On the quasi-stationary distribution for some randomly perturbed transformations of an interval

Fima C. Klebaner*
Department of Mathematics and Statistics
The University of Melbourne
Parkville Australia 3052
F.Klebaner@ms.unimelb.edu.au

Justin Lazar†
Department of Statistics
The University of Melbourne
Parkville Australia 3052
LAZARJ@anz.com

Ofer Zeitouni‡
Department of Electrical Engineering
Technion
Haifa 32000, Israel
zeitouni@ee.technion.ac.il

January 13, 1997. Revised May 15, 1997

Abstract

We consider a Markov chain X_n^ϵ obtained by adding small noise to a discrete time dynamical system and study the chain's quasi-stationary distribution (qsd). The dynamics is given by iterating a function $f : I \to I$ for some interval I when f has finitely many fixed points, some stable and some unstable. We show that under some conditions the quasi-stationary distribution of the chain concentrates around the stable fixed points when $\epsilon \to 0$. As a corollary we obtain the result for the case when f has a single attracting cycle and perhaps repelling cycles and fixed points. In this case the quasi-stationary distribution concentrates on the attracting cycle. The result applies to the model of population dependent branching processes with periodic conditional mean function.

AMS subject classification: primary 60J80 Secondary 60F10

Key words: large deviations, logistic map, branching systems, quasi-stationary distribution.

*The work of this author was supported by the Australian Research Council
†The work of this author was supported by the Australian Postgraduate Research Award
‡The work of this author was partially supported by the Australian Research Council and by a US-ISRAEL BSF grant

1. INTRODUCTION

This paper deals with quasi-stationary distributions for Markov chains X_n^ϵ

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

obtained by adding small noise ξ^ϵ, which is generally state-dependent, to a discrete time dynamical system

$$x_{n+1} = f(x_n). \tag{2}$$

The function f maps some interval I into itself, and we consider the case when f has finitely many fixed points, some of which are stable and some unstable. The dynamical system defined by (2) models a particular physical phenomenon, when the variable of interest is confined to the range with set I. For example, in population dynamics x_n denotes the population density in the n-th generation. The Markov chain defined by (1) models small random perturbations to that phenomenon. Therefore in applications it is often of interest to study the long term behavior of the chain (1) as long as it stays in I. This is done by studying the quasi-stationary distribution (qsd) of the chain,

$$\rho^\epsilon(A) = \lim_{n\to\infty} P(X_n^\epsilon \in A | X_n^\epsilon \in I), \tag{3}$$

defined for any Borel set $A \subset I$.

Our goal in this paper is to show that under some conditions the quasi-stationary distribution of the chain concentrates around the stable fixed points of f when $\epsilon \to 0$.

Our motivation comes from the population density process in density-dependent branching models. A density-dependent branching process is a branching model, where the law of offspring distribution depends on the population density. These processes are stochastic analogues of the deterministic models of population dynamics given by (2), and were studied in [6], [7], where the Law of Large Numbers, the Central Limit Theorem and convergence to a Gaussian process were established. The basic feature of the deterministic models is the decline in the reproduction rate from a constant, when the population density is small, to zero when the population density increases to its threshold. This decline is due to competition between individuals. The logistic model with $f(x) = rx(1-x)$, one of the more famous models, eg. [10], provides a perfect example of this phenomenon, and we shall use this model in our representative examples. Density-dependent branching processes incorporate the same principle, by making the offspring distribution decline with an increase in population density. If $m(x)$ denotes the mean of offspring distribution when the population density is x, then the expected density of the next generation is given by the conditional mean function $f(x) = xm(x)$. The precise definition of the density-dependent branching processes is given in the section on applications, it suffice to say here that these models can be represented in the form of (1) with the dynamics being that of the corresponding deterministic population model with

$f(x) = xm(x)$. Under broad assumptions density-dependent branching processes become extinct with probability one. We show that the qsd exists and describes the long-term behavior of the process prior to extinction.

In [8], the Large Deviation Principle was established for the Markov chain (1), and this reference is central for this paper. It was shown there that if the deterministic system has a stable periodic orbit then the perturbed system will follow approximately a limit cycle for a time exponentially long in the level of noise before switching to another cycle.

We also consider the model with additive state-independent noise, it may also have an independent interest in applications. This model is conceptually the same, but technically it is easier to treat.

It is not hard to establish the existence of qsd under some broad assumptions, cf. Theorem 1, the main tool being the Krein-Rutman theory of positive operators. It is also rather straightforward that any weak limit of the qsd as the noise converges to zero must be an invariant measure of the deterministic system, defined by $\rho(A) = \rho(f^{-1}(A))$, cf. Theorem 2. Therefore if the deterministic system has a unique invariant measure then the qsd concentrates around it. We, however, consider the case when the deterministic model may have many invariant measures. The main result of this paper, Theorem 3, is to establish conditions so that the qsd does not put, in the limit of small noise, mass around the unstable fixed points. The technique of proof relies on an adaptation of the discrete time Freidlin-Wentzell theory, e.g. as developped in [5].

We use the main result to treat the case when f has an attracting stable cycle (an attracting fixed point is a cycle of period 1) to which all the trajectories converge (except unstable cycles and their inverse images), and show that under suitable assumptions the qsd concentrates around the attracting cycle. Thus the limit of the qsd is uniform on the stable periodic orbits.

To put our results into perspective, we point out a general result of Kifer [4], which shows, amongst other things, that for diffeomorphisms under a bit different conditions of hyperbolic sets and a more special form of random perturbations, i.e. perturbation by a diffusion, the quasi-invariant measures concentrate in the limit on the attractors. While our conclusions might possibly be obtained from general arguments (see [4, Section 5]), the case we consider does require some care, as the example provided after the statement of the main Theorem 3 demonstrates.

While working on this paper we learnt of a recent, independent work of Högnäs [3], where he obtained similar results for the stochastic Ricker model, by techniques not far from ours. The main difference of his work with ours lies in the greater generality and applicability of our assumptions. On the other hand, [3] explains how to extend the assumption (A6) below to allow for an application to the situation of stable periodic orbits of period larger than 2. This extension caries over to the results in this paper, see the comment at the end of Section 2.

2. ASSUMPTIONS AND RESULTS

Let $I = [0,1]$, and consider a continuous, piecewise C^2 function $f : I \to I$. $(I, \mathcal{B}, f)$ defines a standard dynamical system. For any $\delta > 0$, let $B_\delta = [\delta, 1-\delta]$. Throughout this paper, we make the following assumption:

(A1) $f(0) = f(1) = 0$, $\max_{x \in I} f(x) < 1$, and f possesses a finite number of unstable fixed points $x_0^* = 0$, $x_1^*, \ldots, x_k^*$ where k may equal zero, and a finite number of stable fixed points $s_1, ..., s_l \in (0,1)$. Further, for all $\delta_0 > 0$ small enough, there exist $\delta_0' > \delta_0 > 0$ such that $f(B_{\delta_0}) \subset B_{\delta_0'}$. (By x_i^* unstable, we mean that $|f'(x_i^*)| > 1$).

(A2) All trajectories of the dynamical system converge to one of the fixed points of f.

Consider the following Markov chain, which is a random perturbation of the dynamical system. Let

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

Here, $\xi^\epsilon(x)$ are random variables, whose law, denoted $Q_{\epsilon,x}^\xi(\cdot)$, depends on ϵ, x only. Let $\Lambda^\xi(x,\lambda) = \log(Ee^{\lambda \xi^\epsilon(x)})$ and define the kernel P_ϵ on $(I, \mathcal{B}(I))$ by

$$P_\epsilon(x, A) = P_x(X_1^\epsilon \in A), \ \ x \in I, \ \ A \in \mathcal{B}(I).$$

We make the following assumption on $Q_{\epsilon,x}^\xi(\cdot)$:

(A3) For all $\epsilon > 0$, and for some nonzero probability measure V_ϵ on $(I, \mathcal{B}(I))$, there exists a non-negative function $m_\epsilon(x,y)$ on $I \times I$ such that for all $x \in I$ and for all $A \in \mathcal{B}(I)$,

$$P_\epsilon(x, A) = \int_A m_\epsilon(x,y) V_\epsilon(dy).$$

Further, there exist integers $n_0(\epsilon)$ and real numbers $a(\epsilon)$ and $b(\epsilon)$ such that for all $x, y \in I$,

$$0 < a(\epsilon) \leq m_\epsilon^{(n_0(\epsilon))}(x,y) \leq b(\epsilon) < \infty,$$

where for any n, $m_\epsilon^{(n)}(x,y)$ is the density of $P_\epsilon^n(x,\cdot)$ with respect to V_ϵ. Assume further that for all $\delta > 0$ small enough, for all $\epsilon > 0$ small enough, $V_\epsilon(B_\delta) > 0$.

(A4) There exists a $\lambda_0 > 0$ such that, for all $|\lambda| \leq \lambda_0$, $\sup_{x \in I, \epsilon} \Lambda^\xi(x, \lambda/\epsilon) < \infty$.

Note that by Chebycheff's inequality, (A4) implies that $\xi^\epsilon(x)$ converge in probability to zero as $\epsilon \to 0$ uniformly in $x \in I$.

Let $U_\delta(x) = (x-\delta, x+\delta) \cap [0,1]$.

(A5) There exists a $\beta > 0$ independent of ϵ such that for all δ small enough, and all ϵ small,

$$\inf_{x \in \cup_{i=0}^{k} U_\delta(x_i^*)} Q_{\epsilon,x}^{\xi}(\epsilon, \infty) > \beta \,,$$

$$\inf_{x \in \cup_{i=0}^{k} U_\delta(x_i^*)} Q_{\epsilon,x}^{\xi}(-\infty, -\epsilon) > \beta \,.$$

(A6) For each $i = 1, ..., k$ there exists an $m_i > 2|f'(x_i^*)|$ such that for all $\delta_i > 0$ small enough, with $A_i = U_{m_i\delta_i}(x_i^*) \backslash U_{2\delta_i}(x_i^*)$, there exists a $j(i) < \infty$ such that $f^{(j)}(A_i) \subset \cup_{i=1}^{\ell} U_{\frac{\delta_i}{2}}(s_i)$.

For $i = 0$ we set $m_0 = 3|f'(0)|$.

For each $i = 0, \ldots, k$, define

$$\tau_{i,1}^\epsilon = \min\{n : X_n^\epsilon \notin U_{2\delta_i}(x_i^*)\}$$

and

$$\tau_{i,2}^\epsilon = \min\{n : X_n^\epsilon \in U_{\delta_i}(x_i^*)\}.$$

In what follows $\Rightarrow$ denotes the weak convergence.

Theorem 1 *Assume (A1)–(A4).*
a). The qsd $\rho^{(\epsilon)}$ (cf. (3)) exists, it is a probability measure on $(I, \mathcal{B}(I))$ and there exists a number $R^{(\epsilon)} > 1$ such that for all $A \in \mathcal{B}(I)$,

$$\rho^{(\epsilon)}(A) = R^{(\epsilon)} \int_I P_\epsilon(x, A)\rho^{(\epsilon)}(dx).$$

b). $\lim_{\epsilon \to 0} R^{(\epsilon)} = 1$.

Let $m_\epsilon^{(1)}(x,y) = m_\epsilon(x,y)$ and $m_\epsilon^{(n)}(x,y)$ be the density of P_ϵ^n for $n \geq 2$, i.e. $P_\epsilon^n(x,dy) = m_\epsilon^{(n)}(x,y)dy$.

Proof of Theorem 1: a). By (A3) and Theorem 10.1 of [2], P_ϵ has a positive eigenvalue $\frac{1}{R^{(\epsilon)}}$, larger in magnitude than any other eigenvalue, and corresponding uniquely defined left eigenfunction $\tilde{\rho}^{(\epsilon)}$

satisfying $\int_I \tilde{\rho}^{(\epsilon)}(y)dy = 1$. Moreover, still by [2] Theorem 10.1, with $\mu^{(\epsilon)}(\cdot)$ denoting the (uniquely defined) corresponding right eigenfunction,

$$m_\epsilon^{(n)}(x,y) = (\frac{1}{R^{(\epsilon)}})^n \mu^{(\epsilon)}(x)\tilde{\rho}^{(\epsilon)}(y)[1+o(1)]\,, \tag{4}$$

where the convergence of the error term in (4) is uniform in $x,\ y$. Then for $A \in \mathcal{B}(I)$ define $\rho^{(\epsilon)}(A) = \int_A \tilde{\rho}^{(\epsilon)}(y)dy$, so that it follows directly that $\rho^{(\epsilon)}$ is a qsd by using (4) in

$$P(X_n^\epsilon \in A | X_n^\epsilon \in I) = \frac{P(X_n^\epsilon \in A)}{P(X_n^\epsilon \in I)}.$$

b). With δ_0, δ_0' as in (A1), it holds that for any $x \in B_{\delta_0}$,

$$P_x(X_1^\epsilon \in B_{\delta_0}^c) \le Q_{\epsilon,x}^\xi(-\infty, \delta_0 - \delta_0') + Q_{\epsilon,x}^\xi(\delta_0' - \delta_0, \infty) \le e^{-\lambda_0|\delta_0 - \delta_0'|/\epsilon}(e^{\Lambda^\xi(x,\lambda_0/\epsilon)} + e^{\Lambda^\xi(x,-\lambda_0/\epsilon)})\,.$$

Hence, by using (A4), for some constant $c > 0$,

$$\sup_{x \in B_{\delta_0}} P_x(X_1^\epsilon \in B_{\delta_0}^c) \le e^{-c/\epsilon}\,. \tag{5}$$

Hence,

$$a_\epsilon \equiv \sup_{x \in B_{\delta_0}} (1 - P_x(X_1^\epsilon \in B_{\delta_0})) \to 0,\ \epsilon \to 0.$$

Then since $\rho^{(\epsilon)}(B_{\delta_0}) > 0$ for all ϵ,

$$\rho^{(\epsilon)}(B_{\delta_0}) = R^{(\epsilon)} \int_I P_x(X_1^\epsilon \in B_{\delta_0})\rho^{(\epsilon)}(dx) \ge R^{(\epsilon)} \int_{B_{\delta_0}} (1 - a_\epsilon)\rho^{(\epsilon)}(dx)$$

implies

$$R^{(\epsilon)} \le \frac{1}{1 - a_\epsilon} \to 1,\ \epsilon \to 0, \tag{6}$$

and since $R^{(\epsilon)} \ge 1$, b) follows. □

Since I is bounded, $\rho^{(\epsilon)}$ is tight. The following identifies the limit points of subsequences of $\rho^{(\epsilon)}$.

Theorem 2 *Assume (A1)–(A4). Then, any weak limit ρ of $\rho^{(\epsilon)}$ is an invariant measure for f, i.e., for any $A \in \mathcal{B}(I)$,*

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

Remark 1: Kifer [5, p. 8] proves this result in the case of stationary distributions of Markov chains under somewhat different assumptions, that are not satisfied here.

Proof of Theorem 2: Let Y_ϵ and Y be random variables with distributions $\rho^{(\epsilon)}$ and ρ respectively, with $Y_\epsilon \Rightarrow_{\epsilon \to 0} Y$. For any $A \in \mathcal{B}(I)$,

$$P(Y_\epsilon \in A) = R^{(\epsilon)} \int_I P_x(X_1^\epsilon \in A)\rho^{(\epsilon)}(dx) = R^{(\epsilon)} P(f(Y_\epsilon) + \xi^\epsilon(Y_\epsilon)) \in A). \tag{7}$$

By continuity of f and (A4), see the comment to (A4), $f(Y_\epsilon) + \xi^\epsilon(Y_\epsilon) \Rightarrow f(Y)$ along the subsequence. Hence for all $A \in \mathcal{B}(I)$ such that $P(Y \in \partial A) = P(f(Y) \in \partial A) = 0$, taking ϵ to zero along the subsequence in (7) along with part b) of Theorem 1 gives

$$\rho(A) = \rho(f^{-1}(A)). \tag{8}$$

Hence we can now show that (8) is true for any interval in I^0 by approximating the interval where necessary by sets that do satisfy it. Hence, by Caratheodory's Extension Theorem, it extends to all $A \in \mathcal{B}(I^0)$. Now, with $0 < \delta < 1 - \max_{x \in I} f(x)$ fixed,

$$\begin{aligned} \rho^{(\epsilon)}((1-\delta, 1]) &= R^{(\epsilon)} \int_I P_x(X_1^\epsilon \in (1-\delta, 1])\rho^{(\epsilon)}(dx) \\ &\leq R^{(\epsilon)} \max_{x \in I} P(\xi^\epsilon(x) > 1 - \delta - \max_{x \in I} f(x)) \to_{\epsilon \to 0} 0, \end{aligned}$$

and so $\rho(\{1\}) = 0$. Hence (8) is true for $A = \{0\}$ and $A = \{1\}$, and so it is true for all $A \in \mathcal{B}(I)$. $\square$

It is rather straight forward to check that any weak limit of $\rho^{(\epsilon)}$ then satisfies $\sum_{i=0}^k \rho(x_i^*) + \sum_{i=1}^l \rho(s_i^*) = 1$. The following is our main result.

Theorem 3 *Assume (A1)–(A6). Then, any weak limit ρ of $\rho^{(\epsilon)}$ satisfies $\rho(\{x_j^*\}) = 0$, for each $j = 0, ..., k$.*

Remark 2 To see that one needs some structural assumptions on $\xi^\epsilon(\cdot)$, consider the case of $f(x) = rx(1-x)$, $1 < r < 3$. For $x \in [\epsilon, 1]$ let $\xi^\epsilon(x)$ be a uniform random variable on $(-\epsilon/2, \epsilon/2)$, while for $x \in [0, \epsilon)$ let the law of $\xi^\epsilon(x)$ be $c_\epsilon U(-rx, -(r-1)x) + (1-c_\epsilon)U(1-\epsilon/2r, 1]$, where $U(A)$ denotes the uniform law on A and $c_\epsilon \to_{\epsilon \to 0} 1$ is chosen such that $E\xi^\epsilon(x) = 0$. It is not hard to check that in this case, the qsd concentrates on the (unstable) point 0.

Proof of Theorem 3: We first need the following preliminaries. Throughout, we use the notations introduced in (A1)–(A6). In addition, $N := N(\epsilon) = K(\log 1/\epsilon)^K$ for some K large enough (and independent of ϵ).

Lemma 1 *For each $i = 1, ..., k$,*

$$\sup_{x \in A_i} P_x(X_{j(i)}^\epsilon \in \left(\cup_{j=1}^\ell U_{\delta_i}(s_j)\right)^c) \to_{\epsilon \to 0} 0\,. \tag{9}$$

Proof: By (A4), $\sup_{x \in I} P(|\xi^\epsilon(x)| > c) \to 0$. This implies that for all $\delta > 0$, $\sup_{x \in I} P_x(|X_1 - f(x)| > \delta) \to_{\epsilon \to 0} 0$. Iterating this and using the fact that $j(i)$ is independent of ϵ, the lemma follows. $\square$

Lemma 2 *[8, Lemma 2.1] For each $i = 1, ..., l$, for each $j = 1, \ldots, l$, for all $k > 0$,*

$$\sup_{x \in U_{\delta_i}(s_i)} P_x(\tau^\epsilon_{j,2} < k/\epsilon) \to_{\epsilon \to 0} 0.$$

Lemma 3 *For each $i = 1, ..., k$, for all $K > 0$,*

$$\sup_{x \in A_i} \sup_{j(i) < t \le K(\log \frac{1}{\epsilon})^K} P_x(X^\epsilon_t \in \cup^k_{m=0} U_{\delta_m}(x^*_m)) \to_{\epsilon \to 0} 0\,.$$

Proof: Omit the i subscript for convenience. Take j as in (A6). Then for all $x \in A$, using the Markov property,

$$\begin{aligned}
\sup_{j<t\le N} P_x(X^\epsilon_t \in U_\delta(x^*)) &\le \sup_{t \le N} P_x(X^\epsilon_{j+t} \in U_\delta(x^*)) \\
&\le \sup_{y \in \cup^\ell_{m=1} U_\delta(s_m)} \sup_{t\le N} P_y(X^\epsilon_t \in \cup^k_{m=0} U_{\delta_m}(x^*_m)) + P_x(X^\epsilon_j \in \left(\cup^\ell_{m=1} U_\delta(s_m)\right)^c) \\
&\le \max_{m=1}^{\ell} \sup_{x \in U_\delta(s_m)} P_x(\tau^\epsilon_2 < N) + P_x(X^\epsilon_j \in \left(\cup^\ell_{m=1} U_\delta(s_m)\right)^c) \to_{\epsilon \to 0} 0,
\end{aligned} \tag{10}$$

where the last limit is due to Lemma 2 and Lemma 1. □

Lemma 4 *For each $i = 0, ..., k$,*
(i) $\sup_{x \in U_{\delta_i}(x^*_i)} P_x(X^\epsilon_{\tau_{i,1}} \notin U_{m_i \delta_i}(x^*_i)) \to_{\epsilon \to 0} 0$, and
(ii) for all K large enough, $\sup_{x \in U_{\delta_i}(x^*_i)} P_x(\tau^\epsilon_{i,1} > K(\log \frac{1}{\epsilon})^K) \to_{\epsilon \to 0} 0$.

Proof: Omit the i subscript for convenience.

(i) By the choice of m in (A6) (m_0 is defined immediately after (A6)), we have that for all δ small enough,

$$\sup_{x \in U_{2\delta}(x^*)} |f(x) - x^*| < m\delta \tag{11}$$

and so there exists a constant $\Delta > 0$ such that

$$\begin{aligned}
\sup_{x \in U_\delta(x^*)} P_x(X^\epsilon_{\tau_1} \notin U_{m\delta}(x^*)) &= \sup_{x \in U_\delta(x^*)} P_x(f(X^\epsilon_{\tau^\epsilon_1 - 1}) + \xi^\epsilon(X^\epsilon_{\tau^\epsilon_1 - 1}) \notin U_{m\delta}(x^*)) \\
&\le \sup_{x \in U_\delta(x^*)} P_x(|\xi^\epsilon(X^\epsilon_{\tau^\epsilon_1 - 1})| \ge \Delta) \\
&\le \sup_{x \in I} P(|\xi^\epsilon(x)| \ge \Delta) \to_{\epsilon \to 0} 0\,.
\end{aligned} \tag{12}$$

(ii) Let $b_\epsilon = \epsilon(\log\epsilon)^2$. Define $\tau = \inf\{t : |X^\epsilon_t - x^*| \ge b_\epsilon\}$. Without loss of generality, we assume that $\inf_{x \in U_\delta(x^*)} f'(x) \ge c > 1$. Then, with $c_1 = 4/\log c$,

$$\inf_{x^* \le x \le b_\epsilon + x^*} P_x(\tau < c_1 \log\log 1/\epsilon) \ge \inf_x P(\xi^\epsilon(x) > \epsilon) \inf_{\epsilon \le x \le b_\epsilon + x^*} P_x(\tau < c_1 \log\log 1/\epsilon - 1)\,.$$

Note that as long as $x_i \in U_\delta(x^*)$, on the event $\{\xi^\epsilon(x_i) \geq 0\}$, $f(x_i) + \xi^\epsilon(x_i) \geq c x_i$. Hence, since

$$\epsilon c^{c_1 \log\log 1/\epsilon} = \epsilon (\log \epsilon)^4 > b_\epsilon \,,$$

one gets, using (A5),

$$\inf_{x^* \leq x \leq b_\epsilon + x^*} P_x(\tau < c_1 \log\log 1/\epsilon) \geq \left(\inf_{x \in U_\delta(x^*)} P_x(\xi^\epsilon(x) > \epsilon) \right)^{-c_1 \log\log \epsilon} \geq \beta^{-c_1 \log\log \epsilon} \,.$$

Arguing similarly, the case $x < x^*$ is handled (for $i \neq 0$). Hence,

$$\inf_{|x^* - x| \leq b_\epsilon} P_x(\tau < c_1 \log\log 1/\epsilon) \geq \beta^{-c_1 \log\log \epsilon} \,.$$

It follows that, with $c_2 = 2c_1$,

$$\inf_{|x^* - x| \leq b_\epsilon} P_x(\tau \geq c_1 \beta^{c_2 \log\log \epsilon} \log\log 1/\epsilon) \leq (1 - \beta^{-c_1 \log\log \epsilon})^{\beta^{c_2 \log\log \epsilon}} \to_{\epsilon \to 0} 0 \,. \tag{13}$$

In what follows, we use c_3, c_4, c_5 to denote some constants which are independent of ϵ. Then, with $\xi_i^\epsilon = \xi^\epsilon(X_i^\epsilon)$, arguing as above,

$$\begin{aligned} \inf_{x \geq b_\epsilon + x^*, x \in U_\delta(x^*)} P_x(\tau_{i,1}^\epsilon \leq \frac{\log \frac{1}{b_\epsilon}}{\log \sqrt{c}}) & \\ &\geq P(|\xi_i^\epsilon| < b_\epsilon/\sqrt{c}\,,\ i = 1, \ldots, \frac{\log \frac{1}{b_\epsilon}}{\log \sqrt{c}}) \\ &\geq \left(1 - \frac{\epsilon c_3}{b_\epsilon}\right)^{-\log b_\epsilon / \log \sqrt{c}} \\ &\geq e^{\frac{c_4 \epsilon}{b_\epsilon} \log b_\epsilon} \geq e^{c_5 / \log \epsilon} \to_{\epsilon \to 0} 1 \,. \end{aligned} \tag{14}$$

Combining (13) and (14), we conclude that for some c_5 independent of ϵ,

$$\inf_{x \in U_\delta(x^*)} P_x(\tau_{i,1}^\epsilon > (\log 1/\epsilon)^{-c_5 \log \beta}) \to_{\epsilon \to 0} 0 \,.$$

This concludes the proof of the lemma. □

Lemma 5 *For any $i = 1, \ldots, k$*

$$\sup_{x \in U_{\delta_i}(x_i^*)} P_x(X_N^\epsilon \in \cup_{j=0}^k U_{\delta_j}(x_j^*)) \to_{\epsilon \to 0} 0 \,.$$

Proof: For $x \in U_{\delta_i}(x_i^*)$, using the Markov property,

$$\begin{aligned} P_x(X_N^\epsilon \in \cup_{j=0}^k U_{\delta_j}(x_j^*)) &\leq P_x(\tau_{i,1}^\epsilon > N) + P_x(X_{\tau_{i,1}}^\epsilon \notin U_{m_i \delta_i}(x_i^*)) + \\ &\quad \sup_{x \in A_i} \sup_{1 < t < N} P_x(X_t^\epsilon \in \cup_{j=0}^k U_{\delta_j}(x_j^*)) \,. \end{aligned} \tag{15}$$

Taking supremum on both sides over $x \in U_{\delta_i}(x_i^*)$, the first two terms on the right converge to zero as ϵ converges to zero by Lemma 4, while the convergence of the third is a consequence of Lemma 3. □

Lemma 6 *For any δ_0 small enough,*

$$\sup_{x\in[0,\delta_0]} P_x(X_N^\epsilon \in [0,\delta_0)) \to_{\epsilon\to 0} 0\,.$$

Proof of Lemma 6: Note that

$$\begin{aligned}\sup_{x\in[0,\delta_0]} P_x(X_N^\epsilon \in [0,\delta_0)) \quad &\le \quad \sup_{x\in[0,\delta_0]} P_x(\tau_{0,1}^\epsilon > N) + \sup_{x\in[0,\delta_0]} P_x(X_{\tau_{0,1}^\epsilon} \notin [2\delta_0, 3|f'(0)|\delta_0]) \\ &\quad + N \sup_{x\in[\delta_0,1-\delta_0]} P_x(X_1^\epsilon \in [0,\delta_0])\,. \end{aligned} \tag{16}$$

The first two terms in (16) converge to 0 with ϵ due to Lemma 4, while the third one is bounded (due to (5)) by $Ne^{-c/\epsilon} \to_{\epsilon\to 0} 0$. □

Proof of Theorem 3: Fix any $j = 0, ..., k$, and pick δ_j small enough so that

$$\rho^{(\epsilon)}(U_{\delta_j}(x_j^*)) \to_{\epsilon\to 0} \rho(\{x_j^*\}).$$

First observe that $R^{(\epsilon)} \le 1 + e^{-\frac{c}{\epsilon}}$, giving

$$(R^{(\epsilon)})^N \to_{\epsilon\to 0} 1,$$

for K large enough. Now,

$$\begin{aligned}\rho^{(\epsilon)}(U_{\delta_j}(x_j^*)) \quad &= \quad (R^{(\epsilon)})^N \int_0^1 P_x(X_N^\epsilon \in U_{\delta_j}(x_j^*))\rho^{(\epsilon)}(dx) \\ &= \quad (R^{(\epsilon)})^N \sum_{i=1}^k \int_{U_{\delta_i}(x_i^*)} P_x(X_N^\epsilon \in U_{\delta_j}(x_j^*))\rho^{(\epsilon)}(dx) + \\ &\quad\quad (R^{(\epsilon)})^N \sum_{i=1}^l \int_{U_{\delta_i}(s_i)} P_x(X_N^\epsilon \in U_{\delta_j}(x_j^*))\rho^{(\epsilon)}(dx) + \\ &\quad\quad (R^{(\epsilon)})^N \int_B P_x(X_N^\epsilon \in U_{\delta_j}(x_j^*))\rho^{(\epsilon)}(dx) + \\ &\quad\quad (R^{(\epsilon)})^N \int_{U_{\delta_0}(x_0^*)} P_x(X_N^\epsilon \in U_{\delta_j}(x_j^*))\rho^{(\epsilon)}(dx) \end{aligned} \tag{17}$$

where $B = [0,1] \setminus \left\{\cup_{i=0}^k U_{\delta_i}(x_i^*) \cup_{i=1}^l U_{\delta_i}(s_i)\right\}$, so that the third term converges to zero as ϵ converges to zero, the first term converges to zero as ϵ converges to zero using lemma 5, and for the second term, for each $i = 1, \ldots, l$,

$$\int_{U_{\delta_i}(s_i)} P_x(X_N^\epsilon \in U_{\delta_j}(x_j^*))\rho^{(\epsilon)}(dx) \leq \sup_{x\in U_{\delta_i}(s_i)} P_x(X_N^\epsilon \in U_{\delta_j}(x_j^*)) \leq \sup_{x\in U_{\delta_i}(s_i)} P_x(\tau_{j,2}^\epsilon \leq N) \to_{\epsilon\to 0} 0, \quad (18)$$

by Lemma 2. For $j = 0$, the last term converges to 0 with ϵ as a consequence of Lemma 6, implying that $\rho^{(\epsilon)}(U_{\delta_0}(x_0^*)) \to_{\epsilon\to 0} 0$. Substituting again in (17) for any $j \neq 0$, this is enough to imply that $\rho^{(\epsilon)}(U_{\delta_j}(x_j^*)) \to_{\epsilon\to 0} 0$. $\square$

Remark 3 We comment here on extensions of Theorem 3 which follow from [3]. As pointed out in the section below, assumption (A6) is hard to check in general. It is used in the proof, however, only to show that points in the neighborhood of unstable fixed points converge in finite time (under the deterministic action) to neighborhood of stable fixed points. This assumption can be replaced by the assumption of non-existence of heteroclinic orbits which include the unstable fixed point. This assumption creates an ordering of the unstable fixed point in such a way that neighborhoods of higher order unstable points are mapped to either stable points or neighborhoods of lower order fixed points, and the lowest order unstable fixed points do satisfy (A6). See [3] for a development of this approach in a particular case.

3. APPLICATIONS AND EXAMPLES

1. The logistic map with normal noise

Consider the case when f is the logistic map, $f(x) = rx(1-x)$, $1 < r \leq 1+\sqrt{6}$.

$$X_n^\epsilon = f(X_{n-1}^\epsilon) + \epsilon\eta_n, \quad (19)$$

where f as above and η_n are iid standard Normal rv's.

(a) Consider first the case when $1 < r < 3$. In this case $x_0^* = 0$ is the only unstable fixed point, $s_1 = 1 - \frac{1}{r}$ is the only stable fixed point and there are no other cycles. Assumptions (A1) and (A2) are satisfied for this case, see [11], and as $k = 0$ (A6) is not required for this case. If $\xi^\epsilon(x) = \epsilon\eta$, where η is the standard Normal variable, then the assumptions (A3), (A4) and (A5) are clearly fulfilled, and the result holds, ie. the qsd. of X_n^ϵ concentrates around the attracting fixed point s_1.

Remark 4 In this case it is possible to show that $\rho(0) = 0$ by using a Lyapunov function approach, see [9].

(b) Consider the case $3 < r < 1+\sqrt{6}$.

In this case f has two unstable fixed points at zero and $1 - \frac{1}{r}$, one stable attracting cycle of period two, and no other cycles. We next show that Theorem 3 applies in this setup, implying that the qsd of the X_n^ϵ in (19) concentrates on the stable attracting cycle.

Define the function $f^*(x) = f^2(x)$, $x \in I$. We shall apply the above results by replacing f with f^*. It is easy to see that f^* has two unstable fixed points at $x_0^* = 0$ and $x_1^* = 1 - \frac{1}{r}$, two stable fixed points s_1 and s_2, and no other cycles.

First we verify assumption (A6) for f^*. We start with

Lemma 7 *Let f be the logistic map with $3 < r < 1+\sqrt{6}$. Then the accumulation points of $f^{-n}(x_1^*)$ are zero and one.*

Proof: f maps $[0,1]$ onto $[0,r/4]$, and $r/4 < 1$! f from $[0,1/2] \to [0,r/4]$ is one-to-one. Put $g(x)$ to be the inverse map from $[0,r/4]$ onto $[0,1/2]$,

$$g(x) = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{x}{r}},$$

and note that it is monotone increasing. Then, $f^{-1}(x_1^*) = \{x_1^*, g(x_1^*)\} = \{x_1^*, 1 - x_1^*\}$. We prove that

$$1 - g^{(2)}(x_1^*) > r/4, \tag{20}$$

hence the "other inverse branch" of f does not contribute, and all further inverse images of x_1^* come from g. The inequality (20) follows from

$$g(x_1^*) < f(r/4). \tag{21}$$

The monotonicity of $x(1-x)$ implies that $g(x_1^*(r))$ is monotone decreasing in r for any $r > 2$. Therefore $\max_{3 \le r \le 1+\sqrt{6}} g(x_1^*(r)) = 1/3$ is achieved at $r = 3$. $f(r/4)$ is monotone decreasing in r on $(3/4, (1+\sqrt{6})/4)$. Therefore $\min_{3 \le r \le 1+\sqrt{6}} f(r/4)$ is achieved at $(1+\sqrt{6})/4)$ and is $f((1+\sqrt{6})/4) = 0.409... > 0.4$. The inequality (21) now follows, hence (20) is established. It follows by symmetry of f that the accumulation points of $f^{-n}(x_1^*)$ are the accumulation points of $\{g^n(x_1^*)\}$ and of $\{1 - g^n(x_1^*)\}$. Since g is monotone increasing, $\{g^n(x_1^*)\}$ form a decreasing sequence. The limit y must satisfy $y = g(y)$, which implies that $y = 0$, and the lemma is proved. $\square$

It is known that the iterates of all the points in I except the unstable fixed points and their inverse images for all $n = 0, 1, 2, \ldots$, converge to the stable cycle, see [11] p.73.

Since $f^{-n}(x_0^*) = \{0, 1\}$ for all n, it now follows from Lemma 7 that for all $\delta > 0$ small enough, for all $x \in U_\delta(x_1^*) \backslash \{x_1^*\}$,

$$f^n(x) \to \{s_1, s_2\}, \ n \to \infty.$$

With A_1 as in the statement of assumption (A6), we have that for each $x \in \bar{A}_1$ there exists an $n(x)$ such that $(f^*)^j(x) \in \cup_{i=1}^2 U_{\frac{\delta}{2}}(s_i)$ for all $j > n(x)$. Using the continuity of f, there exists thus a $r(x) > 0$ such that

$$\forall y \in B_{r(x)}(x),\, \forall j > n(x)+1,\; (f^*)^j(y) \in \cup_{i=1}^2 U_{\frac{\delta}{2}}(s_i)\,.$$

Covering $\bar{A}_1$ by the balls $B_{r(x)}(x)$ and taking a finite subset by compactness, (A6) follows for f^*.

Consider $Y_n^\epsilon = X_{2n}^\epsilon$. It is easily verified that assumption (A3) holds for (19), and that by Theorem 1, the qsd. of X_n^ϵ exists. It is clear that X_{2n}^ϵ and X_n^ϵ have the same qsd. Let

$$\xi^\epsilon(x) = f(f(x)+\epsilon\eta_1) + \epsilon\eta_2 - f^2(x).$$

With this choice of $\xi^\epsilon(x)$ we can write that

$$X_{2(n+1)}^\epsilon = f^2(X_{2n}^\epsilon) + \xi^\epsilon(X_{2n}^\epsilon).$$

To verify assumptions (A4) and (A5) notice that the iterated noise satisfies, by the mean value theorem,

$$\begin{aligned} \xi^\epsilon(x) &= f(f(x)+\epsilon\eta_1) + \epsilon\eta_2 - f^{(2)}(x) \\ &= \epsilon\eta_2 + f'(\theta)\epsilon\eta_1 \\ &= \epsilon\left(\eta_2 + f'(\theta)\eta_1\right), \end{aligned}$$

where $\theta \in (f(x), f(x)+\epsilon\eta_1)$ can be random. Since the derivative f' is bounded on I, $-r \le f'(x) \le r$, $|f'(\theta)| \le r$. Using this (A4) is straightforward. To verify (A5) write with $\beta = P(\eta_2 > 1 + r|\eta_1|)$

$$\begin{aligned} \inf_{U_\delta(x^*)} P(\xi^\epsilon(x) > \epsilon) &= \inf_{U_\delta(x^*)} P(\eta_2 + f'(\theta)\eta_1 > 1) \\ &> \inf_{U_\delta(x^*)} P(\eta_2 > 1 + r|\eta_1|) = \beta > 0. \end{aligned}$$

The other side of (A5) is similar. This completes checking (A1)–(A6).

Theorem 3 implies than the limit of the qsd is of the form $\alpha\delta_{s_1} + (1-\alpha)\delta_{s_2}$, where $0 \le \alpha \le 1$ and δ_{s_i} denotes the point mass at s_i, $i = 1, 2$. Since s_1, s_2 are points of the cycle for f, it follows that $\alpha = 1/2$, ie. the limit is the uniform distribution on the stable cycle. □

Remark 5 The following is useful for checking assumptions (A1), (A2). The dynamical system defined by iterations of $f \in C^0$ is called simple if for any starting point the trajectory converges to one of the cycles of f. f is said to belong to class G_{2^m} if it has a cycle of period 2^m and no cycles of period greater than 2^m. The result [11] p.73 states that if $f \in G_{2^m}$ for some m, then the dynamical system defined by iterations of f starting from any point in I is a simple dynamical system. On the other hand if f is C^1 and if the dynamical system defined by iterations of f starting from any point in I is a simple dynamical system, then $f \in G_{2^m}$, Theorem 3.3 in [11].

2. Density dependent branching processes

In this section we apply the main result to the model of density dependent branching processes. Let $\{X_n^\epsilon\}$ be the population density in a density dependent branching process, defined as follows. For any fixed $x \geq 0$, let $Y_{j,n}(x)$ be independent and identically distributed random variables for all $j,\ n$ with distribution $Y(x)$, where for all $x \geq 0,\ Y(x)$ is non-negative and integer-valued, and for all $x \geq 1,\ Y(x) = 0$. $Y(x)$ represents the law of offspring distribution when the population density is x. Then for fixed $K \in [2, \infty)$ define a population density dependent branching process $\{Z_n^K\}$, $n = 0, 1, 2, \ldots$ with threshold K inductively by taking Z_0^K to be a positive integer less than K and

$$Z_{n+1}^K = \begin{cases} \sum_{j=1}^{Z_n^K} Y_{j,n+1}(\frac{Z_n^K}{K}) & , Z_n^K > 0 \\ 0 & , Z_n^K = 0 \end{cases}$$

where we assume that for any fixed $x,\ K$ and $n+1$, that the $Y_{j,n+1}(x),\ j = 1, 2, \ldots$, are independent of $Z_n^K, Z_{n-1}^K, \ldots, Z_0^K$. Let $X_n^K = \frac{Z_n^K}{K}$ denote the population density.

For this application it is suitable to take $I = (0, 1)$ to be the open interval. Then if $X_n^K \in I$, then, with $f(x) = xEY(x) = xm(x)$,

$$X_{n+1}^K = f(X_n^K) + \frac{1}{K} \sum_{j=1}^{KX_n} (Y_{j,n+1}(X_n) - m(X_n^K)) = f(X_n^K) + \frac{1}{K} \sum_{j=1}^{KX_n} Y'_{j,n+1}(X_n), \qquad (22)$$

where $Y'(x) = Y(x) - m(x)$ denotes the centered offspring distribution, and $X_{n+1}^K = 0$ if $X_n^K \notin I$.

To apply the main result to the chain $\{X_n^K\}$, set $\epsilon = \frac{1}{\sqrt{K}}$, and put

$$\xi^\epsilon(x) = \frac{1}{\sqrt{K}} \sum_{j=1}^{Kx} Y'_j(x) = \epsilon^2 \sum_{j=1}^{x/\epsilon^2} Y'_j(x). \qquad (23)$$

Taking $f(x)$ to be the logistic map $rx(1-x)$ with $1 < r < 1 + \sqrt{6}$, the assumptions (A1), (A2) and (A6) on $f(x)$ were verified in the previous section on the logistic map with additive noise.

To verify the rest of the main assumptions, we make the following assumptions on the offspring distribution $Y(x)$.

Assume (A3) with V_ϵ being a discrete measure on i/K, $i = 1, \ldots, K-1$. Further, let $\Lambda^{Y'}(x, \lambda) = \log E \exp(\lambda Y'(x))$, and assume that

(B1) $\sup_{x \in I, \epsilon} (x/\epsilon^2) \Lambda^{Y'}(x, \lambda\epsilon) < \infty$.

To show (A4) write by using independence

$$\begin{aligned}
\Lambda^{\xi}(x,\lambda/\epsilon) &= \log E(\exp(\lambda\epsilon\sum_{j=1}^{x/\epsilon^2} Y_j'(x)) = \log E(\prod_{j=1}^{x/\epsilon^2}\exp(\lambda\epsilon Y_j'(x)) \\
&= \log\prod_{j=1}^{x/\epsilon^2} E\exp(\lambda\epsilon Y_j'(x)) = x/\epsilon^2\log E(\exp(\lambda\epsilon Y'(x)) \\
&= x/\epsilon^2\Lambda^{Y'}(x,\lambda\epsilon).
\end{aligned}$$

Thus (B1) implies (A4).

To give a specific example, take $Y(x)$ to be a Poisson$(m(x))$ random variable, where $f(x) = xm(x)$. For the logistic dynamics, $m(x) = r(1-x)$ for $x \in I$, and zero for $x \notin I$. Then $\Lambda^{Y'}(x,\lambda) = m(x)(e^{\lambda}-1) - \lambda m(x)$. So that

$$x/\epsilon^2\Lambda^{Y'}(\lambda\epsilon) = f(x)(e^{\lambda\epsilon}-1-\lambda\epsilon)/\epsilon^2.$$

Using expansion for the exponential, (B1) clearly holds. (A3) also holds.

Before we verify (A5) we prove a result on sums of iid rv.'s.

$$\frac{1}{\sqrt{K}}\sum_{j=1}^{Kx} Y_j'(x) = \epsilon\sum_{j=1}^{x/\epsilon^2}(Y_j(x) - m(x)) = \xi^{\epsilon}(x)/\epsilon.$$

Let $\nu(x) = E(|Y'(x)|^3)$ and $\sigma^2(x) = \mathrm{Var}(Y(x)) = E(Y'^2(x))$.

Lemma 8 *Let $A \subset I$ be such that*

$$\sup_{x\in A}\frac{\nu(x)}{\sigma^3(x)\sqrt{x}} < \infty. \tag{24}$$

Then sums $\frac{1}{\sqrt{K}}\sum_{j=1}^{Kx} Y_j'(x)$ converge in distribution as $K \to \infty$ to the $N(0, x\sigma^2(x))$ distribution uniformly in $x \in A$.

Proof: Indeed, fix x, and using the Berry-Esseen inequality ([1], p.542), we have for any real a, with Φ being the standard Normal cdf,

$$|P(\frac{1}{\sqrt{K}}\sum_{j=1}^{Kx} Y_j'(x) \le a\sigma(x)\sqrt{x}) - \Phi(a)| \le \frac{3\nu(x)}{\sigma^3(x)\sqrt{Kx}}. \tag{25}$$

Replacing a by $a/(\sigma(x)\sqrt{x})$ we obtain convergence to the $N(0, x\sigma^2(x))$ distribution, moreover the convergence is uniform in x if condition (24) holds. $\Box$

An easily checked sufficient condition for (24) is (B2).

(B2) Assume that there is a small $\gamma > 0$, such that with $A = (\gamma, 1-\gamma)$, $\sup_{x\in A}\nu(x) < \infty$, and $\inf_{x\in A}\sigma^2(x) > 0$.

It is clear that (B2) implies (24). Now, the uniform convergence of the sums to the normal distribution implies (A5). Indeed, as it was assumed that unstable fixed points x_i^*, $i = 1, 2, \ldots, k$ are in the interior of I, there is $\gamma > 0$ so that A includes all of these points. Now,

$$\inf_{x \in \cup_{i=1}^k U_\delta(x_i^*)} Q^\xi_{\epsilon,x}(\epsilon, \infty) \geq \inf_{x \in A} P(\xi^\epsilon(x) > \epsilon) = \inf_{x \in A} P(\frac{1}{\sqrt{K}} \sum_{j=1}^{Kx} (Y_j(x) - m(x)) > 1) = \beta > 0.$$

(B2) is clearly satisfied for the specific example when $Y(x)$ is a Poisson($m(x)$) rv.

This completes the check of the basic assumptions for branching processes when $f(x) = xm(x)$ has a single attracting fixed point, eg. when f is a logistic map and $1 < r < 3$. In this case the main result implies that the qsd. concentrates around this point. We formulate this as

Theorem 4 *Suppose the offspring distributions satisfy (B1) and (B2). Then* $\rho(x_0^*) = 0$.

When $3 < r < 1 + \sqrt{6}$ such an attracting fixed point does not exist, but there is an attracting two-cycle. The analysis via $f^{(2)}$, similar to that carried out in the additive case and details of which are omitted, shows that the qsd. concentrates on the attracting cycle. These results for the stochastic Ricker model were obtained independently by Högnäs (1996).

References

[1] Feller, W. (1971) *An Introduction to Probability Theory and Its Applications. Vol.2.* Wiley, New York.

[2] Harris, T.E. (1989) *The Theory of Branching Processes.* Dover, New York.

[3] Högnäs, G. (1996) On the quasi-stationary distribution of a stochastic Ricker model. Preprint.

[4] Kifer, Y. (1990) Principal eigenvalues, topological pressure, and stochastic stability of equilibrium states, *Israel J. Math*, 70, pp. 1–47.

[5] Kifer, Y. (1988) *Random Perturbations of Dynamical Systems.* Birkhäuser, Boston.

[6] Klebaner, F.C. (1993) Population-dependent branching processes with a threshold. *Stochastic Process. Appl.* 46, 115-127.

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

[8] Klebaner, F.C. and Zeitouni, O. (1994). The exit problem for a class of density-dependent branching systems. *Ann. Appl. Probab.* Vol. 4, No 4, 1188-1205.

[9] Lazar, J.(1996) *On the Quasi-Stationary Distribution of Randomly Perturbed Discrete Time Dynamical Systems.* Ph.D. Thesis University of Melbourne.

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

[11] Sharkovskii, A.N., Maistrenko, Y.L. and Romanenko E.Y. (1993) *Difference Equations and Their Applications.* Mathematics and Its Applications, Volume 250. Kluwer, Dordrecht.