STABILITY AND APPROXIMATION OF RANDOM INVARIANT MEASURES OF MARKOV CHAINS IN RANDOM ENVIRONMENTS

GARY FROYLAND AND CECILIA GONZÁLEZ-TOKMAN

SCHOOL OF MATHEMATICS AND STATISTICS
UNIVERSITY OF NEW SOUTH WALES
SYDNEY NSW 2052, AUSTRALIA

ABSTRACT. We consider finite-state Markov chains driven by a $\mathbb{P}$-stationary ergodic invertible process $\sigma : \Omega \to \Omega$, representing a random environment. For a given initial condition $\omega \in \Omega$, the driven Markov chain evolves according to $A(\omega)A(\sigma\omega)\cdots A(\sigma^{n-1}\omega)$, where $A : \Omega \to M_d$ is a measurable $d \times d$ stochastic matrix-valued function. The driven Markov chain possesses $\mathbb{P}$-a.e. a measurable family of probability vectors $v(\omega)$ that satisfy the equivariance property $v(\sigma\omega) = v(\omega)A(\omega)$. Writing $v_{\omega} = \delta_{\omega} \times v(\omega)$, the probability measure $\nu(\cdot) = \int v_{\omega}(\cdot) \, d\mathbb{P}(\omega)$ on $\Omega \times \{1, \ldots, d\}$ is the corresponding random invariant measure for the Markov chain in a random environment.

Our main result is that $\nu$ is stable under a wide variety of perturbations of $\sigma$ and $A$. Stability is in the sense of convergence in probability of the random invariant measure of the perturbed system to the unperturbed random invariant measure $\nu$. Our proof approach is elementary and has no assumptions on the transition matrix function $A$ except measurability. We also develop a new numerical scheme to construct rigorous approximations of $\nu$ that converge in probability to $\nu$ as the resolution of the scheme increases. This new numerical approach is illustrated with examples of driven random walks and an example where the random environment is governed by a multidimensional nonlinear dynamical system.

1. Introduction

Let $(\Omega, \mathbb{P})$ be a probability space and $\sigma : \Omega \to \Omega$ an ergodic invertible map. A simple example is where $\Omega$ is a product of probability spaces, $\Omega = \prod_{i=-\infty}^{\infty} X$, $\sigma$ is the left shift $[\sigma(\omega)]_i = \omega_{i+1}$, and $\omega = (\ldots, \omega_{-1}, \omega_0, \omega_1, \ldots) \in \Omega$. If $\mathbb{P} = \prod_{i=-\infty}^{\infty} \xi$, where $\xi$ is a probability measure on $X$, then the dynamics of $\sigma$ generate an iid process on $X$ with distribution $\xi$. While this simple example illustrates the flexibility of this setup, for the purposes of this work, we make no assumptions on $(\Omega, \sigma, \mathbb{P})$ apart from ergodicity and invertibility, and consequently can also handle systems that are very far from iid.

Let $A : \Omega \to M_{d \times d}(\mathbb{R})$ take values in the space of stochastic matrices. We are interested in the linear cocycle generated by stochastic transition matrix function $A$ and driving dynamics $\sigma$. Given an initial probability vector $v(\omega) \in \mathbb{R}^d$, after $n$ time steps, this vector evolves to $v'(\sigma^n\omega) := v(\omega)A(\omega)\cdots A(\sigma^{n-1}\omega)$. If the vector-valued function $v$ satisfies the natural equivariance property $v(\sigma\omega) = v(\omega)A(\omega)$ for $\mathbb{P}$-a.a. $\omega \in \Omega$, we set $\nu(\cdot) = \int v_{\omega}(\cdot) \, d\mathbb{P}(\omega)$, where $v_{\omega} = \delta_{\omega} \times v(\omega)$ is supported on $\{\omega\} \times \{1, \ldots, d\}$, and call $\nu$ a random invariant measure for the random dynamical system $\mathcal{R} = (\Omega, \mathcal{F}, \mathbb{P}, \sigma, A, \mathbb{R}^d)$. Viewed as a skew-product, the map $\tau : \Omega \times \{1, \ldots, d\} \to \Omega$, defined by $\tau(\omega, v(\omega)) := (\sigma\omega, v(\omega)A(\omega))$, preserves $\nu$; that is, $\nu \circ \tau^{-1} = \nu$.
These transition matrix cocycles are also known in the literature as *Markov chains in random environments*, where the $A(\omega)$ govern a Markov chain at configuration $\omega$, and $\sigma$ is thought of as an ergodic stationary stochastic process (whose realisations give rise to the random environment). Markov chains in random environments have been considered in a series of papers by Nawrotzki [22], Cogburn [3, 4] and Orey [24]. The existence of a random invariant measure in our general setup is proven in [4] (Corollary 3.1). Central limit theorems [5] and large deviation results [31] have also been proven in this setting. Our focus in the present paper is on the stability of $\nu$ to various perturbations.

Applications of Markov chains in random environments arise in analyses of time-dependent dynamical systems, such as models of stirred fluids [12] and circulation models of the ocean and atmosphere [6, 1, 2]. In these settings, one can convert the typically low-dimensional nonlinear dynamics into infinite-dimensional linear dynamics by studying the dynamical action on functions on state space, (representing densities of invariant measures with respect to a suitable reference measure), rather than points in state space. The driven infinite-dimensional dynamics is governed by cocycles of Perron-Frobenius or transfer operators. In numerical computations, these operators are often estimated by large sparse stochastic matrices [12] that involve perturbations in the form of discretisations of both the phase space and the driving system, resulting in a finite-state Markov chain in a random environment. Thus, the stability and rigorous approximation of random invariant measures of Markov chains in random environments are important questions for applications in the physical and biological sciences.

Other application areas include multi-timescale systems of skew-product type (see eg. [26]), where the “random environment” is an aperiodic fast dynamics that drives the slow dynamics. In this setting the random invariant measures are a family of probability measures on the slow space, indexed by the fast coordinates. This collection of probability measures on the slow space represent time-asymptotic distributions of slow orbits that have been driven by particular sample trajectories of the fast system. In computations, both the fast and slow dynamics are approximated by discretised transfer operators, leading to a Markov chain in a random environment.

When $\Omega$ consists of a single point, one returns to the setting of a homogeneous Markov chain. The question of stability of the stationary distribution of homogeneous Markov chains under perturbations of the transition matrix has been considered by many authors [27, 21, 14, 13, 29, 30, 20]. These papers developed upper bounds on the norm of the resulting stationary distribution perturbation, depending on various functions of the unperturbed transition matrix, the unperturbed stationary distribution, and the perturbation. Our present focus is somewhat different: for Markov chains in random environments we seek to work with minimal assumptions on both the random environment and stochastic matrix function, and our primary concern is whether one can expect stability of the random invariant measures at all, and if so, in what sense. However, by enforcing stronger assumptions or requiring more knowledge about the driving process and the matrix functions, it may be possible to obtain bounds analogous to the homogeneous Markov chain setting.

In the dynamical systems context, linear random dynamical systems have a long history. When the matrices $A(\omega)$ are invertible, the celebrated multiplicative ergodic theorem (MET) of Oselechts [25] guarantees the $P$-a.e. existence of a measurable splitting of $\mathbb{R}^d$ into equivariant subspaces, within which vectors experience an identical asymptotic growth rate, known...
as a Lyapunov exponent. A recent extension [11] of the Oseledets theorem yields the same conclusion even when the matrices are not invertible. These equivariant subspaces, known as Oseledets spaces, generalise eigenspaces, which describe directions of distinct growth rates under linear dynamics from repeated application of a single matrix \( A \). In the present setting, the maximal growth rate is \( \log 1 = 0 \) and the equivariant \( v(\omega) \) belong to the corresponding Oseledets space.

In related work, Ochs [23] has linked convergence of Oseledets spaces to convergence of Lyapunov exponents in a class of random perturbations of general matrix cocycles. One of his standing hypotheses was that the matrices \( A(\omega) \) were invertible, which is not a natural condition in the setting of stochastic matrices. For products of stochastic matrices, the top Lyapunov exponent is always 0, thus [23] yields convergence of the random invariant measures in probability, provided the matrices are invertible. The type of perturbations that we investigate generalize Ochs’ “deterministic” perturbations in the context of stochastic matrices, which require \( \Omega \) to be a compact topological space and \( \sigma \) to be a homeomorphism. Moreover, the arguments of Ochs do not easily extend to the noninvertible matrix setting. Our approach also enables the construction of an efficient rigorous numerical method for approximating the random invariant measure.

Recent work on random compositions of piecewise smooth expanding interval maps [10] has shown stability of random absolutely continuous invariant measures (the physically relevant random invariant measures), under a variety of perturbations of the maps (and thus the Perron-Frobenius operators), while keeping the driving system unchanged. The results of [10] are obtained using ergodic-theoretical tools that heavily rely on the system and its perturbations sharing a common driving system, and thus do not extend directly to our current setting.

Using an elementary approach, we demonstrate stability of the random invariant measures in the sense of convergence in probability. We show that the random invariant measures are stable to the following types of perturbations:

1. **Perturbing the random environment** \( \sigma \): The base map \( \sigma \) is perturbed to a nearby map.
2. **Perturbing the transition matrix function** \( A \): The matrix function \( A \) is perturbed to a nearby matrix function.
3. **Stochastic perturbations**: The action of the random dynamical system is perturbed by convolving with a stochastic kernel close to the identity.
4. **Numerical schemes**: The random dynamical system is perturbed by a Galerkin-type approximation scheme inspired by Ulam’s method to numerically compute an estimate of the random invariant measure.

An outline of the paper is as follows. In Section 2 we state an abstract perturbation lemma that forms the basis of our results and check the hypotheses of this lemma in the stochastic matrix setting for the unperturbed Markov chain. In Section 3 we provide natural conditions under which the random invariant measure is unique. Section 4 proceeds through the four main types of perturbations listed above, deriving and confirming the necessary boundedness and convergence conditions. Numerical examples are given in Section 5.

## 2. A general perturbation lemma and random matrix cocycles

We begin with an abstract stability result for fixed points of linear operators.
Lemma 2.1. Let \((B, \| \cdot \|)\) be a separable normed linear space with continuous dual \((B^*, \| \cdot \|_*)\). Let \(L : (B^*, \| \cdot \|_*) \odot n (B, \| \cdot \|) \odot L_n : (B^*, \| \cdot \|_*) \odot n = 1, \ldots\) be linear maps satisfying

(a) there exists a bounded linear map \(L' : (B, \| \cdot \|) \odot \) such that \((L')^* = L\),
(b) for each \(n \in \mathbb{N}\) there is a \(v_n \in (B^*, \| \cdot \|_*)\) such that \(L_n v_n = v_n\), which we normalise so that \(\|v_n\|_* = 1\) (where \(\|v_n\|_* = \sup_{f \in B, \|f\| = 1} |v_n(f)|\)),
(c) for each \(n = 1, \ldots\), there exists a bounded linear map \(L'_n : B \odot \) such that \((L'_n)^* = L_n\), satisfying \((L' - L'_n)f \to 0\) as \(n \to \infty\) for each \(f \in B\).

Then there is a subsequence \(v_{n_j} \in B^*\) converging weak-* to some \(\tilde{v} \in B^*\); that is, \(v_{n_j}(f) \to \tilde{v}(f)\) as \(j \to \infty\). Moreover, \(L\tilde{v} = \tilde{v}\).

Proof. The existence of a weak-* convergent subsequence follows from (b) and the Banach-Alaoglu Theorem. To show that \(L\tilde{v} = \tilde{v}\), for \(f \in B\) we write

\[
(L\tilde{v} - \tilde{v})(f) = (L\tilde{v} - L v_n)(f) + (L v_n - L_n v_n)(f) + (v_n - \tilde{v})(f).
\]

Writing the first term of (1) as \(|(\tilde{v} - v_n)(L'f)|\) this term goes to zero by boundedness of \(L'\) and weak-* convergence of \(v_n\) to \(\tilde{v}\). Writing the second term as \(|v_n(L' - L'_n)(f)|\) and applying (c), we see this term vanishes as \(n \to \infty\). The third term goes to zero as \(n \to \infty\) by weak-* convergence of \(v_n\) to \(\tilde{v}\).

Remark 2.2. Condition (a) is equivalent to \(L\) bounded and weak-* continuous; that is, \(L v_i \to L v\) weak-* if \(v_i \to v\) weak-* where \(v, v_i \in B^*\) (see eg. Theorem 3.1.11 [19]).

2.1. Application to Random Matrix Cocycles. We now begin to define the objects \((B, \| \cdot \|)\) and its dual, and the operator \(L\) and its pre-dual \(L'\) in the random matrix cocycle setting. Let \(\mathcal{V}\) denote the Banach space of \(d\)-dimensional bounded measurable vector fields \(v : \Omega \to \mathbb{R}^d\), with norm \(\|v\|_* = \left| \sum_{i=1}^d |v_i| \right|_{L^\infty(\Omega)} = \|v(\cdot)\|_{L^\infty(\Omega)}\). Associated with \(\mathcal{V}\) is the Banach space \(\mathcal{F}\) of \(d\)-dimensional integrable functions \(f : \Omega \to \mathbb{R}^d\), with norm \(\|f\| = \max_{1 \leq i \leq d} |f_i|_{L^1(\Omega)} = |f|_{L^1(\Omega)}\).

Lemma 2.3. \((\mathcal{V}, \| \cdot \|_*)\) and \((\mathcal{F}, \| \cdot \|)\) are Banach spaces and \((\mathcal{V}, \| \cdot \|_*) = (\mathcal{F}, \| \cdot \|)^*\).

Proof. Given Banach spaces \(X_i\), one can identify \((X_1 \oplus \cdots \oplus X_d)^*\) with \((X_1^* \oplus \cdots \oplus X_d^*)\), and with the right identification \(x^*(y) = \sum_{i=1}^d x_i^*(y_i)\), [19, Theorem 1.10.13].

For the norms, note that using the usual formula for \(\| \cdot \|_*\) in terms of \(\| \cdot \|\) one has

\[
\|v\|_* = \sup_{\|f\| = 1} \left| \int v(\omega) \cdot f(\omega) \, d\mathbb{P}(\omega) \right| = \sup_{f : \|f\|_{L^1(\mathbb{P})} \leq 1} \left| \int \sum_{i=1}^d v_i(\omega) f_i(\omega) \, d\mathbb{P}(\omega) \right| \\
\leq \sup_{f : \|f\|_{L^1(\mathbb{P})} \leq 1} \int \sum_{i=1}^d |v_i(\omega)| \max_{1 \leq i \leq d} |f_i(\omega)| \, d\mathbb{P}(\omega) \\
\leq \sum_{i=1}^d |v_i(\omega)| \max_{1 \leq i \leq d} |f_i|_{L^1(\mathbb{P})} = \sum_{i=1}^d |v_i|_{L^1(\mathbb{P})}.
\]
The reverse inequality may be obtained as follows. For each \( j \in \mathbb{N} \), let \( f_{j,i} = \frac{\text{sgn}(v_i(\omega))}{P(\Omega_j)} \mathbbm{1}_{\Omega_j} \),
where \( \Omega_j = \left\{ \omega \in \Omega : \sum_{i=1}^{d} |v_i(\omega)| \geq (1 - 1/j) \left| \sum_{i=1}^{d} |v_i| \right|_{L^\infty(P)} \right\} \). Then, \( \|f_j\| = 1 \) and
\[
\left| \int v(\omega) \cdot f_j(\omega) \ d\mathbb{P}(\omega) \right| = \int \sum_{i=1}^{d} v_i(\omega) f_{j,i}(\omega) \ d\mathbb{P}(\omega)
= \frac{1}{P(\Omega_j)} \int_{\Omega_j} \sum_{i=1}^{d} |v_i(\omega)| d\mathbb{P}(\omega) \geq (1 - 1/j) \left| \sum_{i=1}^{d} |v_i| \right|_{L^\infty(P)}.
\]
Letting \( j \to \infty \), we get that \( \|v\|_* \geq \left| \sum_{i=1}^{d} |v_i| \right|_{L^\infty(P)} \).

Let \( \mathcal{L} : \mathcal{V} \to \mathcal{V} \) be the operator acting as \( (\mathcal{L}v)(\omega) = v(\sigma^{-1}\omega)A(\sigma^{-1}\omega) \), where multiplication on the left of the row-stochastic matrix \( A(\sigma^{-1}\omega) \) is meant.

**Lemma 2.4.** The operator \( \mathcal{L} \) has \( \| \cdot \|_* \)-norm 1.

**Proof.** One has
\[
\|\mathcal{L}v\|_* = \left| \sum_{i=1}^{d} |(\mathcal{L}v)_i| \right|_{\infty} = \left| \sum_{i=1}^{d} |v \circ \sigma^{-1} \cdot A \circ \sigma^{-1}|_i \right|_{\infty} \leq \left| \sum_{i=1}^{d} |v \circ \sigma^{-1}||_i \right|_{\infty} = \|v\|_*.
\]
The inequality holds as \( A(\omega) \) is row-stochastic for each \( \omega \), and the inequality is sharp if \( v \geq 0 \).

The following lemma shows that condition (a) of Lemma 2.1 holds.

**Lemma 2.5.** The operator \( \mathcal{L}' : \mathcal{F} \to \mathcal{F} \) defined by \( \mathcal{L}'f = A(\omega)f(\sigma\omega) \) satisfies \( (\mathcal{L}')^* = \mathcal{L} \) and \( \|\mathcal{L}'\| \leq 1 \).

**Proof.** One has
\[
(\mathcal{L}'v)(f) = \int v(\sigma^{-1}\omega)A(\sigma^{-1}\omega)f(\omega) \ d\mathbb{P}(\omega) = \int v(\omega)A(\omega)f(\sigma\omega) \ d\mathbb{P}(\omega) = v(\mathcal{L}'f),
\]
where \( \mathcal{L}'f = A(\omega)f(\sigma\omega) \). Moreover,
\[
\|\mathcal{L}'f\| = \int \max_{1 \leq i \leq d} |(A(\omega)f(\sigma\omega))_i| \ d\mathbb{P}(\omega) \leq \int |A(\omega)|_{\ell^\infty} |f(\sigma\omega)|_{\ell^\infty} \ d\mathbb{P}(\omega) = \|f\|,
\]
as each \( A(\omega) \) is a row-stochastic matrix with \( |A(\omega)|_{\ell^\infty} = 1 \) (\( | \cdot |_{\ell^\infty} \) is the max-row-sum norm).

Conditions (b) and (c) of Lemma 2.1 will be treated on a case-by-case basis in Section 4 for four natural types of perturbations. To conclude this section we have the following connection between weak-* convergence in \( L^\infty \) and convergence in probability.

**Lemma 2.6.** Let \( g, g_n \in L^\infty(\mathbb{P}) \) for \( n \geq 0 \). Then,
1. \( g_n \to g \) weak-* \( \iff \) \( g_n \to g \) in probability.
2. If additionally, \( |g_n|_{\ell^\infty}, n \geq 0 \), are uniformly bounded then \( g_n \to g \) weak-* \( \iff \) \( g_n \to g \) in probability.

**Proof.** See appendix A.
In this paper our interest is in \( v \in \mathcal{V} \) such that for each \( \omega \in \Omega \), \( v(\omega) \) represents a \( 1 \times d \) probability vector (so \( |v(\omega)|_1 = 1 \) for \( \mathbb{P}\text{-a.e. } \omega \)). Thus, \( \|v\|_* = 1 \) and we will always be in situation (2) of Lemma 2.6 where statements of convergence may be regarded in either a weak-* sense or in probability. In the context of random invariant measures for Markov chains in random environments, \( v_n \to v \) in probability means \( \lim_{n \to \infty} \mathbb{P}(\{|v_n(\omega) - v(\omega)|_1 > \epsilon\}) = 0 \) for each \( \epsilon > 0 \).

3. Uniqueness of fixed points of \( \mathcal{L} \)

In this section, we derive an easily verifiable condition for \( \mathcal{L} \) to have a unique fixed point. Seneta [28] studied the coefficient of ergodicity in the context of stochastic matrices. We refer the reader to references therein for earlier appearances of related concepts.

**Definition 3.1.** Let \( M \) be a stochastic matrix. The coefficient of ergodicity of \( M \) is

\[
\tau(M) := \sup_{\|v\|_1 = 1, \sum_i |v_i| = 0} \|vM\|_1.
\]

One feature of this coefficient is that \( 1 \geq \tau(M) \geq \mu_2 \), where \( \mu_2 \) is the second eigenvalue of \( M \). In particular, when \( \tau(M) < 1 \), the eigenspace corresponding to \( \mu_1 = 1 \) is one-dimensional. In the random case, an analogous statement holds.

**Lemma 3.2** (Uniqueness). Suppose \( \tilde{\tau}(A) := \int \tau(A(\omega))d\mathbb{P}(\omega) < 1 \). Then \( \mathcal{L} \) has at most one fixed point (normalised in \( \|\cdot\|_* \)) in \( \mathcal{V} \).

**Proof.** We argue by contradiction. Suppose there exist \( v_1, v_2 \in \mathcal{V} \) distinct fixed points of \( \mathcal{L} \). The set of \( \omega \in \Omega \) such that \( v_1(\omega) \) and \( v_2(\omega) \) are linearly dependent is \( \sigma \) invariant, so by ergodicity of \( \sigma \) it has measure 0 or 1. The latter is ruled out because \( v_1 \) and \( v_2 \) are distinct normalised fixed points. Hence, \( v_1(\omega) - v_2(\omega) \neq 0 \) for \( \mathbb{P}\text{-a.e. } \omega \). Thus,

\[
\tilde{\tau}(A) \geq \int \frac{\|(v_1 - v_2)(\omega)A(\omega)\|_1}{\|(v_1 - v_2)(\omega)\|_1} d\mathbb{P} \\
= \int \frac{\|(v_1 - v_2)(\sigma\omega)\|_1}{\|(v_1 - v_2)(\omega)\|_1} d\mathbb{P}.
\]

However, the last expression is bounded below by 1, as the following sublemma, applied to \( f(\omega) = \|(v_1 - v_2)(\omega)\|_1 \), shows. \( \square \)

**Sublemma 3.3.** Let \( f \geq 0 \) be such that \( \frac{f(\sigma \omega)}{f(\omega)} \) is a \( \mathbb{P} \)-integrable function. Then,

\[
\int \frac{f(\sigma \omega)}{f(\omega)} d\mathbb{P} \geq 1.
\]

**Proof.** Jensen’s inequality yields

\[
\int \frac{f(\sigma \omega)}{f(\omega)} d\mathbb{P} \geq \exp \left( \int \log f(\sigma \omega) - \log f(\omega) d\mathbb{P} \right) = 1,
\]

where the equality follows from \( \sigma \)-invariance of \( \mathbb{P} \). \( \square \)

**Corollary 3.4.** Suppose there exist \( \hat{\Omega} \subset \Omega \) with \( \mathbb{P}(\hat{\Omega}) > 0 \) and \( n \in \mathbb{N} \), such that for every \( \omega \in \hat{\Omega}, \) the matrices \( A^{(n)}(\omega) := A(\omega) \cdots A(\sigma^{n-1}\omega) \) are positive. Then \( \mathcal{L} \) has a unique fixed point.
Proof. If \( v \) is a fixed point of \( \mathcal{L} \), it is obviously fixed by \( \mathcal{L}^n \). By applying Lemma 3.2 to the random cocycle \( (\sigma^n, A \cdot A \circ \sigma \cdots A \circ \sigma^{n-1}) \) we have that \( v \) is the unique fixed point of \( \mathcal{L}^n \), and thus the unique fixed point of \( \mathcal{L} \).

4. Types of Perturbations

In this section we verify that a variety of natural perturbations satisfy conditions (b) and (c) of Lemma 2.1.

4.1. Perturbing the random environment. We consider a sequence of base dynamics \( \sigma_n, n = 1, \ldots \) that are “nearby” \( \sigma \).

Proposition 4.1. Let \( \{\mathcal{R}_n\}_{n \in \mathbb{N}} = \{\{\Omega, \mathcal{F}, \mathbb{P}_n, \sigma_n, A, \mathbb{R}^d\}\}_{n \in \mathbb{N}} \) be a sequence of random dynamical systems with \( \mathbb{P}_n \) equivalent to \( \mathbb{P} \) for \( n \geq 1 \) that satisfies

\[
\lim_{n \to \infty} \|g \circ \sigma - g \circ \sigma_n\|_{L^1(\mathbb{P})} \to 0 \text{ for each } g \in L^1(\mathbb{P}),
\]

(in other words, the Koopman operator for \( \sigma_n \) converges strongly to the Koopman operator for \( \sigma \) in \( L^1(\mathbb{P}) \)). Then for each \( n \geq 1 \), associated with \( \mathcal{R}_n \), there is an operator \( \mathcal{L}_n \) with a fixed point \( v_n \in \mathcal{V} \). One may select a subsequence of \( \{v_n\}_{n \in \mathbb{N}} \) weak-* converging in \( (\mathcal{V}, \| \cdot \|_* \) to \( \tilde{v} \in \mathcal{V} \). Further, \( \tilde{v} \) is \( \mathcal{L} \)-invariant.

Proof. Firstly, for each \( n \) the existence of a fixed point of \( \mathcal{L}_n \) (defined as \( (\mathcal{L}_n v)(\omega) = v(\sigma_n^{-1} \omega)A(\sigma_n^{-1} \omega) \) is guaranteed at \( \mathbb{P} \) a.a. \( \omega \) by the MET [11] and equivalence of \( \mathbb{P}_n \) to \( \mathbb{P} \). Secondly, by Lemma 2.5 it is clear that \( \mathcal{L}_n \) exists as a bounded operator on \( \mathcal{F} \), defined as \( (\mathcal{L}_n f)(\omega) = A(\omega)f(\sigma_n \omega) \). Thirdly, \( \| (\mathcal{L}' - \mathcal{L}_n')f \| = \int \max_{1 \leq i \leq d} |A(\omega)(f(\sigma_n) - f(\sigma_n \omega))|_i \ d\mathbb{P}(\omega) \) goes to zero as \( n \to \infty \) if \( \| g \circ \sigma - g \circ \sigma_n\|_{L^1(\mathbb{P})} \) for any \( g \in L^1(\mathbb{P}) \), since the entries of \( A(\omega) \) are bounded between 0 and 1.

Remark 4.2.

(i) Note that we do not require \( \mathbb{P}_n \) to converge to \( \mathbb{P} \) in any sense.

(ii) If \( \Omega \) is a manifold and we work with the Borel \( \sigma \)-algebra defined by a metric \( \rho \) on \( \Omega \), then \( \sup_{\omega \in \Omega} \rho(\omega, \sigma_n \circ \sigma^{-1} \omega) \to 0 \text{ as } n \to \infty \), implies \( \|g \circ \sigma - g \circ \sigma_n\|_{L^1(\mathbb{P})} \) for any \( g \in L^1(\mathbb{P}) \) (eg. Corollary 5.1.1 [18]).

(iii) The result here is considerably more general than Ochs [23] applied to stochastic invertible matrices. For perturbations of the base, Ochs considers \( \Omega \) a topological space, \( \sigma \) a homeomorphism, and \( A \) a continuous matrix function. Moreover Ochs has a further requirement that, in our language, \( \sigma \) and \( \sigma_n, n \geq 0 \), all preserve \( \mathbb{P} \). The convergence result in our setting is equivalent to Ochs (convergence in probability).

Example 4.3. Let \( \Omega = T^D = \mathbb{R}^D / \mathbb{Z}^D \), the D-dimensional torus and \( \sigma \) be rigid rotation by an irrational vector \( \alpha \in \mathbb{R}^D \), which preserves D-dimensional volume \( \mathbb{P} \). Let \( \sigma_n(\omega) = \omega + \alpha_n, n \geq 0 \text{ where } \alpha_n \in \mathbb{R}^D \) is irrational, and \( \alpha_n \to \alpha \text{ as } n \to \infty \). Then for any given stochastic matrix function \( A : \Omega \to \mathcal{M}_{d \times d}(\mathbb{R}) \), one has \( v_n \to v \) in probability. If, for example, each matrix \( A(\omega) \) is a random walk on the finite set of states \( \{1, \ldots, d\} \) where there is a positive probability to remain in place and walk both left and right for each \( \omega \), then \( A^{(d-1)}(\omega) \) is a positive matrix for all \( \omega \in \Omega \) and by Corollary 3.4 there is a unique random probability measure. See also Section 5.1 for numerical computations.
4.2. Perturbing the transition matrix function. We consider a sequence of matrix functions $A_n$, $n \in \mathbb{N}$, that are nearby $A$.

**Proposition 4.4.** Let $A_n : \Omega \to \mathcal{M}_{d \times d}(\mathbb{R})$ be a sequence of measurable stochastic matrix-valued functions that converge in measure to $A$; that is

$$\lim_{n \to \infty} \mathbb{P}(\{ \omega \in \Omega : |A_n(\omega) - A(\omega)|_{\ell^\infty} > \epsilon \}) \to 0 \text{ for each } \epsilon > 0.$$  

For each $n \geq 1$, the random dynamical system $\{R_n\}_{n \in \mathbb{N}} := \{(\Omega, \mathcal{F}, \mathbb{P}, \sigma, A_n, \mathbb{R}^d)\}_{n \in \mathbb{N}}$ has an associated operator $L_n$ with a fixed point $v_n \in \mathcal{V}$. One may select a subsequence of $\{v_n\}_{n \in \mathbb{N}}$ weak-* converging in $(\mathcal{V}, \| \cdot \|_*)$ to $\tilde{v} \in \mathcal{V}$. Further, $\tilde{v}$ is $L$-invariant.

**Proof.** Existence of $v_n$ for $\mathbb{P}$ a.e. $\omega$ follows from the MET [11]. As in the proof of Proposition 4.1, the operators $L'_n : \mathcal{F} \odot$ are defined by $L'_n f(\omega) = A_n(\omega)f(\sigma \omega)$ and by Lemma 2.5 these are bounded. We need to check that $\|(L - L'_n)f\| \to 0$ as $n \to \infty$.

$$\|(L - L'_n)f\| = \int \max_{1 \leq i \leq d} \|[A(\omega) - A_n(\omega)]f(\sigma \omega)\| \, d\mathbb{P}(\omega) \leq \int |A(\omega) - A_n(\omega)|_{\ell^\infty} |f(\sigma \omega)|_{\ell^\infty} \, d\mathbb{P}(\omega).$$

Define $A_{\epsilon,n} = \{ \omega \in \Omega : |A(\omega) - A_n(\omega)|_{\ell^\infty} < \epsilon \}$. Then

$$\int |A(\omega) - A_n(\omega)|_{\ell^\infty} |f(\sigma \omega)|_{\ell^\infty} \, d\mathbb{P}(\omega) \leq \epsilon \int_{A_{\epsilon,n}} |f(\sigma \omega)|_{\ell^\infty} \, d\mathbb{P}(\omega) + \int_{A_{\epsilon,n}^c} |A(\omega) - A_n(\omega)|_{\ell^\infty} |f(\sigma \omega)|_{\ell^\infty} \, d\mathbb{P}(\omega) \leq \epsilon \|f\| + 2\int_{A_{\epsilon,n}^c} |f(\omega)|_{\ell^\infty} \, d\mathbb{P}(\omega).$$

Without loss, let $f$ have unit norm $\|f\| = 1$, and select some $\delta > 0$. Then choosing $\epsilon = \delta/2$, there is an $N$ such that for $n \geq N$, $\int_{A_{\epsilon,n}^c} |f(\sigma \omega)|_{\ell^\infty} \, d\mathbb{P}(\omega) < \delta/4$ and thus $\|(L - L'_n)f\| < \delta$ for $n \geq N$. \hfill \Box

**Remark 4.5.** In the context of stochastic matrices, Proposition 4.4 is analogous to Ochs [23], except that we can additionally handle non-invertible matrices.

4.3. Stochastic perturbations. We now consider the situation where the system is subjected to an averaging process. In this section we assume that $\Omega$ is a compact metric space with metric $\rho$. For each $n \geq 1$ let $k_n : \Omega \times \Omega \to \mathbb{R}$ be a nonnegative measurable function satisfying $\int k_n(\omega, \zeta) \, d\mathbb{P}(\zeta) = 1$ for a.a. $\omega \in \Omega$. Define $(\mathcal{L}_n v)(\omega) = \int k_n(\omega, \zeta) v(\sigma^{-1} \zeta) A(\sigma^{-1} \zeta) \, d\mathbb{P}(\zeta).$ We first demonstrate existence of fixed points.

**Proposition 4.6.** If for every $\omega_1, \omega_2 \in \Omega$, $\int |k_n(\omega_1, \zeta) - k_n(\omega_2, \zeta)| \, d\mathbb{P}(\zeta) \leq K_n(g(\omega_1, \omega_2))$, where $K_n$ is a function satisfying $\lim_{x \to 0} K_n(x) = 0$, then $\mathcal{L}_n$ has a fixed point $v_n \in \mathcal{V}$.

**Proof.** Let $S_{d+1}^{d-1} = \{ x \in \mathbb{R}^d : x \geq 0, \sum_{i=1}^d x_i = 1 \}$, and $S_{d+1}^{d-1} = \{ v \in \mathcal{V} : v(\omega) \in S_{d+1}^{d-1}, \omega \in \Omega \}$. For $v \in S_{d+1}^{d-1}$,

$$\sum_{i=1}^d [\mathcal{L}_n v(\omega)]_i = \int \sum_{i=1}^d [v(\sigma^{-1} \zeta) A(\sigma^{-1} \zeta)]_i k_n(\omega, \zeta) \, d\mathbb{P}(\zeta) = \int k_n(\omega, \zeta) \, d\mathbb{P}(\zeta) = 1,$$
for $\mathbb{P}$-a.e. $\omega$; thus $\mathcal{L}_n$ preserves $\mathcal{S}^{d-1}_+$. For $v \in \mathcal{S}^{d-1}_+$ we have

$$
\sum_{i=1}^{d} |[\mathcal{L}_n v(\omega_1) - \mathcal{L}_n v(\omega_2)]_i|
$$

$$
\leq \int \sum_{i=1}^{d} \left| [v(\sigma^{-1} \zeta) A(\sigma^{-1} \zeta)]_i (k_n(\omega_1, \zeta) - k_n(\omega_2, \zeta)) \right| d\mathbb{P}(\zeta)
$$

$$
\leq \int |k_n(\omega_1, \zeta) - k_n(\omega_2, \zeta)| d\mathbb{P}(\zeta)
$$

$$
\leq K_n(\varrho(\omega_1, \omega_2)).
$$

Fixing $n$, let $v_0$ be an arbitrary element of $\mathcal{S}^{d-1}_+$ and define $v^n_m = (1/m) \sum_{l=0}^{m-1} \mathcal{L}^l v_0$. The sequence $v^n_m$ of $\mathbb{R}^d$-valued functions is uniformly bounded coordinate-wise below by 0 and above by 1. Further,

$$
\sum_{i=1}^{d} |[\mathcal{L}^2_n v(\omega_1) - \mathcal{L}^2_n v(\omega_2)]_i|
$$

$$
\leq \int \int \sum_{i=1}^{d} \left| [v(\sigma^{-1} \rho) A(\sigma^{-1} \rho) A(\sigma^{-1} \zeta)]_i \right| (k_n(\sigma^{-1} \zeta, \rho) (k_n(\omega_1, \zeta) - k_n(\omega_2, \zeta))) d\mathbb{P}(\rho) d\mathbb{P}(\zeta)
$$

$$
\leq \int \int k_n(\sigma^{-1} \zeta, \rho) |k_n(\omega_1, \zeta) - k_n(\omega_2, \zeta)| d\mathbb{P}(\rho) d\mathbb{P}(\zeta)
$$

$$
\leq K_n(\varrho(\omega_1, \omega_2)).
$$

By induction, one has the same result for all powers of $\mathcal{L}_n$ and so one has that the sequence $v^n_m$ is equicontinuous coordinate-wise. By Arzela-Ascoli, we can extract a subsequence $v^n_{m_j}$ that converges uniformly to some $\tilde{v}^n$. We show that $\tilde{v}^n$ is a fixed point of $\mathcal{L}_n$ via a triangle inequality of the form $\|\tilde{v}^n - v^n_m\|_* + \|v^n_m - \mathcal{L}_n v^n_m\|_* + \|\mathcal{L}_n v^n_m - \mathcal{L}_n \tilde{v}^n\|_*$. The first term goes to zero by uniform convergence, the second by telescoping, and the last because $\|\mathcal{L}_n\| = 1$ and by uniform convergence of $v^n_m$ to $\tilde{v}^n$.

**Example 4.7.** The assumptions of Proposition 4.6 hold for the following natural random perturbations:

1. $\Omega = \mathbb{T}^D$, $\mathbb{P}$ is absolutely continuous with respect to Lebesgue measure, $d\mathbb{P}/d(\text{Leb})$ is uniformly bounded above, and $k_n \in L^1(\text{Leb})$. In this case, the statement follows from continuity of translations in $L^1(\text{Leb})$ (e.g. [15, Theorem 13.24]). This includes discontinuous $k_n$, for example, $k_n(\omega, \zeta) = 1_{B_n}(\omega - \zeta)$, where $1_{B_n}$ is the characteristic function of an $\epsilon_n$-ball centred at the origin. The operator $\mathcal{L}_n$ is performing a local averaging over an $\epsilon_n$-neighbourhood. Note that there are no assumptions of continuity on $\sigma$.

2. $k_n$ is uniformly continuous and $\mathbb{P}$ is arbitrary. In this case the statement follows immediately from the definition.
Proposition 4.8. Let \( v_n \in \mathcal{V} \) be a fixed point of \( \mathcal{L}_n \) for \( n \geq 0 \) as guaranteed by Lemma 4.6. If for each \( f \in L^1(\mathbb{P}) \),

\[
\lim_{n \to \infty} \int \left| \int k_n(\omega, \zeta) f(\zeta) \, d\mathbb{P}(\zeta) - f(\omega) \right| \, d\mathbb{P}(\omega) = 0,
\]

then one may select a subsequence of \( \{v_n\}_{n \in \mathbb{N}} \) weak-* converging in \( (\mathcal{V}, \| \cdot \|_*) \) to \( \tilde{v} \in \mathcal{V} \). Further, \( \tilde{v} \) is \( \mathcal{L} \)-invariant.

Proof. Firstly, one may check that \( \mathcal{L}_n' f(\omega) = \int k_n(\sigma \omega, \omega) A(\omega) f(\sigma \omega) \, d\mathbb{P}(\zeta). \) Now, for \( f \in \mathcal{F} \),

\[
\| \mathcal{L}_n' f - \mathcal{L}' f \| = \int \left| \int k_n(\sigma \omega, \omega) A(\omega) f(\sigma \omega) \, d\mathbb{P}(\zeta) - A(\omega) f(\sigma \omega) \right| \, d\mathbb{P}(\omega)
\]

\[
= \int \left| A(\sigma^{-1} \omega) \left( \int k_n(\sigma \omega, \zeta) \, d\mathbb{P}(\zeta) - f(\omega) \right) \right| \, d\mathbb{P}(\omega)
\]

\[
\leq \int \left| \int k_n(\sigma \omega, \zeta) \, d\mathbb{P}(\zeta) - f(\omega) \right| \, d\mathbb{P}(\omega),
\]

since \( |A(\sigma^{-1} \omega)|_{\ell^\infty} = 1 \). The result follows from Lemma 2.1. \( \square \)

Remark 4.9. The condition (2) is reminiscent of what has been called a “small random perturbation” by Khas’minskii [16] and later Kifer [17], in the context of deterministic dynamical systems governed by a continuous map \( T : X \to X \). In this setting, one asks about whether limits of invariant measures of stochastic processes formed by small random perturbations are invariant under the deterministic map \( T \). A sufficient condition for this to be the case is: for each continuous \( f : X \to \mathbb{R} \),

\[
\lim_{n \to \infty} \sup_{x \in X} \left| \int_X P_n(x, dy) f(y) - f(Tx) \right| = 0,
\]

where \( P_n : X \times \mathcal{B}(X) \to [0,1] \) is a transition function (\( \mathcal{B}(X) \) is the collection of Borel-measurable sets in \( X \)).

4.4. Perturbations arising from the numerical Ulam scheme on manifolds. Ulam’s method [32] is a common numerical procedure for estimating invariant measures of dynamical systems. In this final subsection we introduce a new modification of the Ulam approach to numerically estimate the random invariant measures \( v(\omega) \) simultaneously for each \( \omega \in \Omega \). We consider our numerical method to be a perturbation of the original random dynamical system and apply our abstract perturbation machinery.

Assume \( \Omega \) is a compact smooth Riemannian manifold, and let \( m \) be the natural volume measure, normalised on \( \Omega \). We suppose that \( \mathbb{P} \equiv m \) and that \( h := d\mathbb{P}/dm \) is uniformly bounded above and below. For each \( n \), let \( \mathcal{P}_n \) be a partition of \( \Omega \) (mod \( m \)) into \( n \) non-empty, connected open sets \( B_{1,n}, \ldots, B_{n,n} \). We require that \( \lim_{n \to \infty} \max_{1 \leq i \leq n} \text{diam}(B_{j,n}) = 0 \).

For each \( n \), we define a projection \( \pi_n : \mathcal{V} \to \mathcal{V} \) by

\[
\pi_n(v) = \sum_{i=1}^{n} \left( \frac{1}{m(B_i)} \int_{B_i} v(\omega) \, dm(\omega) \right) 1_{B_i}.
\]

Using the standard Galerkin procedure we define a finite-rank operator \( \mathcal{L}_n := \pi_n \mathcal{L} \pi_n \).
It will be useful to consider the $m$-predual of $\mathcal{L}$, which we denote $\mathcal{L}_m^\prime$: $\int \mathcal{L} v \cdot f \ dm = \int v \cdot \mathcal{L}_m^\prime f \ dm$. It is easy to verify that $\mathcal{L}_m^\prime f = A(\omega)f(\sigma\omega)h(\omega)/h(\sigma\omega)$. We first consider condition (c) of Lemma 2.1.

**Lemma 4.10.** $\mathcal{L}_n^\prime f = \pi_n(\mathcal{L}_n(\pi_n(f \cdot h)))/h$.

**Proof.** We repeatedly use the fact that $\int \pi_n v \cdot f \ dm = \int v \cdot \pi_n f \ dm$.

\[
\int \pi_n \mathcal{L} \pi_n v(\omega) \cdot f(\omega) \ d\mathbb{P}(\omega) = \int \pi_n \mathcal{L} \pi_n v(\omega) \cdot f(\omega) h(\omega) \ dm(\omega) = \int v(\omega) \cdot \pi_n (\mathcal{L}_m(\pi_n(f(\omega)h(\omega)))) \ dm(\omega) = \int v(\omega) \cdot \pi_n (\mathcal{L}_m(\pi_n(f(\omega)h(\omega))))/h(\omega) \ d\mathbb{P}(\omega).
\]

\[\]

**Lemma 4.11.** $\mathcal{L}_n^\prime$ is bounded.

**Proof.**

\[
\|\mathcal{L}_n^\prime f\| = \int \left| \pi_n \left( \frac{A(\omega)\pi_n(f(\sigma\omega)h(\sigma\omega)h(\omega))}{h(\sigma\omega)} \right) \right|_{\ell^\infty} \ dm(\omega), \text{ by Sublemma A.1}
\]

\[
\leq \frac{1}{\inf_{\omega \in \Omega} h(\omega)} \int \left| \pi_n (f(\sigma\omega)h(\sigma\omega)h(\omega)) \right|_{\ell^\infty} \ dm(\omega), \text{ since } |A(\omega)|_{\ell^\infty} \leq 1
\]

\[
\leq \frac{1}{\inf_{\omega \in \Omega} h(\omega)} \int \left| \pi_n (f(\sigma\omega)h(\sigma\omega)) \right|_{\ell^\infty} \ dm(\omega)
\]

\[
\leq \frac{\sup_{\omega \in \Omega} h(\omega)}{\inf_{\omega \in \Omega} h(\omega)^2} \int \left| f(\sigma\omega)h(\sigma\omega) \right|_{\ell^\infty} \ dm(\omega), \text{ by Sublemma A.1}
\]

\[
\leq \frac{\sup_{\omega \in \Omega} h(\omega)^2}{\inf_{\omega \in \Omega} h(\omega)^4} \int \left| f(\sigma\omega) \right|_{\ell^\infty} \ dm(\omega)
\]

\[
\leq \frac{\sup_{\omega \in \Omega} h(\omega)^2}{\inf_{\omega \in \Omega} h(\omega)^4} \|f\|
\]

\[\]

**Lemma 4.12.** $\int |\mathcal{L}_n^\prime f - \mathcal{L}_n^\prime f|_{\ell^\infty} \ d\mathbb{P} \to 0$ as $n \to \infty$. 

11
Proof. We will use the facts that $\|\pi_n\| \leq (\inf_{\omega \in \Omega} h(\omega))^{-1}$ (Sublemma A.1) and $|A(\omega)|_{\ell^\infty} \leq 1$ for each $\omega \in \Omega$.

$$\int \left| \pi_n \left( \frac{A(\omega)\pi_n(f(\sigma\omega)h(\sigma\omega))h(\omega)}{h(\sigma\omega)} \right) / h(\omega) - A(\omega)f(\sigma\omega) \right|_{\ell^\infty} d\mathbb{P}(\omega)$$

$$= \int \left| \pi_n \left( \frac{A(\omega)\pi_n(f(\sigma\omega)h(\sigma\omega))h(\omega)}{h(\sigma\omega)} \right) - A(\omega)f(\sigma\omega)h(\omega) \right|_{\ell^\infty} dm(\omega)$$

$$\leq \int \left| \pi_n \left( \frac{A(\omega)\pi_n(f(\sigma\omega)h(\sigma\omega))h(\omega)}{h(\sigma\omega)} \right) - \pi_n (A(\omega)f(\sigma\omega)h(\omega)) \right|_{\ell^\infty} dm(\omega)$$

$$+ \int \left| \pi_n (A(\omega)f(\sigma\omega)h(\omega)) - A(\omega)f(\sigma\omega)h(\omega) \right|_{\ell^\infty} dm(\omega)$$

$$\leq \frac{1}{\inf_{\omega \in \Omega} h(\omega)} \int \left| \frac{A(\omega)\pi_n(f(\sigma\omega)h(\sigma\omega))h(\omega)}{h(\sigma\omega)} - A(\omega)f(\sigma\omega)h(\omega) \right|_{\ell^\infty} dm(\omega)$$

$$+ \int \left| (\pi_n(f(\sigma\omega)h(\omega)) - f(\sigma\omega)h(\omega)) \right|_{\ell^\infty} dm(\omega).$$

The second term goes to zero as $n \to \infty$ as $|\pi_n g - g|_{L^1(m)} \to 0$ for any $g \in L^1(m)$. Continuing with the first term,

$$(4) \leq (\inf_{\omega \in \Omega} h(\omega))^{-1} \int \left| \pi_n(f(\sigma\omega)h(\sigma\omega))h(\omega) - f(\sigma\omega)h(\sigma\omega)h(\omega) \right|_{\ell^\infty} dm(\omega),$$

which also goes to zero as $n \to \infty$ as above. \qed

4.4.1. Numerical considerations. We wish to construct a convenient matrix representation of $\mathcal{L}_n$.

**Lemma 4.13.** Let $\pi_n(v) = \sum_{i=1}^n v^i \mathbf{1}_{B_i}$, where $v^i \in \mathbb{R}^d$. Then the action of $\mathcal{L}_n$ can be written

$$\pi_n \mathcal{L} \pi_n(v) = \sum_{j=1}^n \left( \sum_{i=1}^n v^i L_{ij} \right) \mathbf{1}_{B_j},$$

where

$$L_{ij} = \frac{\int_{B_j \cap \sigma B_i} A(\sigma^{-1}\omega) \ dm(\omega)}{m(B_j)}.$$
Proof.

\[
\pi_n \mathcal{L} \pi_n(v) = \sum_{j=1}^{n} \left( \frac{1}{m(B_j)} \int_{B_j} \left( \sum_{i=1}^{n} v^{i} \mathbf{1}_{B_i}(\sigma^{-1}\omega) \right) A(\sigma^{-1}\omega) \, dm(\omega) \right) \mathbf{1}_{B_j}
\]

\[
= \sum_{j=1}^{n} \left( \sum_{i=1}^{n} v^{i} \frac{1}{m(B_j)} \int_{B_j} \mathbf{1}_{\sigma B_i}(\omega) A(\sigma^{-1}\omega) \, dm(\omega) \right) \mathbf{1}_{B_j}
\]

\[
= \sum_{j=1}^{n} \left( \sum_{i=1}^{n} v^{i} \int_{B_j \cap \sigma B_i} A(\sigma^{-1}\omega) \, dm(\omega) \right) \mathbf{1}_{B_j}
\]

\[
= \sum_{j=1}^{n} \left( \sum_{i=1}^{n} v^{i} L_{ij} \right) \mathbf{1}_{B_j}.
\]

We note that if \( A(\omega) = A_i \) (a fixed matrix) for \( \omega \in B_i \cap \sigma^{-1}B_j \), then the expression (6) simplifies:

\[
L_{ij} = \int_{B_j \cap \sigma B_i} A(\sigma^{-1}\omega) \, dm(\omega)
\]

\[
= \int_{B_j \cap \sigma B_i} A(\sigma^{-1}\omega) / h(\omega) \, dP(\omega)
\]

\[
= \frac{1}{m(B_j)} \int_{\sigma^{-1}B_j \cap B_i} A(\omega) / h(\sigma \omega) \, dP(\omega)
\]

\[
= \frac{A_i}{m(B_j)} \int_{B_j \cap \sigma B_i} 1 / h(\omega) \, dP(\omega)
\]

\[
= \frac{A_i}{m(B_j \cap \sigma B_i)}.
\]

One could for example for \( \omega \in B_i \) replace \( A(\omega) \) with \( \bar{A}_i = 1/m(B_i) \int_{B_i} A(\omega) \, dm(\omega) \) for \( i = 1, \ldots, n \). Such a replacement would create an additional triangle inequality term in the proof of Lemma 4.12 to handle the difference \( \bar{A}(\omega) - A(\omega) \), but using the argument of the proof of Lemma 4.4 we see that any replacement that converges to \( A \) in probability, including the \( \bar{A} \) replacement will leave the conclusion of Lemma 4.12 unchanged. Thus, supposing that we have made such a replacement, and denoting \( P_{ij} = \frac{m(B_j \cap \sigma B_i)}{m(B_j)} \), we may write

\[
(7) \quad L_n = \begin{pmatrix}
    P_{11}A_1 & P_{12}A_1 & \cdots & P_{1n}A_1 \\
    P_{21}A_2 & P_{22}A_2 & \cdots & P_{2n}A_2 \\
    \vdots & \vdots & \ddots & \vdots \\
    P_{n1}A_n & P_{n2}A_n & \cdots & P_{nn}A_n
\end{pmatrix},
\]

where each block is a \( d \times d \) matrix. We now show that there is a fixed point of \( L \).
Lemma 4.14. For each $n$, the matrix $L_n$ has a fixed point $v_n$.

Proof. Let $S^d_n = \{x \in \mathbb{R}^{nd} : x \geq 0, \sum_{i=0}^{k} x_i = 1, k = 1, \ldots, n\}$. A fixed point exists by Brouwer: the set $S^d_n$ is convex and compact, and is preserved by $L_n$. To see the latter, the first block of length $d$ of the image of $x = [x^1 \cdots x^n]$ under $L$ is given by $\sum_{i=1}^{n} P_{ij} x^i A_i$, where $x^i$ is the $i^{th}$ block of length $d$. As each $A_i$ is row-stochastic, the sum of the entries of $x_i A_i$ remains 1; further note that $\sum_{i=1}^{n} P_{ij} = 1$ by the definition of $P$, so the summation is simply a convex combination of the $x_i A_i$. \hfill $\Box$

Numerically, one seeks a fixed point $v_n = [v^1 | v^2 | \cdots | v^n] L_{ij} = [v^1 | v^2 | \cdots | v^n]$. One can for example initialise with

$$v^0_n := [(1/d, \ldots, 1/d)(1/d, \ldots, 1/d)\cdots(1/d, \ldots, 1/d)]$$

and repeatedly multiply by the (sparse) matrix $L_n$. We have proved:

Proposition 4.15. Let $v_n \in V$ be constructed as a fixed eigenvector of $L_n$ in (7), and considered to be a piecewise constant vector field on $\Omega$, constant on each partition element $B_i$, $i = 1, \ldots, n$. As $n \to \infty$, one may select a subsequence of $\{v_n\}_{n \in \mathbb{N}}$ weak-* converging in $(V, || \cdot ||_v)$ to $\tilde{v} \in V$. The vector field $\tilde{v}$ is $C$-invariant.

Remark 4.16. The expression (7) is related to the constructions in [9] (Theorem 4.2) and [8] (Theorem 4.8). In [9], the focus was on estimating the top Lyapunov exponent of a random matrix product driven by a finite-state Markov chain, rather than approximating the top Oseledets space. In [8], the matrices $A$ were finite-rank approximations of the Perron-Frobenius operator, mentioned in the introduction, and one sought an equivariant family of absolutely continuous invariant measures of random Lasota-Yorke maps (piecewise $C^2$ expanding interval maps) with Markovian driving. In the present paper, we are able to handle non-Markovian driving and require no assumptions on the transition matrix function beyond measurability. Propositions 4.15 and 4.4 have enabled a very efficient numerical approximation of the random invariant measure by exploiting perturbations in both the random environment and the transition matrix function.

4.4.2. Application to perturbations arising from finite-memory approximations. A theoretical application of the Ulam approximation result is stability under finite-memory approximations of the driving process. Suppose that $\Omega = \prod_{i=-\infty}^{\infty} X$, where $X$ is a probability space, and that $\sigma : \Omega \to \Omega$ is the left-shift defined by $(\sigma \omega)_i = \omega_{i+1}$, where $\omega_i \in X$ is the $i^{th}$ component of $\omega$. Let $P$ be $\sigma$-invariant. The simplest example of such a $P$ is $P = \prod_{i=-\infty}^{\infty} p$, where $p$ is a probability measure on $X$; the dynamics of $\sigma$ now generate an id process on $X$ with distribution $p$. In general, $\sigma$ generates a stationary stochastic process on $X$ with possibly infinite memory.

Let $P_N$ be a partition of $\Omega$ into cylinder sets of length $N$. By applying the constructions above, with $m = P$ one obtains convergence in probability of the random invariant measures of the memory-truncated process to the random invariant measure of the original process.

5. Numerical Examples

In this section we illustrate our results with numerical experiments. §5.1 explores a model of a nearest-neighbor random walk, driven by an irrational circle rotation. §5.2 is more experimental, as we take the standard map as a forcing system to illustrate a relation between the structure of the driving map and the random invariant probability measure.
5.1. A random walk driven by a simple ergodic process. Our driving process is an irrational rotation of the circle. Let $\Omega = S^1$, $\alpha \not\in \mathbb{Q}$, and $\sigma(\omega) = \omega + \alpha \pmod{1}$ for $x \in \Omega$. We set $\mathbb{P}$ to Lebesgue on $S^1$; $\sigma$ preserves $\mathbb{P}$ and is ergodic. For fixed $\omega \in \Omega$, the matrix $A(\omega)$ is a random walk transition matrix on states $\{1, \ldots, d\}$.

For $1 < i < d$, there are possible transitions to states $i-1, i, i+1$ with conditional probabilities $A_{i,i-1}(\omega), A_{i,i}(\omega), A_{i,i+1}(\omega)$ given by

$$
\begin{cases}
0.8 - 1.2\omega, 0.1 + \omega, 0.1 + 0.2\omega, & 0 \leq \omega < 1/2; \\
0.3 - 0.2\omega, 1.1 - \omega, -0.4 + 1.2\omega, & 1/2 \leq \omega \leq 1.
\end{cases}
$$

For $i = 1$, $A_{i,i-1}(\omega) = 0$ and $A_{i,i}(\omega), A_{i,i+1}(\omega)$ are given by

$$
\begin{cases}
0.9 - 0.2\omega, 0.1 + 0.2\omega, & 0 \leq \omega < 1/2; \\
1.4 - 1.2\omega, 1.1 - \omega, -0.4 + 1.2\omega, & 1/2 \leq \omega \leq 1.
\end{cases}
$$

For $i = d$, $A_{i,i+1}(\omega) = 0$ and $A_{i,i-1}(\omega), A_{i,i}(\omega)$ are given by

$$
\begin{cases}
0.8 - 1.2\omega, 0.2 + 1.2\omega, & 0 \leq \omega < 1/2; \\
0.3 - 0.2\omega, 0.7 + 0.2\omega, & 1/2 \leq \omega \leq 1.
\end{cases}
$$

Roughly speaking, the closer $\omega$ is to 0, the greater the tendency to walk left; the closer $\omega$ is to 1, the greater the tendency to walk right; and the closer $\omega$ is to 1/2, the greater the tendency to remain at the current state. The matrix $A$ is a continuous function of $\omega$, except at $\omega = 0$, however, our theoretical results only require $A$ to be a measurable function of $\omega$, so we can also handle very irregular $A$.

Using $n = 5000$ partition elements for $\Omega$ and $d = 10$, we form the (sparse) matrix $L_n$ in (7), and compute the fixed left eigenvector; each of these operations takes less than 1 second in MATLAB. Figure 1 shows a numerical approximation of the random invariant measure using Ulam’s method. The $\omega$-coordinates are along the $x$-axis, and for a fixed vector $v(\omega) \in \mathbb{R}^d$, the 10 components are plotted as differently coloured bars. The value of $v(\omega)_i$ is equal to the height of the $i^{th}$ coloured bar at $x$-coordinate $\omega$; note the total height is unity for all $\omega \in \Omega$.

Let us consider first Figure 1(a), where $\alpha = 1/(20\sqrt{2}) \approx 0.0354$. This value of $\alpha$ represents a relatively slow evolution on the base. The peak probabilities to be in state 1, the left-most state (dark blue), occur around $\omega = 0.5$, after the driven random walk has been governed by many matrices favouring walking to the left (from $\omega = 0$ up to $\omega = 0.5$). Once the driving rotation passes $\omega = 0.5$, the random walk matrices now favour movement to the right, and probability of being in state 1 (dark blue) decreases, while the probability of being in state 10, the right-most state, (dark red) increases, the latter finally reaching a peak around $\omega = 1$. This high probability of state 10 continues for one more iteration of $\sigma$ (white for $\omega \in [0, \alpha]$), but once $\omega$ again passes $\alpha$, the probability of being in state 10 quickly declines as the matrices again favour movement to the left.

Figure 1(b) reduces the resolution of the Ulam approximation from 5000 bins on $\Omega$ to 500. One sees that the result is still very accurate, with only the very fine irregularities beyond the resolution of the coarser grid unable to be captured.

Figures 1(c),(d) show approximations of the random invariant measure with an identical setup to Figure 1(a), except that $\alpha = 1/\pi, 1/\sqrt{2}$, respectively. These rotations are relatively fast and so one does not see the unimodal “hump” shape in Figure 1(a); nevertheless, it is clear that there is a complicated interplay between the driving map $\sigma$ and the resulting invariant measures.
5.2. **Stochastic matrices driven by a two-dimensional chaotic map.** Our driving process is the so-called “standard map” on the torus $\mathbb{T}^2$, given by $\sigma(\omega_1, \omega_2) = (\omega_1 + \omega_2 + 8\sin(\omega_1), \omega_2 + 8\sin(\omega_1)) \pmod{2\pi}$, where we write $\omega = (\omega_1, \omega_2) \in \mathbb{T}^2$. The invertible map $\sigma$ is area-preserving and the number 8 is a parameter that we have chosen to be sufficiently large so that numerically it appears that there are no positive area $\sigma$-invariant sets and that Lebesgue measure is ergodic; we refer the reader to [7] for rigorous results on the standard family, which is still an active research topic. We firstly consider the stochastic matrix

\[ \int v_\omega \, d\mathbb{P}(\omega), \] where $\mathbb{P} = \text{Leb}$ and $v_\omega = \delta_\omega \times v(\omega)$, for the Markov chain in a random environment described in Section 5.1. Shown are cumulative distributions of $v_n(\omega) \in \mathbb{R}^{10}$ vs. $\omega$ for different random environments (different $\alpha$) and different numerical resolutions (different $n$). As $n \to \infty$, the $v_n(\omega)$ converge in probability to the $\mathbb{P}$-a.e. unique collection $\{v(\omega)\}_{\omega \in \Omega}$, which is equivariant: $v(\omega)A(\omega) = v(\sigma\omega)$ $\mathbb{P}$-a.e.
function $A : \mathbb{T}^2 \to \mathcal{M}_{2 \times 2}(\mathbb{R})$ defined by

\begin{equation}
A(\omega_1, \omega_2) = \begin{cases}
   A_l := \begin{pmatrix} 1/2 & 1/2 \\ 0 & 1 \end{pmatrix}, & 0 \leq \omega_1 < \pi; \\
   A_r := \begin{pmatrix} 2/3 & 1/3 \\ 2/3 & 1/3 \end{pmatrix}, & \pi \leq \omega_1 \leq 2\pi.
\end{cases}
\end{equation}

The matrix function $A$ is piecewise constant, and on the “right half” of the torus, the matrix is non-invertible.

We apply the new Ulam-based method using $n = 2^{16}$ partition cells in $\mathbb{T}^2$. Figure 2(a) shows the weight assigned to the first of the two states as a graph over $\Omega = \mathbb{T}^2$. The dark red area is the image of the right half of the torus under $\sigma$; these $\omega$ have just had the non-invertible matrix $A_r$ applied and thus, the probability of the first state is exactly $2/3$, independent of what its value was at $\sigma^{-1}\omega$. Following a similar reasoning, it is possible to characterise the invariant probability measure, as follows. For a point $(\omega_1, \omega_2) \in \mathbb{T}^2$ such that $\sigma^{-k}(\omega_1, \omega_2)$ lies in the right half, but $\sigma^{-j}(\omega_1, \omega_2)$ does not for every $0 < j < k$, the vector $v(\omega_1, \omega_2)$ is given by $(2/3, 1/3)A_{k-1}^t$.

Finally, we use the stochastic matrix function

\begin{equation}
A(\omega_1, \omega_2) = \begin{pmatrix} \sin(\omega_1/2) & 1 - \sin(\omega_1/2) \\ \cos((\omega_2 - \pi/2)/3) & 1 - \cos((\omega_2 - \pi/2)/3) \end{pmatrix},
\end{equation}

which is discontinuous at $\omega_2 = 0$. Figure 2(b) again shows the probability of being in the first state as a graph over $\Omega = \mathbb{T}^2$. In Figure 2(a) and (b), there is the appearance of continuity of probability values along curves. These curves approximate the unstable manifolds of the map $\sigma$. Two points on the
same unstable manifold will by definition have a similar pre-history. Because the matrix function \( A \) is continuous almost everywhere, the probability vectors corresponding to two points on the same unstable manifold will have been multiplied by similar stochastic matrices in the past. This provides an intuitive explanation for why the probability values appear to be continuous along unstable manifolds.

**Appendix A. Proofs**

**Proof of Lemma 2.6.** Let us start with (1). After replacing \( g_n \) with \( g_n - g \), it suffices to show that if \( g_n \in L^\infty \) and \( \int g_n f \, dP \to 0 \) as \( n \to \infty \) for each \( f \in L^1 \), then \( \|g_n\| \to 0 \) in probability. We show the contrapositive: Suppose \( \|g_n\| \) does not converge to 0 in probability. We will show \( \int |g_n| f \, dP \) does not converge to zero for every \( f \in L^1 \), yielding a contradiction.

By assumption, there exists some \( \epsilon > 0 \) such that \( \mathbb{P}( \{ \omega : |g_n(\omega)| \geq \epsilon \} ) \to 0 \). So there exists a sequence of sets \( E_n = \{ |g_n| \geq \epsilon \} \) with \( \mathbb{P}(E_n) \to 0 \) as \( n \to \infty \). Thus, \( \limsup_n \mathbb{P}(E_n) > 0 \).

Note that \( |g_n| \geq \epsilon 1_{E_n} \) for \( n \geq 0 \). Let \( E = \limsup E_n \), note \( \mathbb{P}(E) \geq \limsup_n \mathbb{P}(E_n) > 0 \), and set \( f = 1_E \). Then \( \int |g_n| f \, dP \geq \epsilon \int (1_{E_n} 1_E) \, dP = \epsilon \mathbb{P}(E_n \cap E) \) for all \( n \geq 0 \). To finish, we will show that \( \limsup \mathbb{P}(E_n \cap E) > 0 \).

We proceed by contradiction. Let \( 0 < \delta \) be such that \( \limsup_n \mathbb{P}(E_n) > \delta \). Suppose \( \limsup \mathbb{P}(E_n \cap E) = 0 \). Then, there exists \( N_0 \in \mathbb{N} \) such that for every \( n \geq N_0 \), \( \mathbb{P}(E_n \cap E) < \delta /2 \). From the definition of \( \limsup \), a point \( x \in E \) if and only if there is an infinite sequence \( \{n_j\} \) such that \( x \in E_{n_j} \) for all \( j \geq 1 \). Thus, for every \( x \notin E \) there exists \( t_x \in \mathbb{N} \) such that \( x \notin E_n \) for every \( n \geq t_x \). Let \( t : \Omega \to \mathbb{N} \cup \{ \infty \} \) be the function \( x \mapsto t_x \) if \( x \notin E \) and \( t(x) = \infty \) if \( x \in E \). That is, \( t \) is the supremum of \( n \) such that \( x \in E_n \). Since the sets \( E_n \) are measurable, so is \( t \). Hence, there exists \( N_1 \) such that \( \mathbb{P}(x \in \Omega \setminus E : t_x > N_1) < \delta /2 \).

Let \( n > \max(N_0, N_1) \) be such that \( \mathbb{P}(E_n) > \delta \). On the one hand, we have \( \mathbb{P}(E_n \cap E) < \delta /2 \). On the other hand, \( \mathbb{P}(E_n \setminus E) \leq \mathbb{P}(x \notin E : t_x \geq n) < \delta /2 \). Thus, \( \mathbb{P}(E_n) < \delta /2 + \delta /2 \), which yields a contradiction.

Now we show the remaining part of (2). Let \( \epsilon > 0 \). If \( \mathbb{P}(|g_n - g| > \epsilon) \to 0 \), then \( \int f(g_n - g) \, dP \leq \|f\|_1 + \delta_1(\epsilon) \|g_n - g\|_\infty \), where \( \delta_1(\epsilon) := \sup_{\Omega \in \mathbb{P}(\Omega) \leq \epsilon} \int |f| \, dP \). In particular, \( \delta_1(\epsilon) \to 0 \) as \( \epsilon \to 0 \), and the claim follows. \( \square \)

**Sublemma A.1.** Let \( f \in \mathcal{F} \). Then \( \|\pi_n f\| \leq (\inf_{\omega \in \Omega} h(\omega))^{-1} \|f\| \).

**Proof.** Without loss, we consider the situation where \( P_n \) consists of a single element, namely all of \( \Omega \). The argument extends identically to multiple-element partitions.

\[
\|\pi_n f\| = \int \max_{1 \leq i \leq d} \left| \int f_i \, dm \right| \, dP \\
\leq \int \max_{1 \leq i \leq d} \left( \int |f_i| \, dm \right) \, dP \\
= \int \max_{1 \leq i \leq d} \left( \int |f_i| \, dm \right) \\
\leq (\inf_{\omega \in \Omega} h(\omega))^{-1} \int \max_{1 \leq i \leq d} |f_i| \, dP \\
= (\inf_{\omega \in \Omega} h(\omega))^{-1} \|f\|. \square
\]
ACKNOWLEDGEMENTS

GF and CGT acknowledge support of an ARC Future Fellowship (FT120100025) and an ARC Discovery Project (DP110100068), respectively.

REFERENCES


