ε. We first assert that

$$\lim_{\varepsilon\to 0} \sup_{x\in Z_\theta} P_x(X^\varepsilon_{\tau^\varepsilon} \in Y^c) = 0. \tag{4.18}$$

Indeed note that $X^\varepsilon_{\tau^\varepsilon} = f(X^\varepsilon_{\tau^\varepsilon-1}) + \xi^\varepsilon(X^\varepsilon_{\tau^\varepsilon-1})$ and $X^\varepsilon_{\tau^\varepsilon-1} \in Z_\theta^{2\alpha}$. Since $\alpha \in (0,\theta/2)$ was chosen to satisfy (4.10), $d(f(X^\varepsilon_{\tau^\varepsilon-1}), I_\theta^c) > (\theta'-\theta)/2$. Thus

$$\sup_{x\in Z_\theta} P_x(X^\varepsilon_{\tau^\varepsilon} \in I_\theta^c) \leq \sup_{x\in I_{\theta-2\alpha}} P(\xi^\varepsilon(x) > (\theta'-\theta)/2).$$

By (3) of Assumption 3.1 this implies that

$$\lim_{\varepsilon\to 0} \sup_{x\in Z_\theta} P_x(X^\varepsilon_{\tau^\varepsilon} \in I_\theta^c) \leq \lim_{\varepsilon\to 0} \sup_{x\in I} P(\xi^\varepsilon(x) > (\theta'-\theta)/2) = 0.$$

Clearly $X^\varepsilon_{\tau^\varepsilon} \in (Z_\theta^\alpha)^c$. The assertion (4.18) then follows from the last display and the fact that $I_\theta \cap (Z_\theta^\alpha)^c \subset Y$. Therefore for any x in Z_θ and $N = N(\varepsilon)$ chosen as in the statement of the theorem, using the strong Markov property we obtain

$$\begin{aligned} P_x(X^\varepsilon_N \in F^c) &= E_x[P_x(X^\varepsilon_N \in F^c|\mathcal{F}_{\tau^\varepsilon})] \\ &\le E_x[P_x(X^\varepsilon_N \in F^c|\mathcal{F}_{\tau^\varepsilon})1_{\{\tau^\varepsilon \le N-j\}}] + P_x(\tau^\varepsilon > N-j) \\ &\le E_x[P_{X^\varepsilon_{\tau^\varepsilon}}(X^\varepsilon_{N-\tau^\varepsilon} \in F^c)1_{\{\tau^\varepsilon \le N-j\}}] + P_x(\tau^\varepsilon > N-j) \\ &\le \sup_{y\in Y}\sup_{j\le k\le N} P_y(X^\varepsilon_k \in F^c) + P_x(X^\varepsilon_{\tau^\varepsilon} \in Y^c) + P_x(\tau^\varepsilon > N-j). \end{aligned}$$

Taking the supremum over all $x \in Z_\theta$ and limits as $\varepsilon \to 0$, the first term goes to zero by Theorem 4.4 since $N < e^{T/\varepsilon}$, while the second term goes to zero by (4.18) and the third term goes to zero by Lemma 4.7 since $c_2'/\varepsilon^{c_1} < N - j < e^{T/\varepsilon}$. ■

Remark 4.9 *In the case when (1a), (2a) and (3) of Assumption 3.1 are satisfied, the condition (3.3) holds on the whole of I. Thus in that case the arguments used in Lemmas 4.5, 4.6 and 4.7, and Theorem 4.8 hold with Z_θ replaced by Z and $Z_\theta^{2\alpha}$ replaced by $Z^{2\alpha} \cap I$. Let T be chosen as in Lemma 4.3. Then in this case too we have the result that there exist c_1, $c_2' \in (0,\infty)$ such that for $N(\varepsilon) \in (c_2'/\varepsilon^{c_1}, e^{T/\varepsilon})$,*

$$\lim_{\varepsilon\to 0}\sup_{x\in Z} P_x(X^\varepsilon_{N(\varepsilon)} \in F^c) = 0.$$

It only remains to consider the behaviour of the chain $\{X^\varepsilon_n\}$ with initial conditions in the region $I \setminus I_\theta$ when (1b), (2b) and (3) of Assumption 3.1 hold. Suppose $s_1 \in \partial I$ is a fixed point of f. In Lemma 4.10 we use the fact that s_1 is hyperbolic repelling to derive estimates on the exit time of the chain from a neighbourhood of s_1. In what follows T is as chosen in Lemma 4.3, and c_1, c_2 and θ are as in Theorem 4.8. Recall that $I(\varepsilon) = I^\circ \cap \mathcal{L}^\varepsilon$.

Lemma 4.10 *Suppose (1b), (2b) and (3) of Assumption 3.1 are satisfied. Define*

$$\tilde\tau^\varepsilon \doteq \inf\{n : X^\varepsilon_n \notin U_{2\theta}(s_1)\}. \tag{4.19}$$

Then there exist c_3, $c_4 \in (0,\infty)$ such that

$$\lim_{\varepsilon\to 0}\sup_{x\in U_\theta(s_1)\cap I(\varepsilon)} P_x(\tilde\tau^\varepsilon > c_3/\varepsilon^{c_4}) = 0.$$

Proof. Due to (2.5) of Assumption 2.1 we know that the fixed point s_1 is hyperbolic repelling and by the choice of θ we know that there exists $r > 1$ such that $f'(x) > r > 1$ for all $x \in U_{3\theta}(s_1)$. Here we have assumed without loss of generality that s_1 is the left end point of I. Suppose $x \in I(\varepsilon) \cap U_\theta(s_1)$ and recall the definition $w_\varepsilon = \inf\{d(x,s_1) : x \in I^\circ \cap \mathcal{L}^\varepsilon\}$. If $f^i(x) \in U_{3\theta}(s_1)$ for $i = 0,1,\dots,n-1$, then by the Mean Value theorem

$$f^n(x) - s_1 = f^n(x) - f(s_1) = f(f^{n-1}(x)) - f(s_1) > r(f^{n-1}(x) - s_1).$$

Iterating this procedure n times we have

$$f^n(x) - s_1 > r^n(x - s_1) > r^n w_\varepsilon,$$

where the last inequality follows because $x \in I^\circ \cap \mathcal{L}^\varepsilon$. Now let $M(\varepsilon) = \log(3\theta/w_\varepsilon)/\log r$ so that $f^{M(\varepsilon)}(x) - s_1 > 3\theta$. This implies that if the noise $\xi^\varepsilon(X^\varepsilon_i)$ is positive for $i = 0,1,\dots,M(\varepsilon)-1$, then $X_{M(\varepsilon)} \notin U_{2\theta}(s_1)$. Let $\kappa > 0$ be such that (3.4) of Assumption 3.1 is satisfied, and note that θ was chosen to be less than κ. Then for $x \in U_\theta(s_1) \cap I(\varepsilon) \subset U_\kappa(s_1) \cap I^\circ$, we have

$$\begin{aligned} P_x(\tilde{\tau}^\varepsilon \leq M(\varepsilon)) &\geq P_x(\xi^\varepsilon(X_i^\varepsilon) > 0, \text{ for } i = 0, 1, \ldots, M(\varepsilon) - 1) \\ &\geq [\inf_{x \in I^\circ} P(\xi^\varepsilon(x) > 0)]^{M(\varepsilon)} \\ &\geq \kappa^{M(\varepsilon)}. \end{aligned}$$

Since by (2b) of Assumption 3.1, $w_\varepsilon \to 0$ as $\varepsilon \to 0$, it follows that $\kappa^{M(\varepsilon)} \to 0$ as $\varepsilon \to 0$. Therefore, analogous to the proof of Claim 2 in Lemma 4.7, by the Markov property we conclude that

$$\lim_{\varepsilon \to 0} \sup_{x \in U_\theta(s_1) \cap I(\varepsilon)} P_x\left(\tilde{\tau}^\varepsilon > M(\varepsilon)\kappa^{-2M(\varepsilon)}\right) \leq \lim_{\varepsilon \to 0} \left(1 - \kappa^{M(\varepsilon)}\right)^{\kappa^{-2M(\varepsilon)}} = 0.$$

Substituting $\log(3\theta/w_\varepsilon)/\log r$ for $M(\varepsilon)$ and using the fact from (2b) of Assumption 3.1 that w_ε decays to zero at most polynomially fast with respect to ε, we see that there exist constants c_3, $c_4 \in (0, \infty)$ such that $M(\varepsilon)\kappa^{-2M(\varepsilon)} \leq c_3/\varepsilon^{c_4}$, which proves the lemma. ■

Theorem 4.11 *Suppose (1b), (2b) and (3) of Assumption 3.1 are satisfied. Moreover choose* c_1, $c_2 \in (0, \infty)$ *as in Theorem 4.8 and* c_3, $c_4 \in (0, \infty)$ *as in Lemma 4.10. Then for* $N(\varepsilon) \in (2c_2/\varepsilon^{c_1} + c_3/\varepsilon^{c_4}, e^{T/\varepsilon})$,

$$\lim_{\varepsilon \to 0} \sup_{x \in I(\varepsilon) \setminus I_\theta} P_x(X_{N(\varepsilon)}^\varepsilon \in U_{2\theta}(\partial I)) = 0.$$

Proof. Suppose s_1 is a fixed point of f in ∂I and s_2 is the other end point of I. Let $N = N(\varepsilon) \in (2c_2/\varepsilon^{c_1} + c_3/\varepsilon^{c_4}, e^{T/\varepsilon})$ and choose $\tilde{N} = \tilde{N}(\varepsilon) = c_3/\varepsilon^{c_4} + 1$ so that $N(\varepsilon) - \tilde{N}(\varepsilon) > c_2/\varepsilon^{c_1}$. Finally let $\tilde{\tau}^\varepsilon$ be as defined in (4.19). Then by the Markov property, for $x \in U_\theta(s_1) \cap I(\varepsilon)$

$$\begin{aligned} P_x(X_N^\varepsilon \in U_{2\theta}(\partial I)) &\leq P_x(X_{\tilde{\tau}^\varepsilon}^\varepsilon \in U_\theta(s_2)) + P_x(\tilde{\tau}^\varepsilon > \tilde{N}(\varepsilon)) \\ &\quad + \sup_{N(\varepsilon) - \tilde{N}(\varepsilon) \leq n \leq N(\varepsilon)} \sup_{x \in I_\theta \cap \mathcal{L}^\varepsilon} P_x(X_n^\varepsilon \in U_{2\theta}(\partial I)). \end{aligned} \tag{4.20}$$

Now recall $G = \sup_{x \in I} |f'(x)|$ and therefore by choice of $\theta > 0$, for $x \in U_{2\theta}(s_1)$, $f(x) \in U_{2G\theta}(s_1)$ and $\lambda(I) - (2G + 2)\theta > \theta$, where $\lambda(I)$ is the Lebesgue measure of I. Thus for $x \in U_\theta(s_1) \cap I(\varepsilon)$, since $X_{\tilde{\tau}^\varepsilon - 1}^\varepsilon \in U_{2\theta}(s_1)$,

$$\begin{aligned} P_x(X_{\tilde{\tau}^\varepsilon}^\varepsilon \in U_{2\theta}(s_2)) &= P_x(f(X_{\tilde{\tau}^\varepsilon - 1}^\varepsilon) + \xi^\varepsilon(X_{\tilde{\tau}^\varepsilon - 1}^\varepsilon) \in U_{2\theta}(s_2)) \\ &\leq \sup_{x \in I} P(\xi^\varepsilon(x) > \lambda(I) - (2G + 2)\theta) \\ &\leq \sup_{x \in I} P(\xi^\varepsilon(x) > \theta), \end{aligned}$$

which tends to zero as $\varepsilon \to 0$ due to (3) of Assumption 3.1. This shows that taking the supremum over all $x \in U_\theta(s_1) \cap I(\varepsilon)$ and limits as $\varepsilon \to 0$ in (4.20) the first term goes to zero. The second term goes to zero due to Lemma 4.10 and the last term goes to zero by Theorems 4.4 and 4.8. This concludes the proof for the case when both end points of ∂I are fixed points of f.

Since $f \in \mathcal{C}^1[I, I]$, the only other possibilities are that the end points form a periodic orbit of period two, or that one end point is mapped on to the other fixed end point of f. For the case when there is a periodic orbit of period two in ∂I, one can set $g = f^2$ and define the chain $Y_n^\varepsilon = X_{2n}^\varepsilon$ so that

$$Y_n^\varepsilon = g(Y_{n-1}^\varepsilon) + \phi^\varepsilon(Y_{n-1}^\varepsilon),$$

where

$$\phi^\varepsilon(x) \doteq f(f(x)+\xi^\varepsilon(x)) + \xi^\varepsilon(f(x)+\xi^\varepsilon(x)) - f^2(x).$$

In order to verify that the noise $\phi^\varepsilon(\cdot)$ corresponding to the Markov chain $\{Y_n^\varepsilon\}$ satisfies (2b) of Assumption 3.1, one needs to show that $\phi^\varepsilon(\cdot)$ satisfies (3.4) in neighbourhoods of both end points of I (since they are both fixed points of $g = f^2$ and consequently of g^2). This can be done by using the fact that ξ^ε satisfies (3.4) at both end points of I (since $\{X_n^\varepsilon\}$ satisfies (2b) of Assumption 3.1), along with the monotonicity of f in a sufficiently small neighourhood of the end points of I, which follows from (2.5). Hence an estimate of the form derived in Lemma 4.10 can also be obtained for the exit time from the neighbourhood $I \setminus I_{2\theta}$ of the periodic orbit. Since in this case it is automatic that $Y_{\tilde{\tau}^\varepsilon}^\varepsilon \notin U_{2\theta}(\partial I)$, the theorem follows from (4.20).

The last case when $f(s_2) = f(s_1) = s_1$ is also dealt with in a similar fashion. We omit the details here. ■

We now prove the main theorem which was stated at the beginning of the section.

Proof of Theorem 4.1. Let $T < \infty$ be as chosen in Lemma 4.3. By Theorem 3.1 we know that ρ exists, and we also know that there exists $T' \leq T$ such that $(R^\varepsilon)^{N(\varepsilon)} \to 1$ for all $N(\varepsilon) < e^{T'/\varepsilon}$. Suppose the Markov chain $\{X_n^\varepsilon\}$ satisfies (1a), (2a) and (3) of Assumption 3.1. Choose c_1, c_2 and c_3, c_4 as in Theorem 4.8 and Lemma 4.10 respectively. Then let $N = N(\varepsilon)$ be such that $N \in (c_2/\varepsilon^{c_1}, e^{T'/\varepsilon})$, and recall the expression (4.2) given in Section 4.1. Since $(R^\varepsilon)^{N(\varepsilon)} \to 1$, the first term on the right hand side of (4.2) is zero by Theorem 4.4 and the second supremum by Remark 4.9. From this we infer that $\rho(I) = \rho(\bar{F})$.

Now suppose $\{X_n^\varepsilon\}$ satisfies (1b), (2b) and (3) of Assumption 3.1. Then $I(\varepsilon) = I^\circ \cap \mathcal{L}^\varepsilon$, and, in analogy with (4.2), we use the decomposition $I = Y \cup Z_\theta \cup [I \setminus I_\theta]$ to obtain

$$\rho^\varepsilon(\bar{F}^c) \leq (R^\varepsilon)^N \left[\int_Y P_x(X_N^\varepsilon \in \bar{F}^c)\rho^\varepsilon(dx) + \int_{Z_\theta} P_x(X_N^\varepsilon \in \bar{F}^c)\rho^\varepsilon(dx) + \int_{I(\varepsilon)\setminus I_\theta} P_x(X_N^\varepsilon \in \bar{F}^c)\rho^\varepsilon(dx) \right]. \tag{4.21}$$

Replacing $\bar{F}^c$ by $U_{2\theta}(\partial I)$ in (4.21), for $N = N(\varepsilon) \in (2c_2/\varepsilon^{c_1} + c_3/\varepsilon^{c_4}, e^{T'/\varepsilon})$ we have

$$\liminf_{\varepsilon\to 0} \rho^\varepsilon(U_{2\theta}(\partial I)) \;\leq\; \liminf_{\varepsilon\to 0} (R^\varepsilon)^N \left[\sup_{x\in I_\theta} P_x(X_N^\varepsilon \in U_{2\theta}(\partial I)) + \sup_{x\in I(\varepsilon)\setminus I_\theta} P_x(X_N^\varepsilon \in U_{2\theta}(\partial I)) \right].$$

Since $U_{2\theta}(\partial I) \subset F^c$ by the choice of θ, and $(R^\varepsilon)^{N(\varepsilon)} \to 1$ as $\varepsilon \to 0$, the first term on the right in the last display goes to zero by Theorems 4.4 and 4.8, and the last term decays to zero by Theorem 4.11. Thus $\liminf_{\varepsilon\to 0} \rho^\varepsilon(U_{2\theta}(\partial I)) = 0$. Now taking limits as $\varepsilon \to 0$ in (4.21), and recalling that $\rho(\bar{F}^c) \leq \liminf_{\varepsilon\to 0} \rho^\varepsilon(\bar{F}^c)$ by the Portmanteau theorem, we conclude that

$$\begin{aligned}
\rho(\bar{F}^c) &\leq \liminf_{\varepsilon\to 0} (R^\varepsilon)^N \left[\int_{I_\theta} P_x(X_N^\varepsilon \in F^c)\rho^\varepsilon(dx) + \rho^\varepsilon(U_{2\theta}(\partial I)) \right] \\
&\leq \liminf_{\varepsilon\to 0} \left[(R^\varepsilon)^N \sup_{x\in Y\cup Z_\theta} P_x(X_N^\varepsilon \in F^c) \right] \\
&= 0,
\end{aligned}$$

where the last equality follows from Theorems 4.4 and 4.8.

Therefore we have shown that if $\{X_n^\varepsilon\}$ satisfies Assumptions 2.1 and 3.1 then $\rho(I) = \rho(\bar{F})$. The f-invariance of ρ proved in Theorem 3.2 and the fact that the finite union $\mathcal{A}$ of the periodic attractors of f is the only f-invariant subset of $\bar{F}$ shows that $\rho(\bar{F}) = \rho(\mathcal{A})$, which proves the theorem. ■

5 Applications

In this section we apply Theorem 4.1 to Markov chains obtained as perturbations of the logistic map $f(x) = \mu x(1-x)$ on $[0,1]$ for the set $A \subset (1,4)$ of parameter values μ for which the map is generalized Axiom A. As mentioned in the introduction to this paper the set A is dense in $(1,4)$ [4, p. 223]. The fact that $\mu \in (1,4)$ ensures that $f'(0) = \mu > 1$ and $\sup_{x\in I} f(x) < 1$ and thus for every $\mu \in A$ the corresponding f satisfies Assumption 2.1. Throughout this section f will always be the logistic map with parameter value $\mu \in A$.

5.1 Additive Normal Noise

Consider the Markov chain obtained by perturbing f by additive normal noise. More precisely define $\{X_n^\varepsilon\}$ by $X_0^\varepsilon = x$ and for $n = 0, 1, \ldots,$

$$X_{n+1}^\varepsilon = f(X_n^\varepsilon) + \varepsilon\psi_n, \tag{5.1}$$

where ψ_n are i.i.d. standard normal random variables. We now show that this Markov chain satisfies Assumption 3.1. Since ψ_n is a $\mathcal{N}(0,1)$ random variable, $\varepsilon\psi_n$ has density with respect to Lebesgue measure and the transition kernel π_x^ε of the chain $\{X_n^\varepsilon\}$ satisfies condition (1) since for every $x, y \in I$,

$$\frac{1}{\sqrt{2\pi}\varepsilon}e^{-1/\varepsilon^2} \leq \pi_x^\varepsilon(y) \leq \frac{1}{\sqrt{2\pi}\varepsilon}.$$

Condition (2) holds because with $\beta = e^{-2}/\sqrt{2\pi}$, for all $x \in I$ and for all $y \in [f(x)-\varepsilon, f(x)+\varepsilon]$,

$$\pi_x^\varepsilon(y) \geq \frac{\beta}{\varepsilon}.$$

Direct computations also verify that condition (3) of Assumption 3.1 is satisfied, and thus we conclude from Theorem 4.1 that the limit quasi-stationary distribution of the family of Markov chains satisfying (5.1) is supported on the union of the periodic attractors of f.

5.2 Density Dependent Branching Processes

We now apply Theorem 4.1 to the model of density dependent branching processes that was considered in [12, 17, 18]. Following the definition given in [17], for all $x \geq 0$ let $Y(x)$ be a non-negative integer-valued random variable and for $x \geq 1$ let $Y(x) = 0$. $Y(x)$ represents the law of the offspring distribution when the population density is x. For $x \geq 0$, let $Y_{j,n}(x)$ be i.i.d. random variables distributed with the same distribution as $Y(x)$. Let K be an integer threshold value that represents the maximum population in the system. Then for $K \in [2, \infty)$ choose Z_0^K to be an integer less than K and define a population density branching process $\{Z_n^K\}, n \in \mathbb{N}$, iteratively by

$$Z_{n+1}^K = \begin{cases} \left(\Sigma_{j=1}^{Z_n^K} Y_{j,n+1}\left(\frac{Z_n^K}{K}\right)\right) \wedge K & \text{if } Z_n^K > 0, \\ 0 & \text{if } Z_n^K = 0, \end{cases}$$

where we assume that for any fixed x, K and n that $Y_{j,n+1}$, $j = 1, 2, \ldots,$ are independent of $Z_n^K, Z_{n-1}^K, \ldots, Z_0^K$. Let $l(x) = EY(x)$, and define $\tilde{Y}(x) \doteq Y(x) - l(x)$ to be the centered offspring distribution. Also define $X_n^K = Z_n^K/K$ and let $\varepsilon = 1/\sqrt{K}$. We make the following assumption on the offspring distribution.

Assumption 5.1 *The offspring distribution of* $Y(x)$ *satisfies the following properties.*

1. *There exists* $\mu \in A$ *such that for* $x \in [0,1]$, $EY(x) = \mu(1-x)$. *Moreover* $\sigma^2(x) \doteq Var(Y(x)) = E[\tilde{Y}^2(x)]$ *is bounded away from zero for* x *in any compact subset of* $(0,1)$.

2. *Let* $\mathcal{L} \doteq [\cup_{\varepsilon>0}\mathcal{L}^\varepsilon] \cap I^\circ$, *the set of all rationals in* I°. *Then* $0 < \inf_{x\in\mathcal{L}} P(Y(x) = 1|Y(x) \neq 0) < \sup_{x\in\mathcal{L}} P(Y(x) = 1|Y(x) \neq 0) < 1$.

3. *There exists* $\lambda_0 > 0$ *such that for all* $|\lambda| < \lambda_0$,

$$\sup_{x\in I, \varepsilon>0} \frac{x}{\varepsilon^2} \log E \exp(\lambda\varepsilon\tilde{Y}(x)) < \infty.$$

Note that for each $\varepsilon > 0$ such that $1/\varepsilon^2 \in I\!N$, X_n^ε has state space $\mathcal{L}^\varepsilon \doteq \{i\varepsilon^2, i = 0, \dots, \varepsilon^{-2}\}$.

It is easy to verify that when the offspring distribution satisfies Assumption 5.1, the Markov chain $\{X_n^\varepsilon\}$ satisfies

$$X_{n+1}^\varepsilon = f(X_n^\varepsilon) + \xi^\varepsilon(X_n^\varepsilon), \tag{5.2}$$

where for x, ε such that x/ε^2 is an integer, we define

$$\xi^\varepsilon(x) \doteq \varepsilon^2 \Sigma_{j=1}^{x/\varepsilon^2} \tilde{Y}_j(x).$$

Let $v(x) = E(|\tilde{Y}(x)|^3)$, $\sigma^2(x) = E[\tilde{Y}^2(x)]$, and note that Assumption 5.1 implies that for any $\theta > 0$,

$$\sup_{x\in I_\theta} \frac{v(x)}{\sigma^3(x)\sqrt{x}} < \infty.$$

Part (3) of Assumption 5.1 implies that the distribution of $\xi^\epsilon(x)$ is not concentrated on a sub-lattice of $\mathcal{L}^\varepsilon$. Hence, by the Berry-Esseen theorem, $\xi^\epsilon(x)/\varepsilon$ converges in distribution, uniformly in $x \in I_\theta$, to a non-degenerate Normal variable having mean 0 and variance $x\sigma^2(x)$. Further, by the local CLT for lattice distributions (see, e.g., [6, Theorem 2, p. 540]), the convergence extends to pointwise uniform convergence of the density on lattice points.

Theorem 5.1 *Suppose the offspring distribution* $Y(x)$ *satisfies Assumption 5.1, then the limit quasi-stationary distribution of the associated family of Markov chain* $\{X_n^\varepsilon\}, \varepsilon > 0$, *described above is supported on the finite union of periodic attractors of* $f(x) = xEY(x)$.

Proof. As shown in the previous section, for $\mu \in A$, f satisfies Assumption 2.1. Part (1b) of Assumption 3.1 is trivial, whereas for part (2b) one may take $a_1 = 2$ and then apply part (3) of Assumption 5.1 and the Berry-Esseen and local CLT theorems mentioned above in order to check (3.3). Note that $0 \in \partial I$ is the only fixed point of $f^2(\cdot)$ in ∂I. By yet another application of the Berry-Esseen theorem,

$$P(\xi^\varepsilon(x) > 0) = P\left(\frac{\xi^\varepsilon(x)}{\varepsilon} > 0\right) \to_{\varepsilon\to 0} \frac{1}{2},$$

as long as $x/\epsilon^2 \to \infty$. Hence, there exists a M large enough such that (3.4) holds (with κ a function of M) as soon as $1/2\epsilon^2 > x/\epsilon^2 > M$ whereas, for smaller values of x, (3.4) is an immediate application (reducing further κ, if necessary) of condition (3) in Assumption 5.1. Thus Assumption 3.1 is satisfied for $\{X_n^\varepsilon\}$ and therefore Theorem 5.1 follows from Theorem 4.1. ■

Corollary 5.2 *Let $Y(x)$ be a Poisson process with rate $\mu(1-x)$ for some $\mu \in A$. Then the quasi-stationary distribution of the chain $\{X_n^\varepsilon\}$ defined by (5.2) is supported on the periodic attractors of $f(x) = \mu x(1-x)$.*

Proof. By Theorem 5.1 it suffices to show that the Poisson process Y satisfies Assumption 5.1. This elementary fact (see e.g. [18, Section 3] for a related explanation) is left to the reader. ■

6 An Open Problem

Consider the deterministic dynamical system defined in (2.1). Generalizing the notion of the basin (of attraction) of a periodic orbit that was defined in Section 2, for any set $S \subset I$ we define its basin of attraction to be $B(S) \doteq \{x \in I : f^n(x) \to S \text{ as } n \to \infty\}$. Then we define a (topological) attractor of the dynamical system (1.1) to be a forward invariant set $\mathcal{A}$ such that the closure $\overline{B(\mathcal{A})}$ of its basin of attraction contains intervals, and such that each closed forward invariant subset $\mathcal{A}'$ which is strictly contained in $\mathcal{A}$ has a smaller basin of attraction : $\overline{B(\mathcal{A})} \setminus \overline{B(\mathcal{A}')}$ contains intervals [4, p. 236]. An attractor in a one-dimensional dynamical system takes one of three forms. It is either periodic, soleniodal or is a finite union of intervals on which the map is transitive [4, Theorem III.4.1]. In this paper we derived conditions under which the quasi-stationary distribution is concentrated on the union of the attractors when the attractors are periodic and two-sided. A thorough understanding of the dynamics of the underlying deterministic system (as characterized in Lemma 2.4) was crucial to this characterization. In the case of a unimodal map with a two-sided attracting periodic orbit, almost all trajectories of the deterministic system tend to the unique stable periodic orbit. In other words $B(s) = I$ upto a set of Lebesgue measure zero. This continues to hold when the stable periodic orbit is one-sided since the dynamics in the presence of a one-sided periodic attractor closely parallels that in the presence of a two-sided periodic attractor [10, 2]. Thus we expect in this case too, that the support of the quasi-stationary distribution will lie in the one-sided periodic attractor. However, since our methods heavily rely on the fact that the basin of attraction B is open, they do not seem to easily extend to the case when there is a one-sided periodic attractor. It would be even more challenging to characterize the support of the quasi-stationary distribution in situations where the underlying deterministic dynamics is even more complicated as in the case when there exist attractors that are not periodic.

References

[1] M. Benedicks and L. S. Young. Absolutely continuous invariant measures and random perturbations for certain one-dimensional maps. *Ergod. Th. and Dyn. Syst.*, 12:13-37, 1992.

[2] P. Collett and J. P. Eckmann. *Iterated Maps on the Interval as a Dynamical System.* Birkhauser, Boston, 1980.

[3] R. L. Devaney. *An Introduction to Chaotic Dynamical Systems.* Addison-Wesley, Redwood City, CA, 2nd ed., 1989.

[4] W. de Melo and S. Van Strein. *One-dimensional Dynamics.* Springer Verlag, Berlin, 1993.

[5] N. Dunford and J. T. Schwartz. *Linear Operators Part III : Spectral Operators* Wiley-Interscience, 1971.

[6] W. Feller. *An Introduction to Probability Theory and its Applications. Vol. II.* Wiley, New York, 1971.

[7] P. A. Ferrari, S. Martinez and P. Picco. Existence of non-trivial quasi-stationary distributions in the birth-death chain. *Adv. Appl. Prob.*, 24:795-813, 1992.

[8] P. A. Ferrari, H. Kesten and P. Picco. R-positivity, quasi-stationary distributions and ratio limit theorems for a class of probabilistic automata. *Ann. Appl. Prob.*, 6:577–616, 1996.

[9] P. Gora. Random composing of mappings, small stochastic perturbations, and attractors. *Z. Wahr. verw. Geb.*, 69:137–160, 1985.

[10] J. Guckenheimer. Sensitive dependence on initial conditions for one-dimensional maps. *Comm. Math. Phys.*, 70:133-160, 1979.

[11] M. P. Hassell. Density-dependence in single-species populations. *Journal of Animal Ecology*, 44:283-296, 1974.

[12] G. Hognas. On the quasi-stationary distribution of a stochastic Ricker model. *Stoch. Proc. Appl.*, 70:243-263, 1997.

[13] Y. Kifer. Principal eigenvalues, topological pressure, and stochastic stability of equilibrium states. *Israel J. Math.*, 70:1-47, 1990.

[14] Y. Kifer. *Random Perturbations of Dynamical Systems.* Birkhauser, Boston 1988.

[15] F. C. Klebaner. Population-dependent branching processes with a threshold. *Stoch. Proc. Appl.*, 46:115-127, 1993.

[16] F. C. Klebaner and O. Nerman. Autoregressive approximation in branching processes with a threshold. *Stoch. Proc. Appl.*, 51:1-7, 1994.

[17] F. C. Klebaner and O. Zeitouni. The exit problem for a class of density-dependent branching systems. *Ann. Appl. Prob.*, 4:1188-1205, 1994.

[18] F. C. Klebaner, J. Lazar and O. Zeitouni. On the quasi-stationary distribution for some randomly perturbed transformations of an interval. *Ann. Appl. Prob.*, 8:300:315, 1998.

[19] R. M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459-467, 1976.

[20] D. Ruelle. Small random perturbations of dynamical systems and the definition of attractor, *Comm. Math. Phys.*, 82:137–151, 1981.

[21] E. Seneta and D. Vere-Jones. On quasi-stationary distribution in discrete time Markov chains with a denumerable infinity of states. *Jour. Appl. Prob.*, 3:403–434, 1966.

[22] D. Vere-Jones. Ergodic properties of nonnegative matrices – I. *Pacific J. Math.*, 22:361-386, 1967.

[23] A. M. Yaglom. Certain limit theorems on the theory of branching stochastic processes. *Dokl. Akad. Nauk SSSR*, 56:795-798, 1947.

[24] G. Zohar. *Ph.D. Thesis.* Dept. of Electrical Eng., Technion, Haifa, 1997.