Nonlinearity **12** (1999) 1029–1052. Printed in the UK PII: S0951-7715(99)96749-X

Ulam's method for random interval maps

Gary Froyland

Department of Mathematics and Computer Science, University of Paderborn, Warburger Str. 100, Paderborn 33098, Germany

E-mail: froyland@math.uni-paderborn.de and
URL: http://www-math.uni-paderborn.de/∼froyland

Received 13 August 1998, in final form 18 January 1999
Recommended by L Bunimovich

Abstract. We consider the approximation of absolutely continuous invariant measures (ACIMs) of systems defined by random compositions of piecewise monotonic transformations. Convergence of Ulam's finite approximation scheme in the case of a single transformation was dealt with by Li (1976 *J. Approx. Theory* **17** 177–86). We extend Ulam's construction to the situation where a family of piecewise monotonic transformations are composed according to either an iid or Markov law, and prove an analogous convergence result. In addition, we obtain a convergence *rate* for our approximations to the unique ACIM, and provide rigorous *bounds* for the L^1 error of the Ulam approximation.

AMS classification scheme numbers: 58F11, 65U05, 60G10, 28D99

1. Introduction

We begin by giving a rough description of the setting and our results. Let $\{T_k\}_{k=1}^r$ be a collection of mappings from the unit interval I into itself. Given an initial point $x \in I$, and a (random) sequence $(k_0, k_1, \ldots)$ with $k_N \in \{1, \ldots, r\}$ for $N \geqslant 0$, we produce a (random) orbit by defining the Nth point in the orbit to be $x_N = x_N(k_{N-1}, \ldots, k_0, x) := T_{k_{N-1}} \circ \cdots T_{k_1} \circ T_{k_0} x$. There are two cases we consider. First, the map $T_{k_{N-1}}$ that is applied at time N is selected from the collection $\{T_k\}_{k=1}^r$ independently of the maps that have previously been applied, and according to the same probability law for all time. Such a composition is often referred to as a random iid composition, and the indices $k_0, k_1, \ldots$ arise as random variables of an iid process. Second, the map $T_{k_{N-1}}$ that is applied at time N is chosen so as to depend only on the map applied at the previous time step, and according to the same probability law for all time. In this situation, the indices $k_0, k_1, \ldots$ arise as random variables of a stationary first-order Markov chain, and we call such a composition of maps a Markov random composition. Such random dynamical systems arise in a variety of ways.

We are interested in determining the asymptotic behaviour of such systems in situations where the orbits $\{x_N\}_{N=0}^\infty \subset I$ have the same asymptotic distribution on I for almost all sequences $k_0, k_1, \ldots$ and Lebesgue almost all starting points $x \in I$, as such asymptotic distributions are clearly of physical significance for the system at hand. To this end, we extend a method of Ulam [18] to produce a rigorous approximation method for absolutely continuous probability measures that are 'invariant on average' under the action of the random system.

2. Formalities, background and outline

We formally set up our problem. Let $S_i = \{1, \ldots, r\}$, $i \geqslant 0$, and $\Omega = \prod_{i=0}^{\infty} S_i$. We select a probability measure $\mathbb{P}$ on Ω that is invariant under the left shift $\sigma : \Omega \circlearrowleft$. The space Ω contains infinite sequences of indices for the maps $T_\cdot$, and the shift invariant probability measure $\mathbb{P}$ governs the stationary stochastic process that generates a (random) index at each time step. In the iid case, we select a probability vector $(w_1, \ldots, w_r)$, and define a probability measure ρ on S_i, $i \geqslant 0$, by $\rho(\{k\}) = w_k$. Denote $[a_0, \ldots, a_s]_t = \{\omega \in \Omega : \omega_t = a_0, \omega_{t+1} = a_1, \ldots, \omega_{t+s} = a_s\}$, and define $\mathbb{P}$ by $\mathbb{P}([a_0, \ldots, a_s]_t) = w_{a_0} \cdots w_{a_s}$, consistently extending $\mathbb{P}$ to all of Ω. In the Markov case, we select an irreducible, aperiodic stochastic $r \times r$ matrix W with invariant (normalised) left eigenvector $(w_1, \ldots, w_r)$, and define $\mathbb{P}([a_0, \ldots, a_s]_t) = w_{a_0} W_{a_0,a_1} \cdots W_{a_{s-1},a_s}$, again consistently extending $\mathbb{P}$ to all of Ω.

Define the skew product $\tau : \Omega \times I \circlearrowleft$ by $\tau(\omega, x) = (\sigma\omega, T_{\omega_0}x)$. We form a (random) dynamical system by considering the orbit $\{\mathrm{Proj}_I(\tau^N(\omega, x))\}_{N=0}^{\infty}$ on I where $\omega \in \Omega$, $x \in I$, and Proj_I denotes the canonical projection of $\Omega \times I$ onto I. By putting $x_N = \mathrm{Proj}_I(\tau^N(\omega, x))$, we see that $x_N = T_{\omega_{N-1}} \circ \cdots \circ T_{\omega_0} x$ for $N \geqslant 1$, with $x_0 = x$. Thus the orbit x_N is defined by a random composition of the mappings $T_1, \ldots, T_r$; the orbit is random in the sense that the sequence of maps $T_{\omega_{N-1}} \circ \cdots T_{\omega_0}$ has probability $\mathbb{P}([\omega_0, \ldots, \omega_{N-1}])$ of occurring. We wish to study the asymptotic behaviour of the orbit x_N.

Denote by $\mathcal{M}(\Omega \times I)$ the space of Borel probability measures on $\Omega \times I$.

Definition 2.1. *We shall say that a probability measure $\tilde{\mu} \in \mathcal{M}(\Omega \times I)$ is τ-*invariant *if*

(i) $\tilde{\mu} \circ \tau^{-1} = \tilde{\mu}$, *and*
(ii) $\tilde{\mu}(E \times I) = \mathbb{P}(E)$ *for all measurable* $E \subset \Omega$.

We say that $\mu \in \mathcal{M}(I)$ is invariant on average, *or simply* invariant, *if there exists a τ-invariant probability measure $\tilde{\mu}$ such that $\mu(A) = \tilde{\mu}(\Omega \times A)$ for all measurable $A \subset I$.*

We seek to approximate invariant measures μ that are absolutely continuous with respect to Lebesgue measure m on I, and restrict ourselves to the situation where the T_k are Lasota–Yorke maps.

Definition 2.2. *We shall say that an interval map $T : I \circlearrowleft$ is a* Lasota–Yorke map *if (i) there is a finite partition $0 = b_0 < b_1 < \cdots < b_q = 1$ of I such that $T|_{(b_{\ell-1}, b_\ell)}$, is a C^2 function and may be extended to a C^2 function on $[b_{\ell-1}, b_\ell]$ for $\ell = 1, \ldots, q$, and (ii) $\inf_{x \in I \setminus \{b_0, \ldots, b_q\}} |T'(x)| > 0$. We denote the partition for the map T_k by $0 = b_0^k < b_1^k < \cdots < b_{q_k}^k = 1$.*

Denote by $T_k(b_\ell^{k,-})$ and $T_k(b_\ell^{k,+})$, the values that T_k takes on either side of the break point b_ℓ^k, $\ell = 1, \ldots, q_k - 1$. We define the numbers $\theta_{k,\ell}$, $\ell = 1, \ldots, q_k - 1$, as follows:

$$\theta_{k,\ell} = \begin{cases} 0, & \text{if } T_k(b_\ell^{k,-}) = 0 \text{ or } 1, \text{ and } T_k(b_\ell^{k,+}) = 0 \text{ or } 1, \\ 2, & \text{if } T_k(b_\ell^{k,-}) \neq 0 \text{ or } 1, \text{ and } T_k(b_\ell^{k,+}) \neq 0 \text{ or } 1, \\ 1, & \text{otherwise.} \end{cases}$$

For $\ell = 0$ and $\ell = q_k$, we put $\theta_{k,\ell} = 0$ if $T_k(b_\ell^k) = 0$ or 1, and $\theta_{k,\ell} = 1$ otherwise.

There exists a minimal partition $0 = b_0^* < b_1^* < \cdots < b_{q^*}^* = 1$ such that for each $k = 1, \ldots, r$ and all $\ell = 1, \ldots, q^*$, one has $T_k|_{(b_{\ell-1}^*, b_\ell^*)}$, is a C^2 function and may be extended to a C^2 function on $[b_{\ell-1}^*, b_\ell^*]$. This number q^* will be used in the main theorems.

We sometimes refer to maps $T : I \circlearrowleft$ as C^2 circle maps. Here we mean to identify the half open unit interval I with a circle, and assume that the map T continues to be C^2 under this identification. Clearly in this case, T is C^2 everywhere and there is no partition allowed.

In the case of a single mapping with $\inf_{x\in I\setminus\{b_0,\dots,b_q\}} |T'(x)| > 2$, it was in such a setting that Li [9] first proved convergence (in L^1) of Ulam's approximation to the unique absolutely continuous invariant measure (ACIM), following the Lasota and Yorke [8] proof of the existence of an absolutely continuous invariant measure. The existence of an ACIM for iid random compositions of such mappings has been considered independently in this setting by Pelikan [16] and Morita [12]. In fact, Morita [11] also gives conditions for the existence of an ACIM where $\mathbb{P}$ is any shift invariant probability measure; that is, the sequence of indices $(\omega_0, \omega_1, \dots)$ arises as random variables of any stationary stochastic process.

However, to the author's knowledge, the rigorous numerical approximation of ACIM's for random compositions has not been considered. We restrict ourselves to using collections of Lasota–Yorke type interval maps, as both the proof of existence of ACIM's and the proof of convergence of Ulam's method are well known in the deterministic situation of a single mapping. Related extensions include [1, 4] which contain existence results for ACIM's of higher-dimensional expanding mappings and iid compositions of expanding Jablonski mappings respectively. Extensions of Ulam's method to higher-dimensional deterministic systems include [2, 3, 14].

We firstly consider iid random compositions and begin by proving suitable inequalities regarding the variation of test functions and their images under an appropriate Perron–Frobenius operator, following the construction of [16]. We then form a finite-dimensional approximation of this Perron–Frobenius operator and use the variational inequalities to first prove *convergence* of Ulam's method. Under slightly stricter conditions, we are able to produce a *rate* of convergence, and in the case where each T_k is a C^2 circle map, we obtain a *bound* for the error in our approximation, in terms of fundamental constants of the mappings T_k, $k = 1, \dots, r$. We can also produce a bound for the case of general Lasota–Yorke maps, but it uses properties of the (unknown) ACIM.

We then turn to Markov random compositions. Here, the definition of an appropriate Perron–Frobenius operator is more complicated. We begin by outlining the construction of a suitable operator and showing that fixed points of this operator correspond to invariant measures of the skew product τ, generalizing a result of Ohno [15] for iid compositions. We then form a finite-dimensional approximation of this new operator and obtain similar results to the iid case.

We remark that the constructions described in this paper should be able to be combined with the variational inequalitites of [2, 4] to prove convergence of Ulam's method for higher-dimensional random dynamical systems. Morever, the representations of the appropriate discretised Perron–Frobenius operators for iid and Markov compositions may be used as a new numerical method for (non-rigorously) approximating invariant measures of more general (non-expanding and/or higher dimensional) random dynamical systems.

Sections 3 and 4 contain the main results for the iid and Markov cases, respectively. Section 5 details numerical examples of our results; the proofs are in section 6.

3. Ulam's method for iid compositions

3.1. Invariant measures for iid compositions

We begin by giving a simpler description of invariant measures on I. The following lemmas hold in the very general situation where I is simply a metric space and each T_k is a (Borel) measurable map of I.

Lemma 3.1 (Ohno [13]). *Define an operator* $\mathcal{D} : \mathcal{M}(I) \circlearrowleft$ *by* $\mathcal{D}\nu = \sum_{k=1}^{r} w_k \nu \circ T_k^{-1}$. *A measure* $\mu \in \mathcal{M}(I)$ *is fixed under* $\mathcal{D}$ *iff* $\mathbb{P} \times \mu$ *is* τ*-invariant.*

Thus, it is easy to construct τ-invariant measures as products of the measure $\mathbb{P}$ and measures on I that are fixed by $\mathcal{D}$. The following result tells us that in the case of *absolutely continuous* τ-invariant measures (absolutely continuous with respect to $\mathbb{P} \times m$, where m is Lebesque measure on I), all are of this product form.

Lemma 3.2 (Morita [11]). *Assume that each T_k, $k = 1, \ldots, r$ is non-singular ($m(A) = 0 \Rightarrow m(T^{-1}A) = 0$ for all Lebesgue measurable $A \subset I$), with respect to m. Then any $\mathbb{P} \times m$ absolutely continuous τ-invariant measure has the form $\mathbb{P} \times \mu$ where μ is fixed by $\mathcal{D}$.*

Thus finding an absolutely continuous probability measure μ on I satisfying $\mathcal{D}\mu = \mu$ is the only way to construct an absolutely continuous τ-invariant measure on $\Omega \times I$.

3.2. Statement of results

As we are interested in absolutely continuous invariant measures, we may rewrite the operator $\mathcal{D}$ as an action on density functions. The corresponding action on $f \in L^1(I, m)$ is given by

$$\mathcal{P} f = \sum_{k=1}^{r} w_k \mathcal{P}_k f, \tag{1}$$

where $\mathcal{P}_k : L^1(I, m) \circlearrowleft$ denotes the standard Perron–Frobenius operator for the map T_k, namely, $\mathcal{P}_k f(x) = \sum_{y \in T_k^{-1}x} \frac{f(y)}{|T_k'(y)|}$. Theorem 1 [16] states that an invariant density of bounded variation on I exists provided that $\sum_{k=1}^{r} w_k/|T_k'(x)| \leqslant \zeta < 1$, while theorem 5.1 [12] shows that the same conclusion holds under the slightly different assumption that $\sum_{k=1}^{r} w_k \log\left(1/\inf_{x\in I} |T_k'(x)|\right) < 0$. We will further assume that there is only one invariant density for our iid random composition (that is, there is one fixed point of $\mathcal{P}$). Corollary 7 [16] and lemma 5.4 [12] both state that provided one of the T_k is uniformly expanding and that this T_k itself has a unique invariant density with support of all of I, then the random composition has a unique invariant density, provided one exists. We will therefore assume that:

(i) $\sup_{x\in I} \sum_{k=1}^{r} w_k |1/T_k'(x)| < 1$,
(ii) one of the T_k is uniformly expanding and has a unique invariant density with support all of I.

We will denote the density of the unique ACIM μ as h; so $\mathcal{P}h = h$. We now state our first main result.

Theorem 3.3. *Let $\{T_1, \ldots, T_r\}$ be a collection of Lasota–Yorke maps. Equipartition the unit interval into n subintervals $I_i = [(i-1)/n, i/n)$, and define r stochastic matrices $P_n(k)$, $k = 1, \ldots, r$, by*

$$P_{n,ij}(k) = \frac{m(I_i \cap T_k^{-1} I_j)}{m(I_i)}. \tag{2}$$

Further, define $P_n = \sum_{k=1}^{r} w_k P_n(k)$, and let p_n be a fixed (normalized) left eigenvector of P_n. Define the approximate invariant density

$$h_n = \sum_{i=1}^{n} \frac{p_{n,i}}{m(I_i)} \chi_{I_i}. \tag{3}$$

Then setting

$$\alpha = \sup_{x\in I} \sum_{k=1}^{r} w_k \frac{1}{|T_k'(x)|}, \tag{4}$$

$$\beta = \sum_{k=1}^{r} w_k \frac{\sup_{x\in I} |T_k''(x)|}{\inf_{x\in I} |T_k'(x)|^2} \tag{5}$$

and

$$\eta = \sum_{k=1}^{r} w_k \frac{\sum_{\ell=0}^{q_k} \theta_{k,\ell}}{\inf_{x\in I} |T_k'(x)|}. \tag{6}$$

(i) If $\alpha < \frac{1}{2}$, $\|h_n - h\|_1 \to 0$ as $n \to \infty$.

(ii) If $\alpha + \eta + \beta/2 < 1$, and the endpoints of the equipartition contain all points where there is a break in the C^1 behaviour of h, then there exists a constant $\mathcal{C} < \infty$ such that $\|h_n - h\|_1 \leqslant \mathcal{C} \log n/n$.

(iii) If $\alpha + \beta/2 < 1$ and each T_k is a C^2 circle map, then the constant $\mathcal{C}$ above may be written in terms of fundamental constants of the maps T_k. Set $C = \max_{1\leqslant k\leqslant r} \mathrm{Lip}(\log |T_k'|)$ and $\lambda = \min_{1\leqslant k\leqslant r} \inf_x |T_k'(x)|$ (assuming $\lambda > 1$). Then

$$\|h - h_n\|_1 \leqslant (\mathrm{e}^{C/(\lambda-1)n} - 1) \inf_{0<\delta<1} \left\{ \left(2 + \frac{\delta}{1-\delta}\right) \left(\left[\frac{\log(4n/\delta)}{-\log(\alpha + \beta/2)} \right] + 1 \right) - \frac{1}{2} \right\}, \tag{7}$$

where $[\cdot]$ denotes the integer part.

Remark 3.4. *(i) In case (iii) of theorem 3.3, the bound (7) may be broken up into bounds for $\|h - \pi_n(h)\|_1$ and $\|\pi_n(h) - h_n\|_1$, where $\pi_n(h) = \sum_{i=1}^n (\frac{1}{n} \int_{I_i} h \, \mathrm{d}m) \chi_{I_i}$. The function $\pi_n(h)$ is simply the unique step function (constant on each I_i) with the property that $\int_{I_i} \pi_n(h) \, \mathrm{d}m = \mu(I_i)$; because of this property, $\pi_n(h)$ is the 'best' step function approximation to h on the partition $\{I_i\}$. Unfortunately, Ulam's method usually does not give us this best approximation, and so for each partition set, there is an error in the weight that Ulam's method assigns to it, namely $|\mu(I_i) - p_{n,i}|$. It is easy to see that $\|\pi_n(h) - h_n\|_1 = \sum_{i=1}^n |\mu(I_i) - p_{n,i}|$, and in fact, the bound for this error forms the bulk of the rhs of (7), with*

$$\sum_{i=1}^{n} |\mu(I_i) - p_{n,i}| \leqslant (\mathrm{e}^{C/(\lambda-1)n} - 1) \inf_{0<\delta<1} \left\{ \left(2 + \frac{\delta}{1-\delta}\right) \left(\left[\frac{\log(4n/\delta)}{-\log(\alpha + \beta/2)} \right] + 1 \right) - 1 \right\}.$$

The other term $\|h - \pi_n(h)\|_1$ is simply controlled through estimates on the derivative of h, and this error is bounded by $(\mathrm{e}^{C/(\lambda-1)n} - 1)/2$.

In practice, we find that optimal values of δ often lie between 0.1 and 0.2.

(ii) In case (i) of theorem 3.3, if $\frac{1}{2} < \alpha < 1$ then we may apply Ulam's method by using higher iterates of the random system. There is an integer N such that

$$\alpha_N := \sum_{k_0,k_1,\ldots,k_{N-1}=1}^{r} w_{k_0} w_{k_1} \ldots w_{k_{N-1}} \frac{1}{\inf_{x\in I} |(T_{k_{N-1}} \circ \cdots \circ T_{k_1} \circ T_{k_0})'(x)|} < \frac{1}{2}.$$

In fact, $\alpha_N \leqslant \alpha^N$, so we simply require $N > (\log 2)/(\log 1/\alpha)$. To apply theorem 3.3 (i) in such a situation, it is necessary to construct r^N matrices of the form

$$P_{k_0,\ldots,k_{N-1},n,ij} = \frac{m(I_i \cap (T_{k_{N-1}} \circ \cdots \circ T_{k_0})^{-1} I_j)}{m(I_i)},$$

and define

$$P_n = \sum_{k_0,\ldots,k_{N-1}=1}^{r} w_{k_0} \cdots w_{k_{N-1}} P_{k_0,\ldots,k_{N-1}}.$$

(iii) *In case (ii) of theorem 3.3 the assumption about break points in h is a technical consideration. The same result on the order of convergence is obtained if the equipartition is refined to include the (possibly unknown) breakpoints.*

(iv) *In case (ii) of theorem 3.3, it is possible to write down an expression for $\mathcal{C}$, using the alternative bounds given in lemmas 6.5 (i) and 6.16 (i). The difference is that in case (iii), we may express the bound in terms of properties of the T_k* only. *For general Lasota–Yorke maps, there are properties of h that are not easily expressed in terms of the maps T_k, and so in case (ii), we have a bound which depends on properties of the (unknown) invariant density h.*

(v) *Clearly theorem 3.3 (ii) and (iii) apply also to deterministic systems where there is only one T_k. For related work, see Froyland [3] and Keane et al [6].*

4. Ulam's method for Markov compositions

4.1. Invariant measures of Markov compositions

In this first section we derive an invariance condition in analogy to lemma 3.1 that enables us to define simple invariant measures on I for Markov compositions of mappings. The setting here is again very general: I is simply a metric space and each of the T_k are (Borel) measurable mappings on I.

The following result is a generalization of lemma 3.1 to Markov compositions.

Lemma 4.1. *Let $\{\mu_k\}_{k=1}^r$ be a family of Borel probability measures on I. Define the section $B_\omega = \{x \in I : (\omega, x) \in B\}$ where $B \in \mathfrak{B}(\Omega \times I)$ is a Borel measurable subset of $\Omega \times I$. A measure $\tilde{\mu}$ defined by*

$$\tilde{\mu}(B) = \int_\Omega \mu_{\omega_0}(B_\omega)\, \mathrm{d}\mathbb{P}(\omega) \qquad \text{for all} \quad B \in \mathfrak{B}(\Omega \times I) \tag{8}$$

is τ-invariant iff the family of measures $\{\mu_k\}_{k=1}^r$ is fixed under the transformation

$$(\nu_1, \ldots, \nu_r) \mapsto \left(\sum_{k=1}^r \nu_k \circ T_k^{-1} W_{1k}^*, \ldots, \sum_{k=1}^r \nu_k \circ T_k^{-1} W_{rk}^* \right) \qquad \nu_k \in \mathcal{M}(I), \tag{9}$$

where $W_{lk}^* = W_{kl} w_k / w_l$ is the transition matrix for the reversed Markov chain.

Hence, any probability measure on I of the form:

$$\mu = \sum_{k=1}^r w_k \mu_k \tag{10}$$

that arises as a projection onto $\mathcal{M}(I)$ of a measure of the form (8), is invariant on average iff the μ_k, $k = 1, \ldots, r$ are fixed under the transformation (9).

See section 6.9 for the proof. Lemma 4.1 gives one a simple way to construct invariant measures for our random system. The following extension of lemma 3.2 says that *all* absolutely continuous (with respect to $\mathbb{P} \times m$) τ-invariant measures may be written in the simple form of (8).

Lemma 4.2 (Kowalski [5]). *Assume that each T_k, $k = 1, \ldots, r$ is non-singular with respect to m. Then any $\mathbb{P} \times m$ absolutely continuous τ-invariant measure may be written in the form (8).*

Thus finding an absolutely continuous probability measure μ of the form (10) with the $\{\mu_k\}_{k=1}^r$ being fixed under the action of (9) is the only way to construct an absolutely continuous τ-invariant measure on $\Omega \times I$.

4.2. A new Perron–Frobenius operator and statement of results

For the remainder of the paper we again restrict ourselves to each T_k being a Lasota–Yorke map. We are unaware of any results paralleling those of [16] for Markov compositions of Lasota–Yorke maps, although theorem 2.1 Morita [11] assures the existence of an absolutely continuous invariant measure provided $\sum_{k=1}^{r} w_k \log(1/\inf_{x\in I}|T_k'(x)|) < 0$. However, as in the iid case, the method of proof of [11] is unsuitable for our purposes, so in section 6.1.2 we state and prove variational inequalities in analogy to lemma 6.1.

As we are primarily interested in absolutely continuous measures, we denote the density of ν_j (as in (9)) by $f^{(j)}$. Let $\hat{BV} = \prod_{i=1}^{r} BV$ denote the r-fold product of the space of functions of bounded variation. We endow the space $\hat{BV}$ with the norm $\|(f^{(1)},\ldots,f^{(r)})\| = \max_{1\leqslant k\leqslant r}\|f^{(k)}\| = \max_{1\leqslant k\leqslant r}\{\max\{\mathrm{var}\, f^{(k)}, \|f^{(k)}\|_1\}\}$. Denote by $\mathcal{P}_k : BV \circlearrowleft$, the standard Perron–Frobenius operator for the map T_k. Following (9), we define an operator $\hat{\mathcal{P}} : \hat{BV} \circlearrowleft$ by

$$\hat{\mathcal{P}}(f^{(1)},\ldots,f^{(r)}) = \left(\sum_{k=1}^{r} W_{1k}^{*}\mathcal{P}_k f^{(k)}, \sum_{k=1}^{r} W_{2k}^{*}\mathcal{P}_k f^{(k)}, \ldots, \sum_{k=1}^{r} W_{rk}^{*}\mathcal{P}_k f^{(k)}\right). \tag{11}$$

By lemma 4.1, we may construct an absolutely continuous invariant probability measure μ from a collection $(h^{(1)},\ldots,h^{(r)})$ of densities that is fixed by $\hat{\mathcal{P}}$. We will call the density of μ, $h = \sum_{k=1}^{r} w_k h^{(k)}$ an *invariant probability density* for our Markov random composition.

Theorem 4.3. *Let $\{T_1,\ldots,T_r\}$ be a collection of Lasota–Yorke maps, and assume that the Markov composition has a unique invariant density h. Equipartition the unit interval into n subintervals $I_i = [(i-1)/n, i/n)$, and define r stochastic matrices $P_n(k)$, $k = 1,\ldots,r$, by*

$$P_{n,ij}(k) = \frac{m(I_i \cap T_k^{-1} I_j)}{m(I_i)}. \tag{12}$$

Further, define the $rn \times rn$ matrix

$$S_n = \begin{pmatrix} W_{11}^{*}P_n(1) & W_{21}^{*}P_n(1) & \cdots & W_{r1}^{*}P_n(1) \\ W_{12}^{*}P_n(2) & W_{22}^{*}P_n(2) & \cdots & W_{r2}^{*}P_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ W_{1r}^{*}P_n(r) & W_{2r}^{*}P_n(r) & \cdots & W_{rr}^{*}P_n(r) \end{pmatrix}, \tag{13}$$

and let $s_n = [s_n^{(1)}|s_n^{(2)}|\ldots|s_n^{(r)}]$ be a fixed left eigenvector of S_n, where each $s_n^{(k)}$, $k = 1,\ldots,r$ is a vector of length n satisfying $\sum_{i=1}^{n} s_{n,i}^{(k)} = 1$. Define the approximate invariant density

$$h_n = \sum_{i=1}^{n}\left(\frac{\sum_{k=1}^{r} w_k s_{n,i}^{(k)}}{m(I_i)}\right)\chi_{I_i}. \tag{14}$$

Then setting

$$\alpha_l' = \sum_{k=1}^{r} \frac{W_{lk}^{*}}{\inf_{x\in I}|T_k'(x)|}, \tag{15}$$

$$\beta_l' = \sum_{k=1}^{r} W_{lk}^{*}\frac{\sup_{x\in I}|T_k''(x)|}{\inf_{x\in I}|T_k'(x)|^2}, \tag{16}$$

and

$$\eta_l' = \sum_{k=1}^{r} W_{lk}^{*}\frac{\sum_{\ell=0}^{q_k}\theta_{k,\ell}}{\inf_{x\in I}|T_k'(x)|}, \tag{17}$$

with $\alpha' = \max_{1\leqslant l\leqslant r}\alpha_l'$ and $\beta' = \max_{1\leqslant l\leqslant r}\beta_l'$.

(i) *If* $\alpha' < \frac{1}{2}$, $\|h_n - h\|_1 \to 0$ *as* $n \to \infty$.

(ii) *If* $\max_{1\leqslant l\leqslant r}(\alpha'_l + \eta'_l) + \beta'/2 < 1$, *and the endpoints of the partition* $\{I_1, \ldots, I_n\}$ *contain all points where there is a break in the* C^1 *behaviour of any* $h^{(k)}$, $k = 1, \ldots, r$, *then there exists a constant* $\mathcal{C} < \infty$ *such that* $\|h_n - h\|_1 \leqslant \mathcal{C} \log n/n$.

(iii) *If* $\alpha' + \beta'/2 < 1$ *and each* T_k *is a* C^2 *circle map, then the constant* $\mathcal{C}$ *above may be written in terms of fundamental constants of the maps* T_k. *Set* $C = \max_{1\leqslant k\leqslant r} \mathrm{Lip}(\log|T'_k|)$ *and* $\lambda = \min_{1\leqslant k\leqslant r} \inf_x |T'_k(x)|$ *(assuming* $\lambda > 1$*). Then*

$$\|h - h_n\|_1 \leqslant (\mathrm{e}^{C/(\lambda-1)n} - 1) \tag{18}$$
$$\cdot\left(\left(\max_{1\leqslant k\leqslant r}\sum_{l=1}^{r} W^*_{lk}\right) \times \inf_{0<\delta<1}\left\{\left(2 + \frac{\delta}{1-\delta}\right)\left(\left[\frac{\log(4rn/\delta)}{-\log(\alpha' + \beta'/2)}\right] + 1\right) - 1\right\} + \frac{1}{2}\right),$$

where $[\cdot]$ *denotes the integer part.*

Remark 4.4. *(i) Using the arguments of [8], one may prove the existence of an invariant density* h *if* $\alpha' < 1$. *As we outline in the examples section, uniqueness is guaranteed in cases (ii) and (iii) if the hypotheses on* α' *and* β' *are satisfied.*

(ii) *If* $\frac{1}{2} < \alpha' < 1$, *then we may produce a Ulam approximation using higher iterates of the mappings. It is simple to show by induction that*

$$\alpha'_N := \max_{1\leqslant l\leqslant r}\sum_{k_0}\cdots\sum_{k_{N-1}} W^*_{lk_{N-1}} W^*_{k_{N-1}k_{N-2}}\cdots W^*_{k_1k_0}\frac{1}{\inf_{x\in I}|(T_{k_{N-1}}\circ\cdots T_{k_1}\circ T_{k_0})'(x)|}$$
$$\leqslant (\alpha')^N. \tag{19}$$

We choose N *such that* $\alpha'_N < \frac{1}{2}$, *and construct matrices*

$$P_{k_0,\ldots,k_{N-1},n,ij} = \frac{m(I_i \cap (T_{k_{N-1}}\circ\cdots\circ T_{k_0})^{-1}I_j)}{m(I_i)}.$$

There will be at most r^N *of these matrices, depending on the transitions allowed by* W. *We combine these matrices to form a collection of* r^2 *matrices,*

$$P^{(N)}_n(k_0, k_N) := \sum_{k_1,\ldots,k_{N-1}=1}^{r} W^*_{k_Nk_{N-1}}\cdots W^*_{k_1k_0} P_{k_0,\ldots,k_{N-1},n},$$

and define the N*-step Ulam matrix as*

$$S^{(N)}_n = \begin{pmatrix} P^{(N)}_n(1,1) & P^{(N)}_n(1,2) & \cdots & P^{(N)}_n(1,r) \\ P^{(N)}_n(2,1) & P^{(N)}_n(2,2) & \cdots & P^{(N)}_n(2,r) \\ \vdots & \vdots & \ddots & \vdots \\ P^{(N)}_n(r,1) & P^{(N)}_n(r,2) & \cdots & P^{(N)}_n(r,r) \end{pmatrix}. \tag{20}$$

(iii) *Remark (iv) of remark 3.4 holds in the Markov situation using the bounds of lemmas 6.6 (i) and 6.17 (i).*

5. Examples

We apply the main results to some specific random systems.

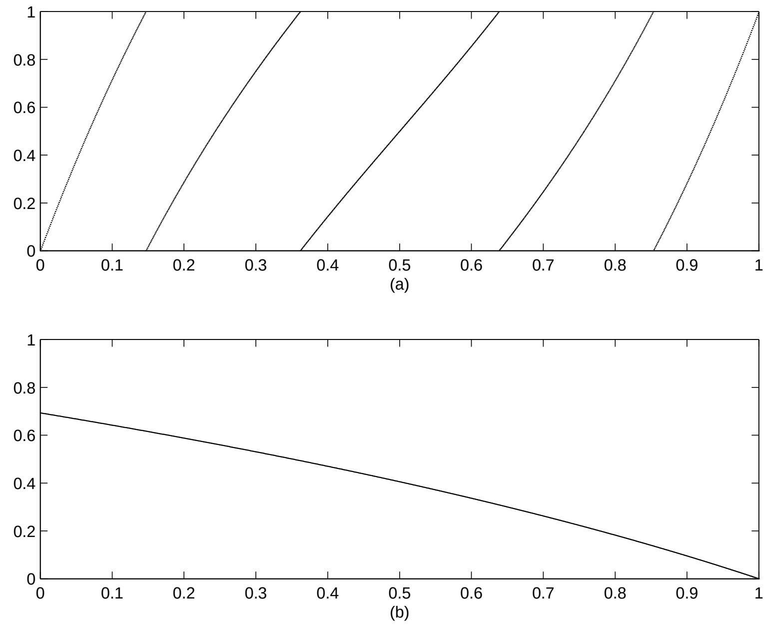

Figure 1. Graphs of (*a*) T_1 and (*b*) T_2.

5.1. Convergence

We first consider statement (i) of theorem 3.3, which guarantees convergence of the Ulam estimates h_n to the unique invariant density h in the L^1 norm.

Define $T_1, T_2 : I \circlearrowleft$ by $T_1 x = 6x^3 - 9x^2 + 8x \pmod 1$ (with $T_1(1) = 1$) and $T_2 x = \log(2 - x)$; see figure 1. The map T_2 is nowhere expanding, however, if it applied sufficiently infrequently, the resulting random system will have a unique absolutely continuous invariant measure; in addition we may approximate this measure via Ulam's method. We put $w_1 = 0.9$ and $w_2 = 0.1$; this leads to a value of $\alpha \approx 0.409\,05$. Our maps satisfy conditions (i) and (ii) at the beginning of section 3.2, so that we are guaranteed a unique invariant density for this iid composition. Theorem 3.3 (i) states that the Ulam approximations h_n converge to this unique invariant density h. If we had chosen w_1 and w_2 such that $\frac{1}{2} < \alpha < 1$, then we could have constructed Ulam matrices from higher powers of T_1 and T_2 as described in remark 3.4 (ii).

Considering now statement (ii) of theorem 3.3, we find that $\theta_{1,\ell} = 0$ for $\ell = 1, \ldots, 6$, and that $\theta_{2,1} = 1$ and $\theta_{2,2} = 0$, yielding $\eta = 0.2$, while $\beta \approx 1.7225$. Thus $\alpha + \eta + \beta/2 \approx 1.4703 > 1$, and so theorem 3.3 (ii) cannot be immediately applied to show that $\|h - h_n\|_1 \to 0$ like $O(\log n/n)$.

Although we have used simple maps for ease of description in this example, we could equally well have chosen general Lasota–Yorke maps with finitely many breaks in continuity.

5.2. Bounds

We consider the case where $T_1, T_2 : S^1 \circlearrowleft$ are C^2 expanding circle maps, and derive bounds for our approximation error $\|h - h_n\|_1$.

Let $T_1 x = 6x^3 - 9x^2 + 8x \pmod 1$ as before, and define $T_2 x = -4x + \sin(2\pi x)/5 \pmod 1$; see figures 2(*a*) and (*b*), respectively.

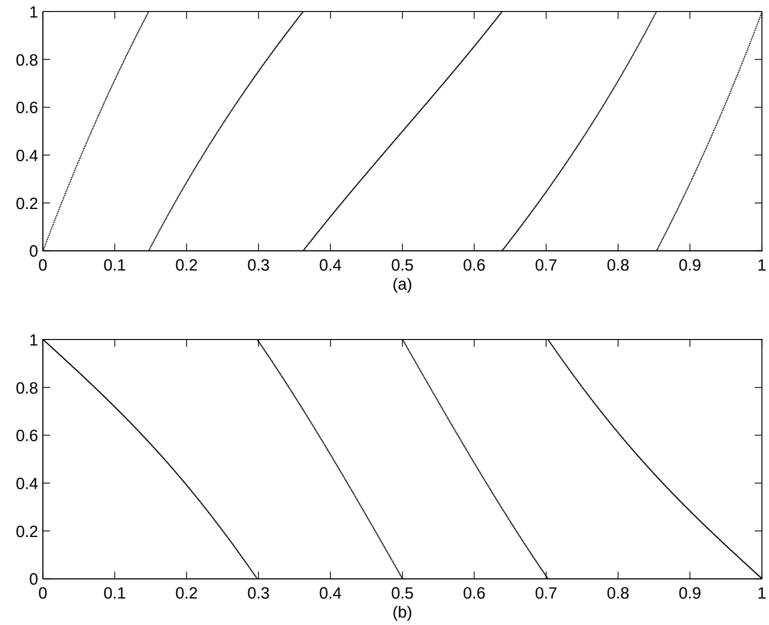

Figure 2. Graphs of (*a*) T_1 and (*b*) T_2.

Table 1. iid error bounds.

Number of partition sets n	Bound for $\|\pi_n(h) - h_n\|_1$	Bound for $\|h - \pi_n(h)\|_1$	Bound for $\|h - h_n\|_1$
100	0.943 91	0.006 55	0.950 46
200	0.513 23	0.003 26	0.516 49
500	0.227 41	0.001 30	0.228 71
1000	0.122 13	0.000 65	0.122 78
2000	0.065 22	0.000 33	0.065 54
5000	0.028 28	0.000 13	0.028 41

5.2.1. iid case. We put $w_1 = \frac{2}{3}$ and $w_2 = \frac{1}{3}$. One has that $\alpha \approx 0.253\,89$, $\beta \approx 1.0748$, so that $\alpha + \beta/2 \approx 0.791\,31$. This tells us that the Perron–Frobenius operator $\mathcal{P}$ is a contraction in the BV norm when restricted to test functions of zero mean (when restricted to the space BV_0; see lemma 6.1 for definitions). It follows immediately that there is a unique invariant density h for this iid composition.

We also compute that $C \approx 2.267\,79$ (using T_1 as it has a larger C value than T_2) and $\lambda \approx 2.743\,36$ (using T_2 as it has a larger λ value than T_1). These values may now be inserted into (7) to obtain a bound for $\|h - h_n\|_1$ for any n. The separate bounds for $\|h - \pi_n(h)\|_1$ and $\|\pi_n(h) - h_n\|_1$ are displayed in table 1.

5.2.2. Markov case. We choose the matrix $W = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix}$; it has a fixed left eigenvector $w = [\frac{2}{3}\ \frac{1}{3}]$ as in the iid case. One may calculate that $\alpha' = \alpha'_1 \approx 0.325\,12$ and that $\beta' = \beta'_1 \approx 0.877\,56$, so that $\alpha' + \beta'/2 \approx 0.763\,90$. Once again, we see that the Perron–Frobenius operator $\hat{\mathcal{P}}$ is a contraction when restricted to $B\hat{V}_0$ (see lemma 6.2 for the definition),

Table 2. Markov error bounds.

Number of partition sets n	Bound for $\|\pi_n(h) - h_n\|_1$	Bound for $\|h - \pi_n(h)\|_1$	Bound for $\|h - h_n\|_1$
100	1.340 28	0.006 55	1.346 82
200	0.723 37	0.003 26	0.726 64
500	0.317 98	0.001 30	0.319 29
1000	0.169 92	0.000 65	0.170 57
2000	0.090 34	0.000 33	0.090 66
5000	0.038 97	0.000 13	0.039 10

so that one is guaranteed a unique invariant density for this Markov composition.

The values for C and λ remain the same as in the iid case, and we may substitute these and the above values into (18) to produce bounds analogous to the iid case; see table 2. Note that the bounds for $\|h - \pi_n(h)\|_1$ do not change.

6. Proofs

6.1. Fundamental inequalities for the Perron–Frobenius operator

6.1.1. iid case. Let BV denote the space of L^1 functions on I which are of bounded variation.

Lemma 6.1. *Let $f \in BV$, suppose that each T_k, $k = 1, \ldots, r$ is a Lasota–Yorke map, and set q^* as in definition 2.2. Define a norm $\|\cdot\|$ on BV by $\|f\| = \max\{\mathrm{var}\, f, \|f\|_1\}$ for $f \in BV$, and set $BV_0 = \{f \in BV : \int f\,\mathrm{d}m = 0\}$. Define α, β and η as in (4), (5) and (6) respectively. Then*

(i) $$\mathrm{var}\,\mathcal{P} f \leqslant 2\alpha\, \mathrm{var}\, f + (2q^*\alpha + \beta)\|f\|_1 \qquad \text{for} \quad f \in BV. \tag{21}$$

(ii) $$\|\mathcal{P} f\| \leqslant (\alpha + \eta + \beta/2)\|f\| \qquad \text{for} \quad f \in BV_0. \tag{22}$$

If in addition, each T_k is a C^2 circle map, then

(iii) $$\|\mathcal{P} f\| \leqslant (\alpha + \beta/2)\|f\| \qquad \text{for} \quad f \in BV_0. \tag{23}$$

We remark that the coefficients $\alpha + \eta + \beta/2$ and $\alpha + \beta/2$ are not the best possible, and we have chosen these values for simplicity of presentation.

Proof. Once part (i) is established, the other bounds follow easily, so we prove this first. Part (i) follows as in the proof of theorem 1 [16], except that a small improvement can be made.

Let $\mathcal{B}_k := \{[b_0^k, b_1^k], [b_1^k, b_2^k], \ldots, [b_{q_k-1}^k, b_{q_k}^k]\}$ and $\mathcal{B}^* := \{[b_0^*, b_1^*], [b_1^*, b_2^*], \ldots, [b_{q^*-1}^*, b_{q^*}^*]\}$ be as in definition 2.2. We write $\mathrm{var}\,\mathcal{P} f \leqslant \sum_{k=1}^r w_k\, \mathrm{var}\,\mathcal{P}_k f$, and proceed to bound $\mathrm{var}\,\mathcal{P}_k f$, $k = 1, \ldots, r$, individually.

One has the standard inequality

$$\mathrm{var}\,\mathcal{P}_k f \leqslant \sum_{B\in\mathcal{B}^*} \mathrm{var}_B \frac{f}{|T_k'|} + \sum_{\ell=1}^{q^*} \left(\left| \frac{f(b_{\ell-1}^*)}{T_k'(b_{\ell-1}^*)} \right| + \left| \frac{f(b_\ell^*)}{T_k'(b_\ell^*)} \right| \right); \tag{24}$$

see [8] for example. We treat the first term:

$$\mathrm{var}_B \frac{f}{|T_k'|} \leqslant \frac{1}{\inf_B |T_k'|} \mathrm{var}_B\, f + \frac{\sup_B |T_k''|}{\inf_B |T_k'|^2} \int_B |f|\,\mathrm{d}m$$

via a standard argument [8]. So

$$\sum_{k=1}^r w_k \sum_{B\in\mathcal{B}^*} \mathrm{var}_B \frac{f}{|T_k'|} \leqslant \sum_{k=1}^r w_k \sum_{B\in\mathcal{B}^*} \frac{1}{\inf_B |T_k'|} \mathrm{var}_B\, f + \sum_{k=1}^r w_k \sum_{B\in\mathcal{B}^*} \frac{\sup_B |T_k''|}{\inf_B |T_k'|^2} \int_B |f|\,\mathrm{d}m$$

$$\leqslant \sum_{B\in\mathcal{B}^*}\left(\max_{B\in\mathcal{B}^*}\sum_{k=1}^{r} w_k \frac{1}{\inf_B |T_k'|}\right)\operatorname{var}_B f + \sum_{k=1}^{r} w_k \frac{\sup_I |T_k''|}{\inf_I |T_k'|^2}\int_I |f|\,\mathrm{d}m$$

$$\leqslant \underbrace{\left(\max_{B\in\mathcal{B}^*}\sum_{k=1}^{r} w_k \frac{1}{\inf_B |T_k'|}\right)}_{\alpha^*}\operatorname{var} f + \sum_{k=1}^{r} w_k \frac{\sup_I |T_k''|}{\inf_I |T_k'|^2}\int_I |f|\,\mathrm{d}m.$$

We may now refine $\mathcal{B}^*$ to make $\alpha^* = \max_{B\in\mathcal{B}^*}\sum_{k=1}^{r} w_k \frac{1}{\inf_B |T_k'|}$ as close to α as we like.

Now for the second term:

$$\sum_{k=1}^{r} w_k \sum_{\ell=1}^{q^*}\left(\left|\frac{f(b_{\ell-1}^*)}{T_k'(b_{\ell-1}^*)}\right| + \left|\frac{f(b_\ell^*)}{T_k'(b_\ell^*)}\right|\right) = \sum_{\ell=1}^{q^*}\left(\sum_{k=1}^{r} w_k \frac{1}{|T_k'(b_{\ell-1}^*)|}\right)|f(b_{\ell-1}^*)|$$
$$+ \sum_{\ell=1}^{q^*}\left(\sum_{k=1}^{r} w_k \frac{1}{|T_k'(b_\ell^*)|}\right)|f(b_\ell^*)|$$
$$\leqslant \alpha \sum_{\ell=1}^{q^*} |f(b_{\ell-1}^*)| + |f(b_\ell^*)|$$
$$\leqslant \alpha(\operatorname{var} f + 2q^*\|f\|_1),$$

since $|f(b_{\ell-1}^*)| + |f(b_\ell^*)| \leqslant 2\inf_{[b_{\ell-1}^*,b_\ell^*]} f + \operatorname{var}_{[b_{\ell-1}^*,b_\ell^*]} f \leqslant 2\int_I f\,\mathrm{d}m(x) + \operatorname{var}_{[b_{\ell-1}^*,b_\ell^*]} f$.

Combining these two terms we have

$$\operatorname{var}\mathcal{P}f \leqslant (\alpha^* + \alpha)\operatorname{var} f + (2q^*\alpha + \beta)\|f\|_1$$

as required.

For the proof of theorem 3.3(i), we shall see that it is important to keep the coefficient of var f small. For the proof of theorem 3.3(ii), however, it is desirable to reduce the combination of both the coefficients of var f and $\|f\|_1$. Because of this, in proving part (ii) of the present lemma, we modify (24) to obtain:

$$\operatorname{var}\mathcal{P}_k f \leqslant \sum_{B\in\mathcal{B}_k}\operatorname{var}_B \frac{f}{|T_k'|} + \sum_{\ell=0}^{q_k}\theta_{k,\ell}\max\left\{\left|\frac{f(b_\ell^{k,-})}{T_k'(b_\ell^{k,-})}\right|, \left|\frac{f(b_\ell^{k,+})}{T_k'(b_\ell^{k,+})}\right|\right\}. \tag{25}$$

The first term of (25) is the same as the first term of (24), and is bounded as before.

When dealing with the second term, we are able to set $\|f\|_\infty \leqslant \operatorname{var} f$ as $f \in BV_0$:

$$\sum_{k=1}^{r} w_k \sum_{\ell=0}^{q_k}\theta_{k,\ell}\max\left\{\left|\frac{f(b_\ell^{k,-})}{T_k'(b_\ell^{k,-})}\right|, \left|\frac{f(b_\ell^{k,+})}{T_k'(b_\ell^{k,+})}\right|\right\}$$
$$\leqslant \left(\sum_{k=1}^{r} w_k \sum_{\ell=0}^{q_k}\frac{\theta_{k,\ell}}{\min\{|T_k'(b_\ell^{k,-})|, |T_k'(b_\ell^{k,+})|\}}\right)\operatorname{var} f$$
$$\leqslant \left(\sum_{k=1}^{r} w_k \frac{\sum_{\ell=0}^{q_k}\theta_{k,\ell}}{\inf_{x\in I}|T_k'(x)|}\right)\operatorname{var} f.$$

Combining these two terms we have

$$\operatorname{var}\mathcal{P}f \leqslant (\alpha^* + \eta)\operatorname{var} f + \beta\|f\|_1.$$

For $f \in BV_0$, one has $\|f\|_1 \leqslant (\frac{1}{2})\operatorname{var} f$ (see lemma 4 [5], for example), so that $\|\mathcal{P}f\| = \operatorname{var}\mathcal{P}f$ and $\|f\| = \operatorname{var} f$. Now:

$$\|\mathcal{P}f\| = \max\{\operatorname{var}\mathcal{P}f, \|\mathcal{P}f\|_1\} = \operatorname{var}\mathcal{P}f \leqslant (\alpha + \eta)\operatorname{var} f + \beta\|f\|_1 \leqslant (\alpha + \eta + \beta/2)\operatorname{var} f$$
$$= (\alpha + \eta + \beta/2)\|f\|,$$

completing the proof of part (ii).

For the case of circle maps, we use the bound of part (ii), and delete the contributions from the branches of monotonicity not being onto (the 'second term' in the preceding argument). This leaves us with $\operatorname{var} \mathcal{P} f \leqslant \alpha \operatorname{var} f + \beta \|f\|_1$. The bound for $\|\mathcal{P}|_{BV_0}\|$ follows as above. □

6.1.2. Markov case

Lemma 6.2. *Let $\hat{f} = (f^{(1)}, \ldots, f^{(r)}) \in \hat{BV}$, suppose that each T_k, $k = 1, \ldots, r$ is a Lasota–Yorke map, and set q^* as in definition 2.2. Set $\hat{BV}_0 = \{\hat{f} \in \hat{BV} : \int f^{(k)}\, \mathrm{d}m = 0$ for all $k = 1, \ldots, r\}$. Define α_l', β_l' and η_l' as in (15), (16) and (17) respectively, setting $\alpha' = \max_{1\leqslant l\leqslant r} \alpha_l'$ and $\beta' = \max_{1\leqslant l\leqslant r} \beta_l'$. Then*

(i)

$$\max_{1\leqslant k\leqslant r} \operatorname{var}(\hat{\mathcal{P}}\hat{f})^{(k)} \leqslant 2\alpha' \max_{1\leqslant k\leqslant r} \operatorname{var} f^{(k)} + \max_{1\leqslant l\leqslant r} (2q^*\alpha_l' + \beta_l') \max_{1\leqslant k\leqslant r} \|f^{(k)}\|_1 \qquad \textit{for } \hat{f} \in \hat{BV}. \tag{26}$$

(ii)

$$\|\hat{\mathcal{P}}\hat{f}\| \leqslant \left(\max_{1\leqslant l\leqslant r}(\alpha_l' + \eta_l') + \beta'/2\right) \|\hat{f}\| \qquad \textit{for} \quad \hat{f} \in \hat{BV}_0. \tag{27}$$

If in addition, each T_k is a C^2 circle map,

(iii)

$$\|\hat{\mathcal{P}}\hat{f}\| \leqslant (\alpha' + \beta'/2)\|\hat{f}\| \qquad \textit{for} \quad \hat{f} \in \hat{BV}_0. \tag{28}$$

Proof. Let $\mathcal{B}_k$ and $\mathcal{B}^*$ be defined as in the proof of lemma 6.1.

First note that

$$\operatorname{var}(\hat{\mathcal{P}}\hat{f})^{(l)} = \operatorname{var}\left(\sum_{k=1}^{r} W_{lk}^* \mathcal{P}_k f^{(k)}\right) \tag{29}$$

$$\leqslant \sum_{k=1}^{r} W_{lk}^* \left(\sum_{B\in\mathcal{B}^*} \operatorname{var}_B \frac{f^{(k)}}{|T_k'|} + \sum_{\ell=1}^{q^*} \left(\left|\frac{f^{(k)}(b_{\ell-1}^*)}{T_k'(b_{\ell-1}^*)}\right| + \left|\frac{f^{(k)}(b_\ell^*)}{T_k'(b_\ell^*)}\right|\right)\right).$$

The first term is treated as follows:

$$\sum_{k=1}^{r} W_{lk}^* \left(\sum_{B\in\mathcal{B}^*} \operatorname{var}_B \frac{f^{(k)}}{|T_k'|}\right) \leqslant \sum_{k=1}^{r} W_{lk}^* \sum_{B\in\mathcal{B}^*} \left(\frac{1}{\inf_{x\in B} |T_k'|} \operatorname{var}_B f^{(k)} + K_k \int_B |f^{(k)}|\, \mathrm{d}m(x)\right), \tag{30}$$

as in the proof of lemma 6.1, where $K_k = \sup_{x\in I} |T_k''(x)|/\inf_{x\in I} |T_k'(x)|^2$. Continuing,

$$(30) \leqslant \sum_{k=1}^{r} W_{lk}^* \frac{1}{\inf_{x\in I} |T_k'(x)|} \operatorname{var} f^{(k)} + \sum_{k=1}^{r} W_{lk}^* K_k \|f^{(k)}\|_1. \tag{31}$$

As for the second term,

$$\sum_{k=1}^{r} W_{lk}^* \sum_{\ell=1}^{q^*} \left(\left|\frac{f^{(k)}(b_{\ell-1}^*)}{T_k'(b_{\ell-1}^*)}\right| + \left|\frac{f^{(k)}(b_\ell^*)}{T_k'(b_\ell^*)}\right|\right) \leqslant \sum_{k=1}^{r} W_{lk}^* \frac{1}{\inf_{x\in I} |T_k'(x)|} (\operatorname{var} f^{(k)} + 2q^* \|f^{(k)}\|_1), \tag{32}$$

as in the proof of lemma 6.1. Combining (31) and (32), we obtain

$$\begin{aligned}
\operatorname{var}(\hat{\mathcal{P}}\hat{f})^{(l)} &\leqslant 2\left(\sum_{k=1}^{r} W_{lk}^* \frac{1}{\inf_{x\in I} |T_k'(x)|}\right) \max_{1\leqslant k\leqslant r} \operatorname{var} f^{(k)} \\
&\quad + \left(\sum_{k=1}^{r} W_{lk}^* \left(K_k + \frac{2q^*}{\inf_{x\in I} |T_k'(x)|}\right)\right) \max_{1\leqslant k\leqslant r} \|f^{(k)}\|_1 \\
&= 2\alpha_l' \max_{1\leqslant k\leqslant r} \operatorname{var} f^{(k)} + (2q^*\alpha_l' + \beta_l') \max_{1\leqslant k\leqslant r} \|f^{(k)}\|_1,
\end{aligned} \tag{33}$$

as required.

For part (ii), we use a modification of the inequality (29):

$$\mathrm{var}\left(\hat{\mathcal{P}}\hat{f}\right)^{(l)} = \mathrm{var}\left(\sum_{k=1}^{r} W^*_{lk}\mathcal{P}_k f^{(k)}\right)$$
$$\leqslant \sum_{k=1}^{r} W^*_{lk}\left(\sum_{B\in\mathcal{B}_k}\mathrm{var}_B\,\frac{f^{(k)}}{|T_k'|} + \sum_{\ell=0}^{q_k}\theta_{k,\ell}\max\left\{\left|\frac{f^{(k)}(b_\ell^{k,-})}{T_k'(b_\ell^{k,-})}\right|,\left|\frac{f^{(k)}(b_\ell^{k,+})}{T_k'(b_\ell^{k,+})}\right|\right\}\right).$$

The first term is bounded as above. As for the second term,

$$\sum_{k=1}^{r} W^*_{lk}\sum_{\ell=0}^{q_k}\theta_{k,\ell}\max\left\{\left|\frac{f^{(k)}(b_\ell^{k,-})}{T_k'(b_\ell^{k,-})}\right|,\left|\frac{f^{(k)}(b_\ell^{k,+})}{T_k'(b_\ell^{k,+})}\right|\right\} \leqslant \left(\sum_{k=1}^{r} W^*_{lk}\frac{\sum_{\ell=0}^{q_k}\theta_{k,\ell}}{\inf_{x\in I}|T_k'(x)|}\right)\max_{1\leqslant k\leqslant r}\mathrm{var}\,f^{(k)} \tag{34}$$

as in the proof of lemma 6.1 (recalling $\|f^{(k)}\|_\infty \leqslant \mathrm{var}\,f^{(k)}$ as $\hat{f}\in \hat{BV}_0$). Combining (31) and (34), we obtain

$$\mathrm{var}\left(\hat{\mathcal{P}}\hat{f}\right)^{(l)} \leqslant \left(\sum_{k=1}^{r} W^*_{lk}\frac{1+\sum_{\ell=0}^{q_k}\theta_{k,\ell}}{\inf_{x\in I}|T_k'(x)|}\right)\max_{1\leqslant k\leqslant r}\mathrm{var}\,f^{(k)} + \left(\sum_{k=1}^{r} W^*_{lk}K_k\right)\max_{1\leqslant k\leqslant r}\|f^{(k)}\|_1 \tag{35}$$
$$= (\alpha_l' + \eta_l')\max_{1\leqslant k\leqslant r}\mathrm{var}\,f^{(k)} + \beta_l'\max_{1\leqslant k\leqslant r}\|f^{(k)}\|_1,$$

and the result follows. Part (iii) follows as in the iid case. □

6.2. *Proof of part (i) of theorems 3.3 and 4.3*

6.2.1. *iid case.*

Proof of theorem 3.3 (i). First, recall that $\alpha < 1$ already guarantees the existence of h; see the discussion at the beginning of section 3.2. Denote $F_n = \{f_n \in BV : f_n = (1/n)\sum_{i=1}^n f_{n,i}\chi_{I_i}$, for some $f_{n,i}\in\mathbb{R}\}$, and let $h_n \in F_n$ be as in (3). Define a mapping $\pi_n : BV \to F_n$ by $\pi_n(f) = (1/n)\sum_{i=1}^n(\int_{I_i} f\,\mathrm{d}m)\chi_{I_i}$. We follow the proofs of lemma 2.7 [9] and theorem 1 [9]. Let $[\pi_n\mathcal{P}]$ represent the matrix representation (with respect to the basis $\{\chi_{I_1},\ldots,\chi_{I_n}\}$) of the projection π_n composed with the Perron Frobenius operator defined in (1). It is easy to show that $[\pi_n\mathcal{P}]_{ij} = P_{n,ij}$, so one has $h_n = \pi_n\mathcal{P}h_n$. Now

$$\mathrm{var}\,h_n = \mathrm{var}(\pi_n\mathcal{P}h_n) \leqslant \mathrm{var}\,\mathcal{P}h_n \leqslant 2\alpha\,\mathrm{var}\,h_n + (2q^*\alpha+\beta)\|h_n\|_1,$$

and so $\mathrm{var}\,h_n \leqslant ((2q^*\alpha+\beta)/(1-2\alpha))\|h_n\|_1$. Thus the sequence $\{h_n\}_{n=n_0}^\infty$ is weakly sequentially compact in L^1 and each convergent subsequence converges to h in L^1. This implies $\lim_{n\to\infty} h_n = h$. □

6.2.2. *Markov case.*

Proof of theorem 4.3 (i). We follow along the lines of the above proof. Denote $\hat{F}_n = \prod_{k=1}^r F_n$ and define the projection $\hat{\pi}_n : \hat{BV} \to \hat{F}_n$ by $\hat{\pi}_n((f^{(1)},\ldots,f^{(r)})) = (\pi_n(f^{(1)}),\ldots,\pi_n(f^{(r)}))$. Note that the matrix representation of $[\hat{\pi}_n\hat{\mathcal{P}}]$ with respect to the basis $\prod_{k=1}^{r}\{\chi_{I_1},\ldots,\chi_{I_n}\}$ is simply S_n, and so $(h_n^{(1)},\ldots,h_n^{(r)}) = \hat{\pi}_n\hat{\mathcal{P}}((h_n^{(1)},\ldots,h_n^{(r)}))$, where $h_n^{(k)} := \sum_{i=1}^{n}(s_{n,i}^{(k)}/m(I_i))\chi_{I_i}$. As in the iid case,

$$\begin{aligned}\max_{1\leqslant k\leqslant r}\operatorname{var} h_n^{(k)} &= \max_{1\leqslant k\leqslant r}\operatorname{var}(\hat{\pi}_n\hat{\mathcal{P}}(h_n^{(1)},\ldots,h_n^{(r)}))^{(k)}\\ &\leqslant \max_{1\leqslant k\leqslant r}\operatorname{var}\hat{\mathcal{P}}(h_n^{(1)},\ldots,h_n^{(r)})^{(k)}\\ &\leqslant 2\alpha'\max_{1\leqslant k\leqslant r}\operatorname{var} h_n^{(k)} + \max_{1\leqslant l\leqslant r}(2q^*\alpha_l' + \beta_l')\max_{1\leqslant k\leqslant r}\|h_n^{(k)}\|_1,\end{aligned}$$

and so $\max_{1\leqslant k\leqslant r}\operatorname{var} h_n^{(k)} \leqslant (\max_{1\leqslant l\leqslant r}(2q^*\alpha_l'+\beta_l')/(1-2\alpha'))\max_{1\leqslant k\leqslant r}\|h_n^{(k)}\|_1$. Thus the sequence $\{(h_n^{(1)},\ldots,h_n^{(r)})\}_{n=n_0}^{\infty} \subset \hat{BV}$ is weakly sequentially compact in $\prod_{k=1}^{r} L^1$ and each convergent subsequence converges to $(h^{(1)},\ldots,h^{(r)})$, the unique fixed point of $\hat{\mathcal{P}}$. □

For parts (ii) and (iii) of theorems 3.3 and 4.3, we need to do much more work, and the remainder of section 6 is devoted to their proof.

6.3. *Error estimates for eigenvectors*

6.3.1. iid case. Following observation 3.4 [3], we note that the matrix $\tilde{P}_{n,ij} := \sum_{k=1}^{r} w_k(\mu(I_i\cap T_k^{-1}I_j)/\mu(I_i))$, has the vector $\tilde{p}_n = [\mu(I_1),\ldots,\mu(I_n)]$ as a fixed left eigenvector (since μ is fixed by $\mathcal{D}$). The vector $\tilde{p}_n$ gives exactly the correct weight to each partition set, so our concern is with the deviation of our approximate vector p_n from $\tilde{p}_n$. Whenever talking about norms on vectors, we shall denote the standard L^1 vector norm as $\|\cdot\|_m$ to avoid confusion with the L^1 norm on functions, which will be denoted $\|\cdot\|_1$. We shall show in section 6.6.1 that the matrix P has a unique (up to scalar multiples) fixed left eigenvector under the conditions of theorem 3.3 (ii) or (iii). Thus, the matrix P is ergodic (irreducible and aperiodic), and we may use the fundamental inequality [17]

$$\|\tilde{p}_n - p_n\|_m \leqslant \|\tilde{P}_n - P_n\|_m\|(I_n - P_n + P_n^{\infty})^{-1}\|_m, \tag{36}$$

where $P_{n,ij}^{\infty} = p_{n,j}$ and I_n is the $n\times n$ identity matrix. We may rewrite (36) as

$$\|\pi_n(h) - h_n\|_1 = \sum_{i=1}^{n}|\mu(I_i) - p_{n,i}| \leqslant \|\tilde{P}_n - P_n\|_m\|Z_n\|_m, \tag{37}$$

with $Z_n = (I_n - P_n + P_n^{\infty})^{-1}$. In section 6.5.1, we derive bounds for $\|\tilde{P}_n - P_n\|_m$, and in section 6.6.1 we produce bounds for $\|Z_n\|_m$. This will provide us with a bound for the difference in the L^1 vector norm between $\tilde{p}_n$ and p_n, and hence a bound for $\|\pi_n(h) - h_n\|_1$.

6.3.2. Markov case. The invariant measure μ may be decomposed as $\sum_{k=1}^{r} w_k\mu_k$, where the μ_k are fixed under (9). We construct matrices

$$\tilde{P}_n(k) = \frac{\mu_k(I_i\cap T_k^{-1}I_j)}{\mu_k(I_i)}, \tag{38}$$

and put them together to form

$$\tilde{S}_n = \begin{pmatrix} W_{11}^*\tilde{P}_n(1) & W_{21}^*\tilde{P}_n(1) & \cdots & W_{r1}^*\tilde{P}_n(1)\\ W_{12}^*\tilde{P}_n(2) & W_{22}^*\tilde{P}_n(2) & \cdots & W_{r2}^*\tilde{P}_n(2)\\ \vdots & \vdots & \ddots & \vdots\\ W_{1r}^*\tilde{P}_n(r) & W_{2r}^*\tilde{P}_n(r) & \cdots & W_{rr}^*\tilde{P}_n(r)\end{pmatrix}. \tag{39}$$

We denote $\tilde{S}^{(kl)}_{n,ij} = W^*_{lk}\tilde{P}(k)_{ij}$, $1 \leqslant i, j \leqslant n$, $1 \leqslant k, l \leqslant r$; that is, the (i, j)th entry of the (k, l)th block. It is easy to check that the vector

$$\tilde{s}_n := [\mu_1(I_1), \ldots, \mu_1(I_n), \mu_2(I_1), \ldots, \mu_2(I_n), \ldots\ldots, \mu_r(I_1), \ldots, \mu_r(I_n)]$$

is a fixed left eigenvector of $\tilde{S}_n$. By $\tilde{s}^{(k)}_{n,i} = \mu_k(I_i)$ we denote the ith entry of the kth block of $\tilde{s}_n$. Under the conditions of theorem 4.3 (ii) or (iii), S_n will have a unique fixed left eigenvector (up to scalar multiples) and we may apply inequality (36) to obtain

$$\|\tilde{s}_n - s_n\|_m \leqslant \|\tilde{S}_n - S_n\|_m \|(I_{rn} - S_n + S^\infty_n)^{-1}\|_m, \tag{40}$$

where $(S^\infty_n)^{(kl)}_{ij} = s^{(l)}_j$. As in the iid case, we denote $\hat{Z}_n = (I_{rn} - S_n + S^\infty_n)^{-1}$ and in later sections will bound both $\|\tilde{S}_n - S_n\|_m$ and $\|\hat{Z}_n\|_m$.

Finally,

$$\|\pi_n(h) - h_n\|_1 = \sum_{i=1}^n \left| \sum_{k=1}^r w_k(\tilde{s}^{(k)}_{n,i} - s^{(k)}_{n,i}) \right| \leqslant \sum_{k=1}^r \sum_{i=1}^n |\tilde{s}^{(k)}_{n,i} - s^{(k)}_{n,i}| = \|\tilde{s}_n - s_n\|_m. \tag{41}$$

Thus we may bound $\|\pi_n(h) - h_n\|_1$ using (40).

6.4. Renyi estimates for the invariant density

This section derives the necessary bounds for the regularity of the invariant density h in terms of fundamental constants of the maps T_k, when each T_k is a $C^{1+\mathrm{Lip}}$ expanding circle map.

6.4.1. iid case.

Lemma 6.3. *Suppose that each T_k is an expanding $C^{1+\mathrm{Lip}}$ map of the circle. Define $\lambda = \min_{1\leqslant k\leqslant r} \inf_{x\in I} |T'_k(x)|$, and $C = \max_{1\leqslant k\leqslant r} \mathrm{Lip}(\log |T'_k|)$. Then*

$$\frac{h(x)}{h(y)} \leqslant \mathrm{e}^{C|x-y|/(\lambda-1)} \qquad \textit{for all} \quad x \in I. \tag{42}$$

Proof. We follow the proof of theorem 5.2.1 in [19]. As in the proof and notation of lemma 5.2.2 [19], we have that

$$\log \frac{|(T_{k_{N-1}} \circ \cdots \circ T_{k_0})'(x)|}{|(T_{k_{N-1}} \circ \cdots \circ T_{k_0})'(y)|} \leqslant \frac{C}{\lambda - 1} |T_{k_{N-1}} \circ \cdots \circ T_{k_0}(x) - T_{k_{N-1}} \circ \cdots \circ T_{k_0}(y)|, \tag{43}$$

provided $|T_{k_i} \circ \cdots \circ T_{k_0}(x) - T_{k_i} \circ \cdots \circ T_{k_0}(y)| < \epsilon_0$ for $i = 0, \ldots, N-1$, where $\epsilon_0 > 0$ is such that $|x - y| < \epsilon_0 \Rightarrow |T_k x - T_k y| \geqslant \lambda|x - y|$ for all $x, y \in I$ and $k = 1, \ldots, r$. Let $\phi_0 \equiv 1$ be an initial density that is to be pushed forward, and denote by $\phi^i_{k_{N-1},\ldots,k_0}$ the push forward of ϕ_0 under $T_{k_{N-1}} \circ \cdots \circ T_{k_0}$ along one of the inverse branches of $T_{k_{N-1}} \circ \cdots \circ T_{k_0}$. By (43), we have that

$$\phi^i_{k_{N-1},\ldots,k_0}(x)/\phi^i_{k_{N-1},\ldots,k_0}(y) \leqslant \mathrm{e}^{C|x-y|/(\lambda-1)}.$$

For a fixed sequence $k_{N-1}, \ldots, k_0$, we may sum over i to obtain

$$\phi_{k_{N-1},\ldots,k_0}(x)/\phi_{k_{N-1},\ldots,k_0}(y) \leqslant \mathrm{e}^{C|x-y|/(\lambda-1)}.$$

We may now combine the contributions from each of the sequences $k_{N-1}, \ldots, k_0$ to obtain

$$\frac{\sum^r_{k_0,\ldots,k_{N-1}=1} w_{k_{N-1}} \ldots w_{k_0} \phi_{k_{N-1},\ldots,k_0(x)}}{\sum^r_{k_0,\ldots,k_{N-1}=1} w_{k_{N-1}} \ldots w_{k_0} \phi_{k_{N-1},\ldots,k_0(y)}} := \frac{\phi_N(x)}{\phi_N(y)} \leqslant \mathrm{e}^{C|x-y|/(\lambda-1)}.$$

Note that $\phi_N(x) = \mathcal{P}^N \phi_0(x)$, so that we have a bound on the distortion of the uniform density after being pushed forward N times under the (random) Perron–Frobenius operator. Clearly, there exist constants such that $A \leqslant \phi_N(x) \leqslant B$ for all $x \in I$ and $N \geqslant 0$. Letting ϕ be a limit of the sequence $\frac{1}{N}\sum_{i=0}^{N-1}\phi_N$, as $N \to \infty$, we see that ϕ is fixed by $\mathcal{P}$ and is bounded above and below by A and B respectively. Furthermore, $\phi(x)/\phi(y) \leqslant \mathrm{e}^{C|x-y|/(\lambda-1)}$. By uniqueness, $\phi = h$. □

6.4.2. Markov case

Lemma 6.4. *Suppose that each T_k is an expanding $C^{1+\mathrm{Lip}}$ map of the circle. Then*

$$\frac{h_k(x)}{h_k(y)} \leqslant \mathrm{e}^{C|x-y|/(\lambda-1)} \qquad \text{for all} \quad x \in I \text{ and each } k = 1, \ldots, r. \tag{44}$$

Proof. Follows as for lemma 6.3 with appropriate modifications. □

6.5. Bounding $\|\tilde{P}_n - P_n\|_m$ and $\|\tilde{S}_n - S_n\|_m$

6.5.1. iid case.

Lemma 6.5. *Let $\tilde{P}_n(k) = \mu(I_i \cap T_k^{-1}I_j)/\mu(I_i)$, and $\tilde{P}_n = \sum_{k=1}^{r} w_k \tilde{P}_n(k)$. Under the assumptions of theorem 3.3:*

(i) $\|\tilde{P}_n - P_n\|_m \leqslant (\mathrm{Lip}\, h/\inf_{x\in I} h)/n$, if each T_k is a general Lasota–Yorke map, and the partition $\{I_1, \ldots, I_n\}$ contains all points of non-Lipschitzness of h.
(ii) $\|\tilde{P}_n - P_n\|_m \leqslant \mathrm{e}^{C/(1-\lambda)n} - 1$, if each T_k is a $C^{1+\mathrm{Lip}}$ map of the circle.

Proof. We treat case (ii) first. Fix k. As in the proof of lemma 3.6 [3], one has

$$|\tilde{P}_{n,ij}(k) - P_{n,ij}(k)| \leqslant P_{n,ij}(k)|1 - (\sup_{x\in I_i} h(x)/\inf_{x\in I_i} h(x))|.$$

Using (42), we have

$$\begin{aligned}\|\tilde{P}_n(k) - P_n(k)\|_m &\leqslant \max_{1\leqslant i\leqslant n}\sum_{j=1}^{n} P_{n,ij}(k)|1 - (\sup_{x\in I_i} h(x)/\inf_{x\in I_i} h(x))| \\ &\leqslant \max_{1\leqslant i\leqslant n}\sum_{j=1}^{n} P_{n,ij}(k)(\mathrm{e}^{C/(\lambda-1)n} - 1) \\ &= \mathrm{e}^{C/(\lambda-1)n} - 1.\end{aligned}$$

The result of (ii) now follows easily. As for case (i), we note that

$$|\tilde{P}_{n,ij}(k) - P_{n,ij}(k)| \leqslant P_{n,ij}(k)|1 - (\sup_{x\in \mathrm{Int}(I_i)} h(x)/\inf_{x\in \mathrm{Int}(I_i)} h(x))|,$$

again as in the proof of lemma 3.6 [3]. Since h is Lipschitz on the interiors of the partition sets by assumption, proceeding as above, we have the bound

$$\|\tilde{P}_n(k) - P_n(k)\|_m \leqslant (\mathrm{Lip}(h)/\inf_{x\in I} h(x))/n,$$

where $\mathrm{Lip}(h)$ is understood to mean the maximum Lipschitz constant calculated over each of the Lipschitz pieces of h separately. The result of (ii) now follows as above. □

6.5.2. Markov case.

Lemma 6.6. *Let S_n be as in (13) and $\tilde{S}_n$ as in (39). Under the assumptions of theorem 4.3:*

(i) $\|\tilde{S}_n - S_n\|_m \leqslant \max_{1\leqslant k\leqslant r}((\sum_{l=1}^r W_{lk}^*)(\mathrm{Lip}\, h_k/\inf_{x\in I} h_k))/n$, *if each T_k is a general Lasota–Yorke map, and the partition $\{I_1, \ldots, I_n\}$ contains all points of non-Lipschitzness of every T_k, $k = 1, \ldots, r$.*

(ii) $\|\tilde{S}_n - S_n\|_m \leqslant (\max_{1\leqslant k\leqslant r} \sum_{l=1}^r W_{lk}^*)(\mathrm{e}^{C/(1-\lambda)n} - 1)$, *if each T_k is a $C^{1+\mathrm{Lip}}$ map of the circle.*

Proof. We treat case (ii) first. Let $\tilde{P}_n(k)$ be defined as in (38) and $P_n(k)$ as in (12). Proceeding as in the proof above, taking the distortion bounds from (44),

$$\|\tilde{P}_n(k) - P_n(k)\|_m \leqslant \max_{1\leqslant i\leqslant n} \sum_{j=1}^n P_{n,ij}(k)|1 - (\sup_{x\in I_i} h_k(x)/\inf_{x\in I_i} h_k(x))| \leqslant \mathrm{e}^{C/(\lambda-1)n} - 1,$$

and noting that

$$\|\tilde{S}_n - S_n\|_m = \max_{1\leqslant k\leqslant r} \sum_{l=1}^r W_{lk}^* \|\tilde{P}_n(k) - P_n(k)\|_m \leqslant \left(\max_{1\leqslant k\leqslant r} \sum_{l=1}^r W_{lk}^*\right)(\mathrm{e}^{C/(\lambda-1)n} - 1),$$

we are done. As for case (i), following the proof of lemma 6.5 we have

$$\|\tilde{P}_n(k) - P_n(k)\|_m \leqslant (\mathrm{Lip}(h_k)/\inf_{x\in I} h_k(x))/n,$$

where $\mathrm{Lip}(h_k)$ is understood to be the maximum Lipschitz constant calculated over each of the Lipschitz pieces of h_k separately, and the result of (i) follows. □

6.6. Bounding $\|Z_n\|_m$ and $\|\hat{Z}_n\|_m$

6.6.1. iid case. We write $Z_n = I_n + \sum_{N=1}^\infty P_n^N - P_n^\infty$, and proceed to bound $\|P_n^N - P_n^\infty\|_m$. *Converting norms.* Define the projection $\Pi_h : BV \to \mathrm{sp}\{h\}$ by $\Pi_h(f) = (\int_I f\, \mathrm{d}m)h$. In this section, we show that the contraction property of $\mathcal{P}$ in the BV norm gives bounds for the rate of convergence of iterates of the Ulam matrix P_n to its eigenvector.

We will identify elements of F_n (as defined in the proof of theorem 3.3 (i)) with the n-tuples $(f_{n,1}, \ldots, f_{n,n})$ that uniquely define the function f_n. Our norm on F_n will be $\|f_n\|_m = \sum_{i=1}^n |f_{n,i}|$, where the $f_{n,i}$ define f_n; if we think of F_n as being isomorphic to $\mathbb{R}^n$, this norm is simply the standard L^1 vector norm. We now show that when restricted to F_n, the norms $\|\cdot\|$ and $\|\cdot\|_m$ are equivalent.

Lemma 6.7. *For $f_n \in F_n$, $\|f_n\|_m \leqslant n\|f_n\|$ and $\|f_n\| \leqslant 2\|f_n\|_m$.*

Proof. First, we note that $\|f_n\|_m = n\|f_n\|_1$. So

$$\|f_n\|_m = n\|f_n\|_1 \leqslant \max\{n\,\mathrm{var}\, f_n, n\|f_n\|_1\} = n\|f_n\|. \tag{45}$$

Second, note that $\mathrm{var}\, f_n \leqslant 2\|f_n\|_m$. So

$$\|f_n\| = \max\{\mathrm{var}\, f_n, \|f_n\|_1\} \leqslant \max\{2\|f_n\|_m, (1/n)\|f_n\|_m\} \leqslant 2\|f_n\|_m. \tag{46}$$

□

Lemma 6.8. $\|P_n^N - P_n^\infty\|_m \leqslant 4n\|\mathcal{P}|_{BV_0}\|^N$.

Proof. Using the notation introduced in the proof of theorem 3.3 (i), recall that $[\pi_n \mathcal{P}_k] = P_n(k)$ and so $[\pi_n \mathcal{P}] = P_n$.

Denote $F_{n,0} = \{f_n \in F_n : \sum_{i=1}^n f_{n,i} = 0\}$. We begin by relating $\|P_n^N - P_n^\infty\|_m$ and $\|P_n^N|_{F_{n,0}}\|_m$. In what follows, we simultaneously consider f_n as a step function, and as the n-tuple $[f_{n,1}, \ldots, f_{n,n}]$; in the latter case the action of matrices is understood to be left multiplication.

$$\begin{aligned}\|P_n^N - P_n^\infty\|_m &= \sup_{f_n \in F_n} \frac{\|(P_n^N - P_n^\infty) f_n\|_m}{\|f_n\|_m} \\ &= \sup_{f_n \in F_n} \frac{\|P_n^N (f_n - P_n^\infty f_n)\|_m}{\|f_n\|_m} \\ &\leqslant \|P_n^N|_{F_{n,0}}\|_m \sup_{f_n \in F_n} \frac{\|f_n - P_n^\infty f_n\|_m}{\|f_n\|_m} \\ &\leqslant \|P_n^N|_{F_{n,0}}\|_m \sup_{f_n \in F_n} \frac{2\|f_n\|_m}{\|f_n\|_m} \\ &= 2\|P_n^N|_{F_{n,0}}\|_m.\end{aligned}$$

Now we link this result with the bounds that we have for the Perron–Frobenius operator:

$$\begin{aligned}2\|P_n^N|_{F_{n,0}}\|_m &= 2 \sup_{f_{n,0} \in F_{n,0}} \frac{\|[\pi_n \mathcal{P}]^N f_{n,0}\|_m}{\|f_{n,0}\|_m} \\ &\leqslant 2 \sup_{f_{n,0} \in F_{n,0}} \frac{n\|(\pi_n \mathcal{P})^N f_{n,0}\|}{\|f_{n,0}\|/2} \\ &\leqslant 4n\|(\pi_n \mathcal{P})^N|_{F_{n,0}}\| \\ &\leqslant 4n\|\pi_n \mathcal{P}|_{F_{n,0}}\|^N \\ &\leqslant 4n\|\mathcal{P}|_{F_{n,0}}\|^N \\ &\leqslant 4n\|\mathcal{P}|_{BV_0}\|^N.\end{aligned}$$ □

Corollary 6.9. *Under the hypotheses of theorem 3.3 (ii) or (iii),* $\|P_n^N - P_n^\infty\|_m \leqslant 4n\gamma^N$, *where* $\gamma = \alpha + \eta + \beta/2$ *or* $\gamma = \alpha + \beta/2$, *respectively.*

A bound for $\|Z_n\|_m$. The following results are modified versions of the arguments of section 6 [3].

Lemma 6.10. *Under the hypotheses of theorem 3.3 (ii) or (iii), setting* $\gamma = \alpha + \eta + \beta/2$ *or* $\gamma = \alpha + \beta/2$, *respectively,*

$$\|Z_n\|_m \leqslant \inf_{0<\delta<1} \left\{ \left(2 + \frac{\delta}{1-\delta}\right) \left(\left[\frac{\log(4n/\delta)}{-\log \gamma} \right] + 1 \right) - 1 \right\}, \tag{47}$$

where $[\cdot]$ *denotes the integer part.*

Before proving this, we state a result from [10]†.

Lemma 6.11 (theorem 16.2.4 [8]). *Suppose that* P_n *is an* $n \times n$ *irreducible, aperiodic stochastic matrix with fixed left eigenvector* p_n. *As before, define* $P_{n,ij}^\infty = p_{n,j}$. *Select a number* $0 < \delta < 1$ *and let* m_n *be such that*

$$P_{n,ij}^{m_n} \geqslant (1-\delta) p_{n,j} \qquad \textit{for all} \quad 1 \leqslant i, j \leqslant n. \tag{48}$$

† Theorem 16.2.4 [10] is incorrect as stated. Equation (16.21) should read $\|P^n(x,\cdot) - \pi(\cdot)\| \leqslant 2$ for $n < m$, and $\|P^n(x,\cdot) - \pi(\cdot)\| \leqslant \rho^{[n/m]}$ for $n \geqslant m$. The appropriate modifications should be made in lemma 6.6 [3].

Then

$$\|P_n^N - P_n^\infty\|_m \leqslant \begin{cases} 2, & \text{if } N < m_n, \\ \delta^{[N/m_n]}, & \text{if } N \geqslant m_n. \end{cases}$$

Proof of lemma 6.10. We now find an appropriate m_n to satisfy (48) for P_n. A sufficient condition for (48) to be satisfied is that $|P_{n,ij}^{m_n} - p_{n,j}| \leqslant \delta p_{n,j}$ for all $1 \leqslant i, j \leqslant n$. Summing over j and maximising over i gives $\|P_n^{m_n} - P_n^\infty\|_m \leqslant \delta \Rightarrow$ (48) holds. From corollary 6.9, we see that provided $4n\gamma^{m_n} \leqslant \delta$, (48) will hold. A simple rearrangement shows that $4n\gamma^{m_n} \leqslant \delta$ if

$$m_n \geqslant [\log(4n/\delta)/(-\log\gamma)] + 1. \tag{49}$$

So now

$$\|Z_n\|_m = \left\| I + \sum_{N=1}^{\infty}(P_n^N - P_n^\infty)\right\|_m \leqslant 1 + \sum_{N=1}^{m_n-1} 2 + \sum_{N=m_n}^{\infty} \delta^{[N/m_n]} = \left(2 + \frac{\delta}{1-\delta}\right)m_n - 1. \tag{50}$$

Inserting (49) into (50), we obtain the result. □

6.6.2. Markov case.

Converting norms. We wish to study the rate of convergence of S_n^N to the limiting matrix S_n^∞ defined in (40) as $N \to \infty$, in terms of the $\|\cdot\|_m$ norm. At the moment, we have information regarding the convergence of $\hat{\mathcal{P}}^N|_{B\hat{V}_0}$ to $\hat{0}$ from lemma 6.2, and in this section, we link these two types of convergence.

Lemma 6.12. $\|S_n^N - S_n^\infty\|_m \leqslant 4rn\|\hat{\mathcal{P}}|_{B\hat{V}_0}\|^N$.

Proof. Let $\hat{F}$ and $\hat{\pi}_n$ be as in the proof of theorem 4.3 (i). Define $\hat{F}_{n,0} = \{(f_n^{(1)}, \ldots, f_n^{(r)}) \in \hat{F}_n : \sum_{i=1}^n f_{n,i}^{(k)} = 0 \text{ for all } k = 1, \ldots, r\}$.

Following the initial chain of inequalities in the proof of lemma 6.8, one has

$$\|S_n^N - S_n^\infty\|_m \leqslant 2\|S_n^N|_{\hat{F}_{n,0}}\|_m.$$

We define an intermediate vector norm $\|\cdot\|_{m'}$ as $\|s_n\|_{m'} = \max_{1\leqslant k\leqslant r}\|s_n^{(k)}\|_m$. Clearly, $\|\cdot\|_{m'} \leqslant \|\cdot\|_m \leqslant r\|\cdot\|_{m'}$.

Lemma 6.13. *Simultaneously thinking of the nr-tuple*

$$\hat{f}_n = (f_{n,1}^{(1)}, \ldots, f_{n,n}^{(1)}, f_{n,1}^{(2)}, \ldots, f_{n,n}^{(2)}, \ldots\ldots, f_{n,1}^{(r)}, \ldots, f_{n,n}^{(r)})$$

as representing a $1 \times nr$ vector and an element of $B\hat{V}$, we have the relations $\|\hat{f}_n\|_{m'} \leqslant n\|\hat{f}_n\|$ and $\|\hat{f}_n\| \leqslant 2\|\hat{f}_n\|_{m'}$.

Proof. Noting that $\|\hat{f}_n\|_{m'} = \max_{1\leqslant k\leqslant r}\|f_n^{(k)}\|_m$ and that $\|\hat{f}_n\| = \max_{1\leqslant k\leqslant r}\|f_n^{(k)}\|$, the result follows directly from lemma 6.7. □

Recall that the matrix form of $\hat{\pi}_n\hat{\mathcal{P}}$ with respect to the basis $\prod_{k=1}^r\{\chi_{I_1}, \ldots, \chi_{I_n}\}$ is simply S_n:

$$2\|S_n^N|_{\hat{F}_{n,0}}\|_m = 2\sup_{\hat{f}_{n,0}\in\hat{F}_{n,0}} \frac{\|[\hat{\pi}_n\hat{\mathcal{P}}]^N\hat{f}_{n,0}\|_m}{\|\hat{f}_{n,0}\|_m} \leqslant 2\sup_{\hat{f}_{n,0}\in\hat{F}_{n,0}} \frac{r\|[\hat{\pi}_n\hat{\mathcal{P}}]^N\hat{f}_{n,0}\|_{m'}}{\|\hat{f}_{n,0}\|_{m'}}$$
$$\leqslant 2r\sup_{\hat{f}_{n,0}\in\hat{F}_{n,0}} \frac{n\|(\hat{\pi}_n\hat{\mathcal{P}})^N\hat{f}_{n,0}\|}{\|\hat{f}_{n,0}\|/2} \leqslant 4rn\|\hat{\mathcal{P}}|_{B\hat{V}_0}\|^N,$$

with the final inequality following as in the proof of lemma 6.8. □

Corollary 6.14. *Under the hypotheses of theorem 4.3 (ii) or (iii),* $\|S_n^N - S_n^\infty\|_m \leqslant 4rn\gamma^N$*, where* $\gamma = \max_{1\leqslant l\leqslant r}(\alpha'_l + \eta'_l) + \beta'/2$ *or* $\gamma = \alpha' + \beta'/2$*, respectively.*

A bound for $\|\hat{Z}_n\|_m$*.*

Lemma 6.15. *Under the hypotheses of theorem 4.3 (ii) or (iii), setting* $\gamma = \max_{1\leqslant l\leqslant r}(\alpha'_l + \eta'_l) + \beta'/2$ *or* $\gamma = \alpha' + \beta'/2$*, respectively,*

$$\|\hat{Z}_n\|_m \leqslant \inf_{0<\delta<1}\left\{\left(2+\frac{\delta}{1-\delta}\right)\left(\left[\frac{\log(4rn/\delta)}{-\log\gamma}\right]+1\right)-1\right\}. \tag{51}$$

Proof. Follows as in the proof of lemma 6.10. □

6.7. Taking care of the difference $\|h - \pi_n(h)\|_1$

6.7.1. iid case.

Lemma 6.16. *Under the assumptions of theorem 3.3:*

(i) $\|h - \pi_n(h)\|_1 \leqslant \mathrm{Lip}(h)/2n$*, if each* T_k *is a general Lasota–Yorke map,*
(ii) $\|h - \pi_n(h)\|_1 \leqslant (\mathrm{e}^{C/(\lambda-1)n} - 1)/2$*, if each* T_k *is a* $C^{1+\mathrm{Lip}}$ *map of the circle.*

Proof. First assume that each T_k is a $C^{1+\mathrm{Lip}}$ map of the circle, so that we may use the estimates of lemma 6.3:

$$\begin{aligned}\|h-\pi_n(h)\|_1 &= \sum_{i=1}^n \int_{I_i}\left|h - \frac{1}{n}\int_{I_i} h\,\mathrm{d}m\right|\mathrm{d}m \leqslant \sum_{i=1}^n \frac{1}{n}(\sup_{x\in I_i} h(x) - \inf_{x\in I_i} h(x))/2 \\ &= \frac{1}{n}\sum_{i=1}^n \inf_{x\in I_i} h(x)\cdot(\sup_{x\in I_i} h(x)/\inf_{x\in I_i} h(x) - 1)/2 \leqslant (\mathrm{e}^{C/(\lambda-1)n} - 1)/2,\end{aligned} \tag{52}$$

as required; the inequality (52) follows as in the proof of lemma 4 [5], for example.

In the case of general Lasota–Yorke maps, the result follows by noting that (52) $\leqslant \sum_{i=1}^n \frac{1}{n}(\mathrm{Lip}(h)/n)/2$. □

6.7.2. Markov case.

Lemma 6.17. *Under the assumptions of theorem 4.3:*

(i) $\|h - \pi_n(h)\|_1 \leqslant \sum_{k=1}^r w_k\,\mathrm{Lip}(h_k)/2n$*, if each* T_k *is a general Lasota–Yorke map,*
(ii) $\|h - \pi_n(h)\|_1 \leqslant (\mathrm{e}^{C/(\lambda-1)n} - 1)/2$*, if each* T_k *is a* $C^{1+\mathrm{Lip}}$ *map of the circle.*

Proof.

$$\|h-\pi_n(h)\|_1 = \left|\sum_{k=1}^r w_k h_k - \pi_n\left(\sum_{k=1}^r w_k h_k\right)\right|_1 \leqslant \sum_{k=1}^r w_k\|h_k - \pi_n h_k\|_1,$$

and the results now follow as in the proof of lemma 6.16. □

6.8. Proof of parts (ii) and (iii) of theorems 3.3 and 4.3

6.8.1. iid case.

Proof of theorem 3.3 (ii) and (iii). This follows immediately from (37), and lemmas 6.5, 6.10 and 6.16. □

6.8.2. Markov case.

Proof of theorem 4.3 (ii) and (iii). The results follow immediately from (40) and (41) and lemmas 6.6, 6.15 and 6.17. □

6.9. General Markov proofs

Definition 6.18. *Define an operator* $\hat{\mathcal{D}}^* : C(S_0 \times I, \mathbb{R}) \circlearrowleft$ *by*

$$(\hat{\mathcal{D}}^* g)(\omega_0, x) = \sum_{\omega_1=1}^{r} g(\omega_1, T_{\omega_0} x) W_{\omega_0 \omega_1}. \tag{53}$$

We call a probability measure $\xi \in \mathcal{M}(S_0 \times I)$ $\hat{\mathcal{D}}$-invariant *if*

$$\int_{S_0 \times I} g(\omega_0, x)\, \mathrm{d}\xi(\omega_0, x) = \int_{S_0 \times I} (\hat{\mathcal{D}}^* g)(\omega_0, x)\, \mathrm{d}\xi(\omega_0, x) \qquad \textit{for all} \quad g \in C(S_0 \times I, \mathbb{R}). \tag{54}$$

The following lemma characterizes τ-invariant measures on $\Omega \times I$ in terms of $\hat{\mathcal{D}}$-invariant measures on the simpler space $S_0 \times I$.

Lemma 6.19. *Let* $A \in \mathfrak{B}(S_0 \times I)$ *(the space of Borel measurable sets on* $S_0 \times I$*) and* $B \in \mathfrak{B}(\Omega \times I)$*. Define the sections* $A_{\omega_0} = \{x \in I : (\omega_0, x) \in A\}$ *and* $B_\omega = \{x \in I : (\omega, x) \in B\}$*. Let* $\{\mu_{\omega_0}\}_{\omega_0=1}^r$ *be a collection of Borel probability measures on* I*. Define a probability measure* $\xi \in \mathcal{M}(S_0 \times I)$ *by*

$$\xi(A) = \int_{S_0} \mu_{\omega_0}(A_{\omega_0})\, \mathrm{d}\rho(\omega_0), \tag{55}$$

and a probability measure $\tilde{\mu} \in \mathcal{M}(\Omega \times I)$ *by*

$$\tilde{\mu}(B) = \int_{\Omega} \mu_{\omega_0}(B_\omega)\, \mathrm{d}\mathbb{P}(\omega). \tag{56}$$

Then ξ *is* $\hat{\mathcal{D}}$*-invariant iff* $\tilde{\mu}$ *is* τ*-invariant.*

Proof. Let $g : \Omega \times I \to \mathbb{R}$ be any continuous function and define

$$\hat{g}(\omega_0, x) = \left(\int_{[\omega_0]} g(\omega, x)\, \mathrm{d}\mathbb{P}(\omega) \right) / \mathbb{P}([\omega_0]),$$

where $[a_0] = \{\omega \in \Omega : \omega_0 = a_0\}$.

$$\begin{aligned}
&\int_{\Omega \times I} g(\tau((\omega_0 \omega_1 \omega_2 \cdots), x))\, \mathrm{d}\tilde{\mu}(\omega, x) \\
&\quad = \int_{\Omega} \int_I g((\omega_1 \omega_2 \ldots), T_{\omega_0} x)\, \mathrm{d}\mu_{\omega_0}(x)\, \mathrm{d}\mathbb{P}(\omega) \\
&\quad = \sum_{\omega_0=1}^{r} \sum_{\omega_1=1}^{r} \int_I \left(\frac{1}{\mathbb{P}([\omega_0 \omega_1])} \int_{[\omega_0 \omega_1]} g((\omega_1 \omega_2 \cdots), T_{\omega_0} x)\, \mathrm{d}\mathbb{P}(\omega) \right) \mathrm{d}\mu_{\omega_0}\, \mathbb{P}([\omega_0 \omega_1]) \\
&\quad = \sum_{\omega_0=1}^{r} \sum_{\omega_1=1}^{r} \int_I \hat{g}(\omega_1, T_{\omega_0} x)\, \mathrm{d}\mu_{\omega_0}(x)\, W_{\omega_0 \omega_1} w_{\omega_0} \\
&\quad = \int_{S_0 \times I} (\hat{\mathcal{D}}^* \hat{g})(\omega_0, x)\, \mathrm{d}\xi(\omega_0, x).
\end{aligned}$$

Since the first expression equals $\int_{\Omega\times I} g(\omega, x)\,\mathrm{d}\tilde{\mu}(\omega, x)$ iff $\tilde{\mu}$ is τ-invariant, and the final expression equals $\int_{S_0\times I} \hat{g}\,\mathrm{d}\xi(\omega_0, x)$ iff ξ is $\hat{\mathcal{D}}$-invariant, we are done. □

Proof of lemma 4.1. We show that the measure ξ in (55) is $\hat{\mathcal{D}}$-invariant iff the family $\{\mu_{\omega_0}\}_{\omega_0=1}^r$ is fixed under the transformation (9). The result will then follow from lemma 6.19.

Suppose that ξ is $\hat{\mathcal{D}}$-invariant, and choose $g(\omega_0, x) = \chi_{\{j\}\times A}(\omega_0, x)$ for some $j \in S_1$ and $A \in \mathfrak{B}(I)$. On one hand, we have:

$$\int_{S_1\times I} \chi_{\{j\}\times A}(\omega_1, x)\,\mathrm{d}\xi(\omega_1, x) = w_j\mu_j(A). \tag{57}$$

On the other, we have:

$$\begin{aligned}\int_{S_0\times I} \hat{\mathcal{D}}^*\chi_{\{j\}\times A}(\omega_0, x)\,\mathrm{d}\xi(\omega_0, x) &= \sum_{\omega_0=1}^{r}\int_I \sum_{\omega_1=1}^{r} \chi_{\{j\}\times A}(\omega_1, T_{\omega_0}x)W_{\omega_0\omega_1}\,\mathrm{d}\mu_{\omega_0}(x)\, w_{\omega_0} \\ &= \sum_{\omega_0=1}^{r} \mu_{\omega_0}(T_{\omega_0}^{-1}A)\ W_{\omega_0 j}w_{\omega_0}.\end{aligned} \tag{58}$$

By a standard approximation argument, the property (54) also holds for simple functions such as $\chi_{\{j\}\times A}$, so that we may equate (57) and (58), proving that $\hat{\mathcal{D}}^*$-invariance implies that the family $\{\mu_{\omega_0}\}_{\omega_0=1}^r$ is fixed by the transformation (9).

For the converse, suppose that $\{\mu_{\omega_0}\}_{\omega_0=1}^r$ are fixed by (9), and choose $g \in C(S_0 \times I, \mathbb{R})$. It is a simple matter to verify directly that

$$\int_{S_0\times I} (\hat{\mathcal{D}}^* g)(\omega_0, x)\,\mathrm{d}\xi(\omega_0, x) = \int_{S_0\times I} g(\omega_0, x)\,\mathrm{d}\mu_{\omega_0}\,\mathrm{d}\rho(\omega_0).$$

Thus $\{\mu_{\omega_0}\}_{\omega_0=1}^r$ being fixed by (9) implies $\hat{\mathcal{D}}$-invariance of ξ.

The fact that $\mu = \sum_{k=1}^r w_k\mu_k$ is invariant follows immediately from definition 2.1. □

Acknowledgments

The author is grateful to an anonymous referee and to Rua Murray; both for pointing out an error in an earlier draft. This research was supported by a Japan Society for the Promotion of Science Postdoctoral Fellowship (Research grant from the Ministry of Education, Science, and Culture No 09-97004), and carried out at the Department of Mathematical Engineering and Information Physics, The University of Tokyo.

References

[1] Boyarsky A and Lou Y S 1992 Existence of absolutely continuous invariant measures for higher dimensional random maps *Dyn. Stab. Syst.* **7** 233–44

[2] Ding J and Zhou A 1996 Finite approximations of Frobenius–Perron operators. A solution of Ulam's conjecture to multi-dimensional transformations *Physica* D **92** 61–8

[3] Froyland G 1998 Approximating physical invariant measures of mixing dynamical systems in higher dimensions *Nonlinear Anal. Theory Methods Appl.* **32** 831–60

[4] Gora P and Boyarsky A 1989 Absolutely continuous invariant measures for piecewise expanding C^2 transformations in R^N *Israel J. Math.* **67** 272–86

[5] Hunt B R 1996 Estimating invariant measures and Lyapunov exponents *Ergod. Theor. Dynam. Syst.* **16** 735–49

[6] Keane M, Murray R and Young L-S 1998 Computing invariant measures for expanding circle maps *Nonlinearity* **11** 27–46

[7] Kowalski Z S 1993 Ergodic properties of skew products with Lasota–Yorke type maps in the base *Stud. Math.* **106** 45–57

[8] Lasota A and Yorke J A 1973 On the existence of invariant measures for piecewise monotonic transformations *Trans. Am. Math. Soc.* **186** 481–8

[9] Li T-Y 1976 Finite approximation for the Frobenius–Perron operator. A solution to Ulam's conjecture *J. Approx. Theory* **17** 177–86

[10] Meyn S P and Tweedie R L 1993 Markov chains and stochastic stability *Communications and Control Engineering* (London: Springer)

[11] Morita T 1985 Asymptotic behavior of one-dimensional random dynamical systems *J. Math. Soc. Japan* **37** 651–63

[12] Morita T 1985 Random iteration of one-dimensional transformations *Osaka J. Math.* **22** 489–518

[13] Morita T 1988 Deterministic version lemmas in ergodic theory of random dynamical systems *Hiroshima Math. J.* **18** 15–29

[14] Murray R 1997 Discrete approximation of invariant densities *PhD Thesis* University of Cambridge

[15] Ohno T 1983 Asymptotic behaviors of dynamical systems with random parameters *Publ. Res. Inst. Math. Sci. (Kyoto University)* **19** 83–98

[16] Pelikan S 1984 Invariant densities for random maps of the interval *Trans. Am. Math. Soc.* **281** 813–25

[17] Schweitzer P J 1968 Perturbation theory and finite Markov chains *J. Appl. Probab.* **5** 401–4

[18] Ulam S M 1964 *Problems in Modern Mathematics* (New York: Interscience)

[19] Young L-S 1995 Ergodic theory of differentiable dynamical systems *Real and Complex Dynamical Systems* ed B Branner and P Hjorth (Dordrecht: Kluwer) pp 293–336