SPECTRAL GALERKIN METHODS FOR TRANSFER OPERATORS IN UNIFORMLY EXPANDING DYNAMICS

CAROLINE WORMELL

School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia

Abstract. We present spectral methods for numerically estimating statistical properties of uniformly-expanding Markov maps. We prove bounds on entries of the Fourier and Chebyshev basis coefficient matrices of transfer operators, and show that as a result statistical properties estimated using finite-dimensional restrictions of these matrices converge at classical spectral rates: exponentially for analytic maps, and polynomially for multiply differentiable maps.

Our proof suggests two algorithms for the numerical computational statistical properties of uniformly expanding Markov maps: a rigorously-validated algorithm, and a fast adaptive algorithm. We give illustrative results from these algorithms, demonstrating that the adaptive algorithm produces estimates of many statistical properties accurate to 14 decimal places in less than one-tenth of a second on a personal computer.

1. Introduction

Many dissipative chaotic dynamical systems have associated time-invariant statistical quantities that describe limits of long-time averages over the dynamics. Of particular importance is the absolutely continuous invariant measure (acim), over which the long-time behaviour of the system can be studied using ergodic theory. Many of the structurally simplest chaotic systems are endowed with acims [24]. Given an acim of a mixing system and looking at sufficiently smooth observables, one can derive other statistical quantities of interest such as diffusion coefficients. In low-dimensional systems, these kinds of statistical quantities are commonly calculated by solving linear problems involving the transfer operator [3, 9, 2].

This paper demonstrates that these linear problems can be solved numerically very efficiently using spectral methods. Our systems of study will be Markov uniformly expanding maps in one dimension. We will prove that...
Finite-dimensional approximations of these linear problems in spectral bases have solutions which converge to the true statistical quantities in a strong norm as the order of the approximation is increased. We will give rates of this convergence: in particular, we will show that estimates of statistical quantities converge exponentially for analytic maps. Because the convergence is so fast, highly accurate estimates of almost any statistical quantity of interest may be obtained conveniently and often in only a fraction of a second.

The improved speed and accuracy opens up many areas of one-dimensional dynamics for efficient and reliable numerical exploration. As an example, in [12] the authors make use of a Fourier spectral method to explore a particular rate of convergence in linear response theory, directing a subsequent proof. We further illustrate the potential of our method in this paper with some representative examples.

Transfer operators are associated with flows (e.g. of a dynamical system) and track the evolution of (probabilistic) mass under that flow. Given a discrete-time dynamical system \( f : \Lambda \to \Lambda \) with \( f \) non-singular with respect to Lebesgue, one can define a transfer operator associated with \( f \) as an endomorphism on \( L_1(\Lambda, \text{Leb}) \) with explicit action \([3, 7]\)

\[
(\mathcal{L}B)(x) = \sum_{f(y) = x} \frac{1}{|f'(y)|} B(y).
\]

Typically, one restricts transfer operators to smaller, invariant subspaces \( E \subseteq L_1(\Lambda, \text{Leb}) \), where \( E \) is typically a space of functions with some kind of regularity. If the acim \( \rho \) of \( f \) exists and is contained in \( E \), it is contained in the 1-eigenspace of \( \mathcal{L} \). rewrite noting that you actually change the norm e.g. BV

In this paper, we will choose \( E \) to be the space of bounded variation \( BV \), on which the maps we consider have spectral gaps. However, the theoretical approach in this paper is mostly agnostic to the Banach space \( E \) on which one wishes to prove convergence of estimates: one can obtain analogous results in any Banach space on which the transfer operator has a spectral gap and whose norm can be in some way related to spectral basis coefficients.

Transfer operator problems cannot in general be solved analytically, and must therefore be solved numerically. To do this on a computer, infinite-dimensional algebraic objects such as \( E \) and \( \mathcal{L} \) must be approximated by corresponding finite-dimensional objects. Typically, one projects \( E \) onto a series of finite approximation spaces \( (E_N)_{N \in \mathbb{N}} \), using suitable projection maps \((\mathcal{P}_N)_{N \in \mathbb{N}}\). Each space \( E_N \) is defined as the span of a basis set \( \{e_{j,N} \mid j \in Z_N\} \), with \( Z_N \) a finite index set.

One then approximates \( \mathcal{L} \) by the finite-dimensional operator

\[
\mathcal{L}_N = \mathcal{P}_N \mathcal{L} |_{E_N} : E_N \to E_N.
\]
One often-used family of bases are hat functions $e_{k,N}(x) = 1_{\Lambda_{k,N}}(x)$, for some partitions $\bigcup_{k=1}^{N} \Lambda_{k,N}$ of the phase space $\Lambda$. This approach is known as Ulam’s method [9]. While Ulam’s method holds for a wide array of chaotic maps, including rough and non-Markovian maps, hat functions are a relatively inefficient approximation scheme. In particular, the convergence rates of hat function approximations is at best $O(\log N/N)$, even for functions with a high degree of regularity [5]. Furthermore, approximations of operators only converge in $B(BV, L_1)$, which makes solving more complex transfer operator problems relatively difficult: for example, rigorous estimation of diffusion coefficients using Ulam’s method, even to one significant figure, demands a very large amount of computing power [2].

By contrast, approximations in spectral bases such as Fourier exponential functions or Chebyshev polynomials are extremely efficient for sufficiently regular approximands [22]. Algorithms to solve infinite-dimensional operator problems, known as spectral (Galerkin) methods, often converge as fast as exponentially.

The classical example of a spectral basis supported on a periodic domain is the Fourier exponential basis. The elements of the basis on the canonical domain $[0, 2\pi)$ are the complex exponentials $c_k(\theta) = e^{ik\theta}$, $k \in \mathbb{Z}$, and the finite approximation basis sets $Z_N := Z \cap [-N, N]$ are just subsets of this full basis. The projections $P_N$ are the orthogonal projections in the Hilbert space $L_2([0, 2\pi])$, in which the $c_k$ are orthogonal.

On non-periodic intervals, the most widely-used spectral basis for approximation is the Chebyshev polynomial basis. On its canonical domain $[-1, 1]$, the basis elements are the Chebyshev polynomials

$$T_k(x) = \cos \left( k \cos^{-1} x \right), \quad k \in \mathbb{N} =: \mathbb{Z},$$

and the finite basis sets are $Z_N := Z \cap [0, N]$. For Chebyshev polynomials we set $P_N$ to be the orthogonal projection in the weighted Hilbert space $L_2([-1, 1], (1 - x^2)^{-1/2})$. The Chebyshev polynomial basis can be seen as the Fourier cosine basis under the domain transformation $x = \cos \theta$, and its theory is closely linked to Fourier series theory [22].

Because of their mathematical simplicity, our spectral methods use Fourier and Chebyshev spectral bases for maps defined on periodic and non-periodic intervals respectively.

In [10], a spectral method for problems involving transfer operators of continuous time systems was proposed, using the infinitesimal generator of the (continuous-time) transfer operator. Numerical results using this method showed fast (super-polynomial) convergence of acim estimates to the true distributions. However, the theory developed in the paper was based on elliptic PDE theory, had as its concept the introduction of a small diffusion. This concept is in general
t was therefore unable to explain the speed of the convergence. In general, current techniques in spectral methods are not well-suited to transfer operator problems.

Standard mixing properties were used to prove convergence of eigenvalue and eigenvector estimates of transfer operators of circle maps in a wavelet basis in [4]. The authors suggested as directions for further research the establishment of quantitative convergence rates and extension to other manifolds, and also suggested that the method be implemented on a computer. This paper carries out these suggested projects in one dimension, looking only at resolvent data at the eigenvalue one. For simplicity of implementation we have chosen to use spectral basis functions instead of wavelet bases, although the two are mutually convertible.

There are several further directions for research. Numerical results indicate that the actual rates of convergence are slightly better than what we prove in this paper (see Section 5.2). It would be possible to improve our theoretical results by using the spectral bases’ natural Hilbert spaces instead of $BV$: proving convergence on appropriate scales of Hilbert spaces would also prove convergence of linear response estimates. Of some use would also be the rigorous justification of convergence of eigenvalues and eigenfunctions, and possibly even dynamical determinants (c.f. [4]). Beyond Markovian dynamics, there is the possibility of extending our results to induced maps of other one-dimensional systems, since our theory applies to maps with countably infinite branches and is more or less independent of the way matrix elements of $L_N$ are estimated. In particular, it may be possible to accurately estimate statistical properties of induced maps with exponential return time tails, such as those of the logistic map. Finally, the theory also appears to extend to higher dimensions, possibly including maps with contracting directions such as the standard map.

The paper is set out as follows. In Section 2.1 we will set up the problem by defining the classes of maps we consider, and in Section 2.2 we will state our main theorems. In Section 3 we outline algorithms for acim finding that we will use to demonstrate the possibilities of our spectral methods. In Section 4 we prove the theoretical results of the paper. In Section 5 we show some numerical results that illustrate the theoretical results and demonstrate the practical potential of the spectral method.

2. Set-up and main theorems

In this section, we summarise the paper’s chief mathematical developments. We will first set up the problem, introducing the maps under consideration. We then present the important results of the paper: Theorem 1 characterising an operator that explicitly solves many typical transfer operator problems; the main theorem, Theorem 2 which gives convergence of spectral operator estimates; its corollaries, Corollaries 1 and 2 which give
convergence of acims and other statistical properties; and finally the Theorems 3 and 4 which bound the magnitude of transfer operator spectral coefficient matrix entries $L_{jk}$, which are central to the proof of Theorem 2. We then finally present and discuss two algorithms that implement our spectral method in a rigorously validated framework and a fast adaptive setting respectively.

2.1. Systems under consideration. In this section we will first introduce the two classes of chaotic maps that we will consider in the paper: one defined on periodic intervals, and one defined on non-periodic intervals. We will then introduce and discuss some generalised distortion conditions that maps in these classes may optionally satisfy and which determine the rate of convergence of spectral approximations.

2.1.1. Maps on periodic domains. The first class $U_P$ we consider will be $C^2$ circle maps $f$ on the canonical periodic interval $[0, 2\pi)$, satisfying the uniform expansion condition

\[(E) \quad \lambda := \inf_{x \in \Lambda} |f'(x)| > 1, \]

and having bounded distortion $[14, 11, 8]$:

\[(DD_1) \quad \sup_{x \in \Lambda} \frac{|f''(x)|}{|f'(x)|^2} < \infty. \]

We will treat the inverses of maps in $U_P$ using lifts. The definition of $U_P$ implies that $f \in U_P$ can be lifted to a map $\hat{f}$ from $[0, 2\pi)$ to the periodic interval $[0, 2\beta\pi)$, where

\[\beta = \frac{1}{2\pi} \int_0^{2\pi} f'(x) dx \in \mathbb{N}^+. \]

The lift $\hat{f}$ is a bijection, with inverse map $v : [0, 2\beta\pi) \to [0, 2\pi)$. This is illustrated in Figure 1(a). For consistency with the maps defined in the next section, we also use the notation $v_\iota(x) = v(x + 2\iota\pi)$ for $x \in [0, 2\pi]$, $\iota \in I := \{0, 1, \ldots, \beta - 1\}$.

2.1.2. Maps on non-periodic domains. The second major class $U_{NP}$ are full-branch Markov maps defined on the non-periodic interval $\Lambda = [-1, 1]$. These may have possibly a countably infinite number of branches. We denote the extension of the $f|_{\Omega_\iota}$ to $\overline{\Omega_\iota}$ by $\hat{f}_\iota : \overline{\Omega_\iota} \to \Lambda$ and require that these are at least $C^2$. A map of this kind is illustrated in Figure 1(b).

We denote the branches of the inverse of $f$ by of $v_\iota := \hat{f}_\iota^{-1} : [-1, 1] \to \overline{\Omega_\iota}$.

Maps in $U_{NP}$ must satisfy three further conditions. Firstly as with maps in $U_{NP}$ they must satisfy bounded distortion condition $[DD_1]$, which may be reformulated as

\[(DD_1) \quad \sup_{x \in [-1, 1], i \in I} \frac{|v''_\iota(x)|}{|v'_\iota(x)|} < \infty. \]
To control high oscillatory behaviour of spectral basis functions near the endpoints of the interval we also introduce the uniform C-expansion condition: a map $f : [-1, 1] \rightarrow [-1, 1]$ is uniformly C-expanding if its associated C-expansion parameter

$$(\text{CE}) \quad \lambda = \inf_{x \in \cup \cup I \in I} \frac{\sqrt{1-x^2}}{\sqrt{1-f(x)^2}} |f'(x)| > 1.$$ 

This is equivalent to requiring $\cos \circ f \circ \cos^{-1}$ being uniformly-expanding with parameter $\lambda > 1$.

When one restricts to Markov maps obeying [DD], uniform C-expansion is, in effect, just uniform expansion. In particular, for every C-expanding $f$, all iterates $f^n$ for $n$ sufficiently large are uniformly expanding. Proofs and further discussion of C-expansion can be found in [A].

Finally, because of the singular $x = \cos \theta$ transformation from Fourier series to Chebyshev series, we will also introduce a constraint on maps in $U_{NP}$ regarding the size and position of the branch domain intervals $\mathcal{O}_i$. This partition spacing condition requires that

$$(P) \quad \sup \left\{ \frac{|\mathcal{O}_i|}{d(\mathcal{O}_i, \sigma)} : i \in I, \sigma \in \partial \Lambda, \sigma \notin \mathcal{O}_i \right\} = \Xi < \infty.$$ 

If $|I| < \infty$, this condition is trivially satisfied.
2.1.3. **Non-expanding maps.** We further consider maps that satisfy all the conditions of $U_P$ (resp. $U_{NP}$) but for which the associated expansion parameters (resp. $CE$) are merely positive, rather than strictly greater than one. These maps have similar structure to those in $U_P$ (resp. $U_{NP}$), except that their dynamics are not necessarily chaotic. We denote the class of such maps $U_P$ (resp. $U_{NP}$).

2.1.4. **Distortion conditions.** To obtain good convergence results we will optionally impose the following generalised distortion conditions on our maps. These generalised distortion conditions are equivalent to uniform bounds on derivatives of the distortion $\log |v|'$.

The first family of generalised distortion conditions are of the form

\[(DD_r) \sup_{i \in I, x \in \Lambda} \left| \frac{v_t^{(n+1)}(x)}{v_t'(x)} \right| = C_n < \infty, \quad n \leq r,
\]

for $r \in \mathbb{N}^+$. Note that the first of these conditions $(DD_1)$ is definitionally satisfied by maps in $U_P$ and $U_{NP}$.

Our second family of distortion conditions extend $(DD_1)$ to hold over an analytic neighbourhood of $\Lambda$. For periodic intervals, the neighbourhood is the closed complex strip $\Lambda = \{x + iy | x \in [0, 2\beta\pi], |y| \leq \delta\}$, for a given $\delta > 0$, on which $v'$ is assumed to be holomorphic. For non-periodic intervals, the neighbourhood is $\Lambda_{\delta}$, defined to be a Bernstein ellipse\(^1\) of parameter $e_\delta$, on which the $v_i$ are assumed to be holomorphic.

The distortion condition itself is that

\[(AD_\delta) \begin{cases} \sup_{z \in \Lambda_{\delta}} \left| \frac{v'(z)}{v(z)} \right| = C_{1,\delta} < \infty, & \Lambda = [0, 2\pi), \\ \sup_{i \in I, z \in \Lambda_{\delta}} \left| \frac{v_i''(z)}{v_i'(z)} \right| = C_{1,\delta} < \infty, & \Lambda = [-1, 1]. \end{cases}\]

If no stipulation is put on $C_{1,\delta}$, this distortion condition for a periodic (respectively a finite-branched non-periodic map) is satisfied simply if $v' \neq 0$ on $\Lambda_{\delta}$ (resp. $v_i' \neq 0$ on $\Lambda_{\delta}$).

To succinctly formulate the main theorems in the next section, we will associate to these distortion conditions **standard spectral rates of convergence**. We define these rates of convergence $\kappa(\cdot)$ as function classes:

\[\kappa(DD_r) = \{x \mapsto C(1 + x)^{-r} : C > 0\},\]

\[\kappa(AD_\delta) = \{x \mapsto Ce^{-\zeta x} : C > 0, \quad \zeta \in (0, \delta)\} .\]

2.2. **Main results.** We can now formulate the main theoretical results of this paper, beginning by introducing a novel operator derived from the transfer operator that explicitly generates acims and other statistical properties.

\(^1\)A Bernstein ellipse of parameter $\rho > 1$ is an ellipse in the complex plane centred at 0 with semi-major axis of length $\frac{1}{2}(\rho + \rho^{-1})$ along the real line and semi-minor axis $\frac{1}{2}(\rho - \rho^{-1})$. 
We define the solution operator inverse
\[ K = \text{id} - \mathcal{L} + u\mathcal{S} \]
and the solution operator
\[ \mathcal{S} = K^{-1} = (\text{id} - \mathcal{L} + u\mathcal{S})^{-1}, \]
where the functional \( \mathcal{S} \) is the total Lebesgue integral on \( \Lambda \) and \( u \) is a function on \( \Lambda \) such that \( \mathcal{S} u = 1 \).

Many statistical properties can be computed using resolvent data of \( \mathcal{L} \) at its eigenvalue 1: the solution operator inverse is a well-conditioned, invertible perturbation of \( \text{id} - \mathcal{L} \) which allows the resolvent data to be recovered.

For any transfer operator \( \mathcal{L} \) with a spectral gap (i.e. with a simple eigenvalue at 1 and the remaining spectrum bounded inside a disk of radius less than 1), the solution operator therefore solves for two important quantities, according to the following theorem:

**Theorem 1.** Let \( \mathcal{L} : E \to E \) be a transfer operator with a spectral gap. Choose \( u \in E \) with \( \mathcal{S} u = 1 \).

Then \( \mathcal{S} = (\text{id} - \mathcal{L} + u\mathcal{S})^{-1} \) is well-defined and bounded as an operator on \( E \), and
(a) If \( \rho \) is the unique acim with \( \mathcal{S} \rho = 1 \),
\[ \rho = \mathcal{S} u. \]
(b) For any \( \phi \in \ker \mathcal{S} \),
\[ \sum_{n=0}^{\infty} \mathcal{L}^n \phi = \mathcal{S} \phi. \]

**Remark 1.** As a result of Theorem 1, many important statistical quantities can be simply expressed using the solution operator and acim. For example, the Green-Kubo formula for diffusion coefficients given in (??) can be rewritten using Theorem 1(b) as
\[ \sigma_f^2(A) = \int_{\Lambda} A (2\mathcal{S} - \text{id})(\text{id} - \rho\mathcal{S})(\rho A) \, dx. \]

This closed formula enables effective rigorous calculation of diffusion coefficients.

We now provide some notation to enable us to state the main theorem, which proves the convergence of the spectral methods. We define the finite-dimensional subspaces \( (E_N)_{N \in \mathbb{N}^+} \)
\[ E_N = \begin{cases} \text{span}\{e_{-N}, \ldots, e_N\}, \Lambda = [0, 2\pi] \\ \text{span}\{T_0, \ldots, T_N\}, \Lambda = [-1, 1] \end{cases} \]
and the orthogonal projections \( \mathcal{P}_N \) onto the \( E_N \) in the relevant Hilbert space. We also define the spectral Galerkin operator discretisations
\[ \mathcal{L}_N = \mathcal{P}_N \mathcal{L}|_{E_N} \]
and
\begin{equation}
S_N := K^{-1}_N := (\text{id} - \mathcal{L}_N + u \mathcal{S}|_{E_N})^{-1},
\end{equation}
where the function $u$ is taken to be in $E_N$. (Typically, $u = 1/|\Lambda|$.)

Our main theorem can then be formulated as follows:

**Theorem 2.** Suppose $f \in U_P$ or $U_{NP}$, and satisfies a distortion bound $(DD)$ (resp. $(AD)$). Then there exist functions $K, \bar{K} \in \kappa(DD)$ (resp. $\kappa(AD)$) such that for sufficiently large $N$ and all $\phi \in E_N$,
\begin{align*}
\|\mathcal{L}_N \phi - L \phi\|_{BV} &< N\sqrt{NK(N)}\|\phi\|_{BV}, \\
\|S_N \phi - S \phi\|_{BV} &< N\sqrt{NK(N)}\|\phi\|_{BV}.
\end{align*}

For ease of expression, in the rest of this section we use the notation $(D)$ to denote either of $(DD)$ or $(AD)$.

Theorem 2 together with Theorem 1 directly implies the convergence of estimates of statistical quantities. In particular, the following corollary gives spectral convergence of the acim.

**Corollary 1.** Suppose $f \in U_P$ or $U_{NP}$, and satisfies a distortion bound $(D)$. Let $\rho_N = S_N u$. Then there exists $K \in \kappa(D)$ such that for all $N$ sufficiently large
\begin{equation}
\|\rho_N - \rho\|_{BV} < N\sqrt{NK(N)}.
\end{equation}

The next corollary gives strong convergence of $\sum_{n=1}^{\infty} \mathcal{L}^n$, and consequently many important statistical estimates (see (9) for an example).

**Corollary 2.** Suppose $f \in U_P$ or $U_{NP}$, and satisfies a distortion bound $(D)$. Then there exists $K \in \kappa(D)$ such that for $N$ large enough and all $\phi \in E_N \cap \ker \mathcal{S}$,
\begin{equation}
\left\| S_N \phi - \sum_{n=0}^{\infty} \mathcal{L}^n \phi \right\|_{BV} < N\sqrt{NK(N)}\|\phi\|_{BV}.
\end{equation}

Since the operators $\mathcal{L}_N$ and $S_N$ are endomorphisms on $E_N$, Theorem 2 and Corollary 2 show that the spectral method converges in operator norm within $E_N$. When attempting to estimate, for example, $S \phi$ for some $\phi \notin E_N$, one can simply substitute $\phi$ for its spectral discretisation $P_N \phi$, and propagate through the calculation the error arising from this substitution.

Critical to proving Theorem 2 are the following bounds on the entries $L_{jk}$ of the transfer operator matrix. We state two analogous theorems for transfer operators on periodic and non-periodic domains: the situation is illustrated in Figure 2. Abstractly, these results reformulate the characterisation of the transfer operator of a uniformly-expanding map as the sum of a strictly upper-triangular operator and a compact operator developed by [16, 4] in the context of $C^\infty$ circle maps in wavelet bases. The important development of our approach is the large amount of quantitative information
generated, which allows us to prove convergence rates and provide rigorous concrete bounds for specific maps.

**Theorem 3.** Suppose $f$ is in the class $\bar{U}_P$ satisfying some distortion bound (D), with $\lambda_1 \leq f' \leq \lambda_2$. Suppose $L$ is the matrix representing the transfer operator of $f$ in a Fourier exponential basis.

Then for every $p_1 > \lambda_1^{-1}$ and $p_2 < \lambda_2^{-1}$ there exists $K \in \kappa(D)$ such that for $j/k > p_1$ or $k = 0$,

$$|L_{jk}| \leq K(|j - p_1k|),$$

and for $j/k < p_2$ or $k = 0$,

$$|L_{jk}| \leq K(|j - p_2k|).$$
Remark 2. One can prove similar results for transfer operators with general weights (c.f. (1)):

\[(L_g \phi)(x) = \sum_{f(y) = x} g(y) \phi(y).\]

This class of operator includes transfer operators and composition operators \(C_v : \phi \mapsto \phi \circ v\).

From the previous results, we are able to prove extremely accurate rigorous bounds on maps in \(U_P\) and \(U_{NP}\) satisfying sufficiently strong distortion conditions. In particular, we prove the following bound on the Lanford map:

**Theorem 5.** Consider the Lanford map \(f : [0, 1] \to [0, 1], f(x) = 2x + \frac{1}{2} x (1 - x) \mod 1\).

(a) The Lanford map’s Lyapunov exponent \(L_{\text{exp}} := \int_\Lambda \log |f'| \rho \, dx\) lies in the range

\[L_{\text{exp}} = 0.657 \ 661 \ 780 \ 006 \ 597 \ 677 \ 541 \ 582 \ 413 \ 823 \ 832 \ 065 \ 743 \ 241 \ 069 \]
\[580 \ 012 \ 201 \ 953 \ 952 \ 802 \ 691 \ 632 \ 666 \ 111 \ 554 \ 023 \ 759 \ 556 \ 459 \]
\[752 \ 915 \ 174 \ 829 \ 642 \ 156 \ 331 \ 798 \ 026 \ 301 \ 488 \ 594 \ 89 \pm 2 \times 10^{-128}.\]

(b) The diffusion coefficient for the Lanford map with observable \(\phi(x) = x^2\) lies in the range

\[\sigma_f^2(\phi) = 0.360 \ 109 \ 486 \ 199 \ 160 \ 672 \ 898 \ 824 \ 186 \ 828 \ 576 \ 749 \ 241 \ 669 \ 997 \]
\[797 \ 228 \ 864 \ 358 \ 977 \ 865 \ 838 \ 174 \ 403 \ 103 \ 617 \ 477 \ 981 \ 402 \ 783 \]
\[211 \ 083 \ 646 \ 769 \ 039 \ 410 \ 848 \ 031 \ 999 \ 960 \ 664 \ 7 \pm 6 \times 10^{-124}.\]

These bounds are derived in Section 5.1.

3. Algorithms

Our results suggest a variety of possible algorithms to capture, given a map, statistical properties that can be expressed as \(S\phi\) for some \(\phi\), such as acims (7) and diffusion coefficients (9). We present two possible algorithms a practitioner might wish to use to calculate invariant measures: one that gives rigorous bounds on statistical properties but is somewhat cumbersome for exploratory use, and one that gives accurate but non-validated estimates that is much more convenient to use. In this section we describe the two algorithms, and then explain how in both algorithms we calculate elements of the transfer operator matrix. We will demonstrate the algorithms in Section 5.

\(^2\)Available theoretical bounds typically scale exponentially with the distortion bound \(C_1\) (see (1) and Appendix D). However, at least in the analytic case, the spectrally fast convergence dominates the large theoretical bounds. It is only necessary to control floating-point error using an appropriately high numerical precision. Alternatively and more generally, one may apply the approach of (11).
Algorithm 1 is a traditional fixed-order spectral method, implemented in interval arithmetic. It requires as input the map inverses $v_i$ and their derivatives, a spectral order $N$, various bounds associated with elements of the transfer operator, and a bound on the norm of the solution operator $S$; it then outputs an estimate for the acim $\rho$ with a rigorously validated $BV$ error.

**Input:** Map inverses and derivatives $v_i, v'_i$, $i \in I$; spectral order $N$; aliasing bounds $A_{jk}^{(N)}$ for $j, k = 1, \ldots, N$; bound $b^S \geq \|S\|_{BV}$; bound $b^L \geq \|L\|_{BV}$ (see Lemma 3); bounds $b^L_{jk} \leq |L_{jk}|$.

**Output:** High-precision floating-point vector $\tilde{\rho}$ containing spectral coefficients of acim estimate; rigorous $BV$ error bound $\tilde{\epsilon}_{\text{obs}}$.

1. Check that $b^E_N b^S < 1$: if this is not the case increase $N$;
2. Set the number of floating-point bits to be greater than $-\log_2(N^4 \cdot b^E_N)$;
3. Initialise $N \times N$ matrix of intervals $L^{(N)}$;
4. for $k \leftarrow 1$ to $N$ do
   5. Calculate interpolant values $q^{(k,N)} = \{L(b_k(x_{i,N}))\}_{i=1}^N$ in interval arithmetic, using (1);
   6. Calculate spectral coefficients of the interpolant $p^{(k,N)} = \text{FFT}(q^{(k,N)}) \ (\text{DCT}(q^{(k,N)}) \text{ in the Chebyshev case});$
   7. for $j \leftarrow 1$ to $N$ do
      8. Calculate spectral coefficient matrix entry $L_{jk}^{(N)}$ as $q_{j}^{(k,N)}$ plus aliasing error $[-A_{jk}^{(N)}, A_{jk}^{(N)}]$;
      9. Refine interval estimate $L_{jk}^{(N)}$ by intersecting it with $[-b^L_{jk}, b^L_{jk}]$;
   end
10. end
11. Calculate $u^{(N)} = \{[\delta_{j0}/|A|, \delta_{j0}/|A|]\}_{j=1}^N$;
12. Calculate row vector of intervals $S^{(N)} = (Sb_j)^{N}_{j=1}$ using standard formulae [22];
13. Calculate the spectral coefficient matrix of $S_N^{-1}$, $K^{(N)} = I - L^{(N)} + S^{(N)}u^{(N)}$, where $I$ is an $N \times N$ identity matrix;
14. Calculate $\rho^{(N)} = K^{(N)}\setminus u^{(N)}$;
15. Calculate $\tilde{\rho} = \{\text{midpoint}(\rho_j^{(N)})\}_{j=1}^N$;
16. Calculate a bound $\tilde{\epsilon}_{\text{interval}} \geq \|\rho^{(N)} - \tilde{\rho}\|_{BV}$;
17. Calculate $\epsilon_{\text{finite}} = 1/(1/(b^E_N b^S) - 1)$;
18. Calculate $\tilde{\epsilon}_{\text{obs}} = \tilde{\epsilon}_{\text{interval}} + \epsilon_{\text{finite}}$;

**Algorithm 1:** Rigorous algorithm to capture invariant measures.

By contrast, Algorithm 2 is an adaptive-order spectral method that is not rigorously validated: it uses an adaptive QR factorisation of the solution...
Input: Map \( f \); map derivative \( f' \) (optional; may be calculated automatically using dual number routines [21]); tolerance \( \epsilon \)

Output: Adaptive order \( k_{\text{opt}} \); floating-point vector \( \hat{\rho} \) containing spectral coefficients of acim estimate \( \hat{\rho}_{N_{\text{opt}}} \)

# Extendable vectors encode an infinite vector with finitely many non-zero entries, ragged matrices' columns are extendable vectors. These will encode infinite-dimensional objects approximating \( u \), \( K \).

Initialise empty ragged matrix \( H \), which will hold Householder vectors for row-reduction;

Initialise empty ragged matrix \( \hat{K} \), which will hold row-reduced coefficients of solution operator inverse \( K \);

Calculate extendable vector \( \hat{u} = (\delta_{1j} / |\Lambda|)_{j \geq 1} \) containing coefficients of \( u \) that will be progressively row-reduced;

Set \( k \), the number of columns of matrix \( \hat{K} \), to be 0;

repeat # Loop between calculating columns of \( \hat{K} \) and row-reducing

Increment \( k \) by 1;

Set the interpolation order \( M = 4 \);

repeat # Calculating optimum order interpolant of \( Lb_k \)

Set \( M \leftarrow 2M \);

Calculate values of the interpolant \( q^{(k)} = \{ L(b_k)(x_{l,M}) \}_{l=1}^M \) using (1) with Newton iteration for the transfer operator;

Calculate spectral coefficients of the interpolant \( p^{(k)} = \text{FFT}(q^{(k)}) \) (DCT(\( q^{(k)} \)) in the Chebyshev case);

until the interpolant has converged according to the reasoning in [1];

Calculate \( \kappa^{(k)} \), which will become the \( k \)th column of \( \hat{K} \), as an extendable vector \( \{ \delta_{jk} + \mathcal{S}_k \delta_{j1} \}_{j \geq 1} - p^{(k)} \), where \( \mathcal{S}_k = (\mathcal{S}_k b_k) \) is calculated from Chebyshev and Fourier integral formulae [22];

Apply previous Householder transformations encoded as column vectors of \( H \) to \( \kappa^{(k)} \);

Calculate Householder vector \( h \) that will row-reduce \( \kappa^{(k)} \) considered as the \( k \)th column of \( \hat{K} \);

Apply \( h \) to \( \kappa^{(k)} \);

Right-concatenate \( \kappa^{(k)} \) onto \( \hat{K} \);

# Note \( \hat{K} \) is row-reduced and so upper-triangular

Apply \( h \) to \( \hat{u} \);

Right-concatenate \( h \) onto \( H \);

until \( \max\{ u_j \}_{j \geq k+1} \leq \epsilon / |\Lambda|^{-1} \) # i.e. negligible benefit from larger \( k \);

Set \( N_{\text{opt}} = k \);

Calculate \( \hat{\rho} = \{ \hat{K}_{jk} \}_{j,k=1}^{N_{\text{opt}}} \{ \hat{u}_j \}_{j=1}^{N_{\text{opt}}} \) via backsolving;

Algorithm 2: Algorithm to capture invariant measures using adaptive interpolation and infinite-dimensional adaptive QR solver.
operator inverse $\mathcal{K}$ to solve the linear problem and test for convergence [20, 15]. It requires as input only an algorithm to calculate the map $f$ and outputs an estimate for $\rho$ whose error is not rigorously bounded but is of the order of $\|S\|_{BV} \epsilon^{1-\theta}$, where $\epsilon$ is the floating-point precision and $\theta$ is a small number depending on the order of differentiability of $f$.

Algorithm 2 is extremely well-suited for numerical exploration. Because the only required input is the map itself, Algorithm 2 requires a minimum of drudge work on the part of the user. It is typically also extremely fast: just with a personal computer, Algorithm 2 gives estimates of statistical quantities of a simple analytic map accurate to 14 decimal places in less than one-tenth of a second (see Section 5). Because our spectral methods are very accurate in an easily verifiable way, an adaptive, non-validated method is also highly reliable. We have consequently made an implementation of Algorithm 2 available in the open-source Julia package Poltergeist [23].

In the presentation of the algorithms and the following discussion we assume that the Fourier and Chebyshev spectral bases have been relabeled as $(b_k)_{k \in \mathbb{N}^+}$. We also implicitly assume that the Fourier exponential basis has been transformed to sines and cosines so that real functions have real spectral coefficients.

In both algorithms, one calculates $L_N$ by columns, using that the $k$th column of $L_N$ consists of the first $N$ spectral coefficients of $L_{b_k}$. The most effective way to estimate these coefficients is by calculating an interpolant. The idea of this is as follows. Using (1) one evaluates the function $L_{b_k}$ at $N$ special interpolation nodes $x_{l,N}$: in the Fourier case these interpolation nodes are evenly-spaced on the periodic interval (in the Chebyshev case respectively, Chebyshev nodes of the first kind) [22, 6]. One then applies the Fast Fourier Transform (resp. Discrete Cosine Transform) to the vector $((L_{b_k}(x_{l,N}))_{l=1,...,N}$. The resulting length-$N$ vector contains the spectral coefficients of the unique function $p^{(k,N)} \in E_N$ which matches $L_{b_k}$ at the interpolation nodes. The so-called interpolant $p^{(k,N)}$ is a close approximation of $L_{b_k}$: the difference between the $j$th spectral coefficient of $p^{(k,N)}$ and that of $L_{b_k}$ (the so-called aliasing error) is guaranteed to be smaller than some bound $A_{jk;N}$. This bound can be determined from aliasing formulæ standard in approximation theory [22] combined with bounds on higher-order spectral elements of $L_{b_k}$ (e.g. from Theorems 3-4).

These algorithms generalise very easily to other transfer operator problems of the form $\psi = S\phi$: see for example the formula for diffusion coefficients [9]. This can be done by formulating the problem as $\mathcal{K}\psi = \phi$ and thus substituting $\rho$ and $u$ (when it is not constituting the solution operator) for $\psi$ and $\phi$ respectively in the algorithms.

4. Proofs of results

Our attack on the theorems in Section 2.2 is structured as follows.
We begin by proving Theorem 1 characterising the solution operator. This proof uses standard linear-algebraic properties of transfer operators. We then turn to proving the entry bound results (Theorems 3 and 4). These results stem from more general properties of Fourier series representations of composition operators (Lemma 1), which we prove using oscillatory integral techniques. Because it is necessary to make a non-diffeomorphic cosine transformation to obtain Fourier basis functions from Chebyshev polynomials (see (2)), some work is required to prove appropriate bounds on derivatives after the transformation.

We then go on to prove Theorem 2. We consider a perturbation of the transfer operator \( L \) so that it becomes block-upper-triangular in the relevant spectral basis (in the Fourier case, under the basis order \( c_0, c_1, c_{-1}, c_2, c_{-2}, \ldots \)). In particular, we choose our perturbation \(-E_N\) so that the finite matrix \( L_N \) forms the first block on the diagonal. Since the solution operator is composed only from these kinds of upper block-diagonal operators, the first diagonal block can thus be approximated only from knowledge of \( L_N \) (Lemma 3).

Using that the BV-norm of the perturbation \(-E_N\) can be bounded using spectral matrix coefficients (Lemma 4), we obtain the main result.

We begin with the proof of Theorem 1, which gives the properties of the solution operator \( S = (I - L + u\mathcal{F})^{-1} \) (see (6)).

**Theorem 1** Split \( E \) as \( V \perp \oplus V \) where \( V \perp = \text{span}\{u\} \) and \( V = \ker \mathcal{F} \). We now consider the action with respect to this splitting of the putative solution operator inverse, \( K = \text{id} - L + u\mathcal{F} \).

Since \( \mathcal{F}(L - \text{id}) = 0 \), we have for any element \( \phi \in V \) that
\[
K\phi = (\text{id} - L)\phi + u\mathcal{F}\phi = (\text{id} - L)\phi \in V.
\]

Similarly, for any scalar \( \alpha \) we have
\[
K\alpha u = (\text{id} - L)(\alpha u) + u\mathcal{F}\alpha u = \alpha u + (\text{id} - L)(\alpha u),
\]

Identify the space \( E \) with the product space \( V \perp \times V \), which has an equivalent norm: the solution operator inverse \( K \) has the matrix form
\[
K = \begin{pmatrix}
\text{id} & 0 \\
\text{id} - L|_{V \perp} & \text{id} - L|_{V}
\end{pmatrix}.
\]

From the spectral gap assumption, the spectral radius of the transfer operator restricted to \( V \) is strictly less than 1: consequently the operator \( Q := (\text{id} - L|_V)^{-1} = \sum_{n=0}^{\infty} L^n|_V \) is bounded as an endomorphism on \( V \).

Hence, we are able to invert \( K \) to get that
\[
(13) \quad S = K^{-1} = \begin{pmatrix}
\text{id} & 0 \\
-Q(\text{id} - L)|_{V \perp} & Q
\end{pmatrix}.
\]

All elements of this array are bounded operators, so the operator itself is bounded.

Our claim was that \( Su = \rho \). The definition of an acim implies that \( S^{-1}\rho = K\rho = (\text{id} - L)\rho + u\mathcal{F}\rho = u \), which proves the claim.
We also claimed that $\mathcal{S}\phi = \mathcal{Q}\phi$ for $\phi \in V$. This follows directly from (13).

**Remark 3.** The solution operator can be written as the following expression

\begin{equation}
\mathcal{S} = u\mathcal{I} + \sum_{n=0}^{\infty} \mathcal{L}^n(d + (\mathcal{L}u - 2u)\mathcal{I}).
\end{equation}

We now set about proving Theorems 3 and 4, which place bounds on the magnitudes of the entries of transfer operator matrices in Fourier and Chebyshev bases.

We begin by proving similar kinds of bounds on the coefficients of a matrix associated with a more general operator $\mathcal{M}$ on the periodic interval $[0, 2\pi)$. $\mathcal{M}$ can be viewed as a generalised transfer operator (12) where instead of using the inverse of the map, one uses a general function $v$ which may be non-injective. Bounds on elements of the Fourier basis transfer operator matrix for $\mathcal{M}$ imply similar bounds on transfer operators in Fourier and Chebyshev bases.

**Lemma 1.** Let $v$ be a differentiable function from the periodic interval $[0, 2\pi\beta)$, $\beta \in \mathbb{Z}^+$ to the periodic interval $[0, 2\pi)$ such that $v'(0, 2\pi\beta)) = \tilde{\mu} = [\mu_2, \mu_1]$, and let $h$ be a continuous function on the periodic interval $[0, 2\beta\pi)$.

Let $\mathcal{M}$ be the corresponding bi-infinite matrix in the Fourier complex exponential basis.

Then:

(a) The elements of $\mathcal{M}$ are bounded uniformly by $\|h\|_1 / 2\pi$.

(b) Suppose that for $n = 1, \ldots, r$, $\sup |v^{(n+1)}| \leq \Upsilon_n < \infty$ and $\sup |h^{(n+1)}| / h \leq H_n < \infty$. Then there exist constants $W_{r,n}$ such that for $j \notin k\tilde{\mu}$,

\begin{equation}
|M_{jk}| \leq \frac{\|h\|_1}{2\pi} \sum_{n=0}^{r-1} \frac{W_{r,n}|k|^n}{d(j, k^{\tilde{\mu}})^{n+r}}.
\end{equation}

Each $W_{r,n}$ is bounded by a linear combination of $H_l$, $l \leq r - n$, whose coefficients are polynomials in $\Upsilon_l$, $l \leq r - n$.

(c) Suppose $v$ and $h$ extend analytically to the complex strip $\Lambda^\beta_\delta = [0, 2\beta\pi) + i[-\delta, \delta]$, and on this strip $\sup |h'/h| \leq H_{1,\delta} < \infty$ and $\sup |v''| \leq \Upsilon_1, \delta < \infty$.

Choose any $\tilde{p} = [p_1, p_2]$ such that $\tilde{\mu} \subset \int \tilde{p}$.

Define $\zeta = \min \left\{ 2\Upsilon_1^{-1}d(\tilde{\mu}, \mathbb{R}\setminus\tilde{p}), \delta \right\}$.

Then $\zeta > 0$ and

\begin{equation}
|M_{jk}| \leq \frac{\|h\|_1}{2\pi} e^{\zeta(H_{1,\delta} - d(j, \tilde{p}))}.
\end{equation}
Lemma 1. The matrix element $M_{jk}$ is the $j$th Fourier coefficient of the function $M_{ek}$, so using the orthogonality of Fourier bases in $L^2$ and (15), we have that

$$M_{jk} = \frac{1}{2\pi} \int_0^{2\beta} \sum_{b=1}^\beta h(x + 2\pi b) e^{ikv(x+2\pi b)} e^{-ijx} dx,$$

which using the $2\pi$-periodicity of $e^{ijx}$ we can rewrite as a single integral

(18) $$M_{jk} = \frac{1}{2\pi} \int_0^{2\beta} h(x) e^{i(kv(x)-jx)} dx.$$

We obtain (a) from this equation simply by taking absolute values.

For (b), we use that the integrand in (18) is oscillatory when the derivative of $kv(x) - jx$ is bounded away from zero, that is, when $j/k \notin [\mu_2, \mu_1]$. As a result, we can improve the bound we got in the first part by repeatedly integrating by parts.

Starting from (18), we separate the integrand into two terms

$$\left( \frac{h(x)}{i(kv'(x) - j)} \right) \left( i(kv'(x) - j)e^{i(kv(x)-jx)} \right),$$

so as to integrate by parts, differentiating the left term and integrating the right. Because the right term integrates to zero, the boundary terms in the integration by parts formula cancel, and we are left with an integral of the same form as (18) on which we can repeat the process. Thus we obtain a family of expressions

$$M_{jk} = (-1)^n \frac{1}{2\pi} \int_0^{2\beta} h_n(x) e^{i(kv(x)-jx)} dx, \quad n \leq r,$$

with each $h_n$ being $(r - n)$-times differentiable and defined by the recurrence relation

$$h_0 = h, \quad h_{n+1} = -i \left[ \frac{h_n}{j - kv'} \right]' .$$

We find by induction that

$$h_n = i^n \sum_{l=0}^n \frac{k^l w_{n,l}(x)}{(j - kv'(x))^{n+l}},$$

with $w_{n,l}$ having the recurrence relation

- $w_{n,l} = w_{n-1,l} + (n + l - 1)v'' w_{n-1,l-1}, \quad 0 < l < n,$
- $w_{n,0} = w_{n-1,0}, \quad n > 0,$
- $w_{n,n} = 2nv'' w_{n-1,n-1}, \quad n > 0,$
- $w_{0,0} = h.$
By induction, we see that each $w_{n,l}$ has the form

$$w_{n,l} = \sum_{l'=0}^{n-l} \omega_{n,l,l'}(v''(x), \ldots, v^{(n-l+2)}(x)) h^{(l')},$$

where $\omega_{n,l,l'}$ are degree $l$ homogeneous polynomials with positive coefficients. (The $\omega_{n,0,l'}$ are constants as a result, and thus issues of existence of derivatives do not arise.)

Setting $W_{n,l} = \sup_{x \in [0, 2\pi]} |w_{n,l}(x)| \leq \sum_{l'=0}^{n-l} \omega_{n,l,l'}(\Upsilon_1, \ldots, \Upsilon_{n-l+1}) H^{(l')}$, we have

$$|M_{jk}| \leq \frac{1}{2\pi} \int_0^{2\beta\pi} |h(x)| |k|^l \sum_{l=0}^{n} W_{n,l} |h(x)| |j - kv'(x)|^{n+l} dx,$$

from which (16) follows by Hölder’s inequality.

For (c), we use the $2\beta\pi$-periodicity of the integrand of (18) to move the contour of integration. When $j/k < p_2$, we shift the contour of integration by $-i\zeta \text{sgn} k$ in the complex plane so

$$(19) \quad M_{jk} = \frac{1}{2\pi} \int_0^{2\beta\pi} h(x - i\zeta \text{sgn} k) e^{ikv(x - i\zeta \text{sgn} k) - ij(x - i\zeta \text{sgn} k)} dx.$$ 

We now use our bounds on derivatives of $h$ and $v$ to bound elements of this expression, beginning with the argument of the exponential.

Applying Taylor’s theorem to $\Im v(x + i\xi)$, we have

$$\Im v(x - i\zeta \text{sgn} k) = -\zeta \text{sgn} k v'(x) - \frac{1}{2}\zeta^2 \Im v''(\xi)$$

for some $\xi \in \Lambda_{\beta}^{\delta}$. This gives us that

$$\Re \left( i k v(x - i\zeta \text{sgn} k) \right) \leq \zeta k \text{sgn} k |v'(x)| + |k| \frac{1}{2} \zeta^2 \Upsilon_{1,\delta}$$

$$\leq \zeta |k| \left( \mu_1 + \frac{\zeta \Upsilon_{1,\delta}}{2} \right)$$

$$\leq \zeta |k| p_1,$$

where the last inequality results from the definition of $\zeta$ in the statement of the lemma.

We can bound $h(x - i\zeta \text{sgn} k)$ by using that the Lipschitz constant of $\log h$ on $\Lambda_\delta^{\beta}$ is $\sup |h'/h| \leq H_{1,\delta}$. As a result,

$$|h(x - i\zeta \text{sgn} k)| \leq |h(x)| e^{i|\zeta \text{sgn} k| H_1} = |h(x)| e^{iH_{1,\delta}}.$$ 

Thus when we take absolute values on (19) we obtain that

$$|M_{jk}| \leq \frac{1}{2\pi} \int_0^{2\beta\pi} |h(x)| e^{iH_{1,\delta}} e^{i\zeta \text{sgn} k(kp_1 - j)} dx.$$
Using that $\text{sgn } k = \text{sgn}(j - p_1 k)$ for $j > p_1 k$ and Hölder’s inequality yields (17).

The proof of (c) for $j/k < p_2$ is analogous, with the contour shifted in the opposite direction.

□ □

Given Lemma 1, Theorem 3 is an elementary result. It is necessary only to check that the conditions for the theorem imply the conditions for the lemma, and vice versa for the results.

**Theorem 3.** From (1), the transfer operator $L$ of a map $f \in \bar{U}_P$ has action

$$L \phi(x) = \sum_{n=1}^{b} \sigma v'(x + 2b\pi)\phi(v(x + 2b\pi)),$$

where $\sigma = \text{sgn } v'(0)$. (Note that $v$ is monotonic and so $\sigma v' = |v'|$).

Since $\lambda_2^{-1} < |v'| < \lambda_1^{-1}$, we can apply Lemma 1 with $h = \sigma v'$.

Suppose that $f$ satisfies (DD). Then we can set $\Upsilon_n = C_n$ for all $n \leq r$, as the definition of $C_n$ in (DD) and of $\Upsilon_n$ in Lemma 1 are the same. We can also set

$$|v^{n+1}| \leq \frac{|v^n| |v'|}{2\pi} \leq \frac{C_n}{\min\{|\lambda_1|, |\lambda_2|\}} = H_n < \infty.$$

This gives us what we need for Lemma 1(b), and so there exist $W_{r,n}$ such that

$$L_{jk} \leq \|v'\|_1 \sum_{n=0}^{r} \frac{W_{r,n}|k|^n}{|j - \lambda_m^{-1} k|^{n+r}}$$

where $\lambda_m$ is $\lambda_1$ for $j/k > \lambda_1^{-1}$ and $\lambda_2$ for $j/k < \lambda_2^{-1}$.

We can eliminate the sum by using that for $j/k > p_1$,

$$\frac{|k|^r}{|j - \lambda_m^{-1} k|^r} = \frac{1}{(p_1 - \lambda_1^{-1})^r},$$

and similarly for $p_2$. Furthermore, since $v'$ does not change sign, $\|v'\|_1 = |v(2\pi\beta) - v(0)| = 2\pi$.

Thus there exists a constant $C$ depending on the distortion constants $C_r$, expansion bounds $\lambda_{1,2}$ and constants $p_{1,2}$ such that for $j/k \notin [p_2, p_1]$ or $k = 0$,

$$L_{jk} \leq \frac{C}{|j - \lambda_m^{-1} k|} \leq \frac{C}{|j - p_m^{-1} k|},$$

which implies the bound for maps in (DD) from Theorem 3.

Similarly, suppose that $f$ satisfies (DD). Then $\Upsilon_{1,\delta} = C_{1,\delta} < \infty$, and

$$\sup_{v \in \Lambda_{\delta}} |v'| \leq e^{3C_{1,\delta}} \sup_{x \in [0, 2\beta\pi]} |v'(x)| < \infty,$$

and hence by Lemma 1(c) there exists $C > 0$ and $\zeta \in (0, \delta]$ such that for $j/k > p_1$, $|L_{jk}| < Ce^{-\zeta |j - p_m k|}$, and similarly for $j/k < p_2$. □ □
Theorem 4 also follows from Lemma 1, since we can piggyback off the relation between Chebyshev polynomials and Fourier series:

\[ T_k(\cos \theta) = \frac{1}{2} e^{ik\theta} + \frac{1}{2} e^{-ik\theta}. \]

However, because the cosine function on \([0, 2\pi]\) is two-to-one with critical points at 0 and \(\pi\), the proof is less straightforward than for Theorem 3. In particular, we will have to address how to turn the transfer operator of a map in \(\tilde{U}_{NP}\) into the sum of operators of the form (15), with regard to the two-to-one nature of the transformation. We will then need to examine how distortion bounds translate quantitatively under this transformation. Once we have done these, the bounds follow easily.

**Theorem 4.** From the definition of transfer operators (1) and the orthogonality relation for the Chebyshev basis, we obtain the following formula for Chebyshev basis matrix elements of transfer operators of maps in \(\tilde{U}_{NP}\):

\[
L_{jk} = \frac{t_j}{\pi} \sum_{s \in I} \sigma_s \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} v'_s(x) T_k(v_s(x))) T_j(x) \, dx,
\]

where \(\sigma_s = \text{sgn} \, v'_s\), \(t_j = 2 - \delta_{j0}\), and the sum is taken over the branches of the map. Under the transformation \(x = \cos \theta\) and using that \(T_k(x) = \cos(k \cos^{-1} x)\), we find that \(L_{jk}\) is related to a Fourier basis matrix entry for a weighted transfer operator:

\[
(20)\quad L_{jk} = \frac{t_j}{\pi} \sum_{s \in I} \sigma_s \int_{0}^{\pi} v'_s(\cos \theta) \cos(k \cos^{-1} v_s(\cos \theta))) \cos j\theta \, d\theta.
\]

Based on this, we set \(h_s = v'_s \circ \cos\) for each \(s \in I\). These functions \(h_s\) are \(2\pi\)-periodic.

Defining \(\nu_{s+} := \cos^{-1} \circ v_s \circ \cos\) and \(\nu_{s-} = 2\pi - \nu_{s+}\), we find

\[
L_{jk} = \frac{t_j}{\pi} \sum_{s \in I} \sigma_s \int_{0}^{\pi} h_s(\theta) \cos(k \nu_{s+}(\theta))) \cos j\theta \, d\theta
\]

\[
= \frac{t_j}{4\pi} \sum_{s \in I} \sigma_s \int_{0}^{\pi} h_s(\theta) \sum_{\pm} \left( e^{i(k\nu_{s\pm}(\theta))} + e^{i(k\nu_{s\pm}(\theta))} \right) d\theta.
\]

Continuing \(\nu_{s\pm}\) differentiably to the non-periodic interval \([0, 2\pi]\) and using that the integrands are symmetric about \(\pi\), we can finally rewrite the transfer operator in the form

\[
(21)\quad L_{jk} = \frac{t_j}{8\pi} \sum_{s \in I, \pm} \sigma_s \int_{0}^{2\pi} h_s(\theta) \left( e^{i(k\nu_{s\pm}(\theta))} + e^{i(k\nu_{s\pm}(\theta))} \right) d\theta.
\]

If neither nor both of \(v_s(-1)\) and \(v_s(1)\) are a singular point of the \(\cos^{-1}\) transformation (i.e. \(-1\) or \(1\)), then the \(\nu_{s\pm}\) are differentiably defined on the periodic interval \([0, 2\pi]\). If one of these values is, then \(\nu_{s\pm}\) will continue across the critical points on either side to \(\nu_{s-}\) and so their concatenation \(\nu_s\) is a differentiable map on \([0, 4\pi]\). Thus, if we define the sets
\( I_c = \{ \nu \in I : |v_\nu(\{\pm 1\}) \cap \{\pm 1\}| = 1 \} \) and \( I' = (I \setminus I_c \times \{+,-\}) \cup I_c \), and set \( \beta_\nu = 1 + 1_{1c}(\nu') \), we have

\[
L_{jk} = \frac{t_j}{8\pi} \sum_{\nu' \in I'} \sigma_i \int_0^{2\pi} \beta_\nu \lambda_{\nu'}^{-1} h_\nu'(\theta + 2\pi b) e^{iku_\nu'(\theta + 2\pi b)} \left( e^{-ij\theta} + e^{ij\theta} \right) d\theta.
\]

Clearly, the summands are two-element sums of Fourier coefficient matrix elements of operators of the form \( L_{jk} \). The following lemma, whose proof is for ease of exposition in Appendix C, shows that that the relevant bounds on \( \nu_\nu' \) and \( h_\nu' \) hold uniformly for all \( \nu \):

**Lemma 2.** Suppose \( f \in \bar{U}_{NP} \) with partition spacing constant \( \Xi \) and \( I' \) is defined as above. Then

(a) If the \( v_\nu \) satisfy \( (DD_r) \) with the same distortion constants \( C_n, n \leq r \), then for \( n \leq r \) there exist \( \Upsilon_n, H_n < \infty \) depending only on \( C_m, m \leq n \) and \( \Xi \) such that

\[
\sup_{\theta \in [0,2\pi], \nu' \in I'} \nu_\nu'^{(n+1)}(\theta) \leq \Upsilon_n
\]

and

\[
\sup_{\theta \in [0,2\pi], \nu' \in I'} \left| \frac{h_\nu'(\theta)}{h_\nu'(\theta)} \right| \leq H_n < \infty.
\]

(b) If the \( v_\nu \) obey \( (AD_\delta) \) with the same distortion constant \( C_{1,\delta} \), then there exists \( \zeta \in (0, \delta], \Upsilon_{1,\zeta}, H_{1,\zeta} < \infty \) depending only on \( \zeta, C_{1,\delta} \) and partition spacing constant \( \Xi \) such that

\[
\sup_{\theta \in \Lambda_{\zeta', \nu'} \cup I'} \nu_\nu'(\theta) \leq \Upsilon_{1,\zeta}
\]

and

\[
\sup_{\theta \in \Lambda_{\zeta', \nu'} \cup I', \pm} \left| \frac{h_\nu'(\theta)}{h_\nu'(\theta)} \right| \leq H_{1,\zeta} < \infty.
\]

Setting \( \tilde{\mu} = [-\lambda^{-1}, \lambda^{-1}] \) and \( \tilde{p} = [-p, p] \), Lemma 3 means we can apply Lemma 1 to each summand in (22). Up to a constant factor \( G \) to be discussed later we have that if \( f \) satisfies \( (DD_r) \) then there exist \( W_{r,n} \) such that for \( j > pk \geq 0 \),

\[
|L_{jk}| \leq 2 \frac{t_j}{8\pi} G \sum_{n=0}^{r} \frac{W_{r,n} k^n}{(j - \lambda^{-1} k)^{n+r}}.
\]

Similarly, if \( f \) satisfies \( (AD_\delta) \) there exists \( \zeta' \in (0, \zeta] \) such that for \( j > pk \geq 0 \),

\[
|L_{jk}| \leq 2 \frac{t_j}{8\pi} G e^{\zeta'(H_{1,\zeta} - (j - pk))}
\]

which gives the decay rates stated in Theorem 4 by the same means as in the proof of Theorem 3.
However, we need to check that the constant factor

\[ G = \sum_{\iota' \in I'} \int_0^{2\beta_1 \pi} |v'_{\iota'}(\cos \theta)|d\theta \]

is in fact finite.

We convert back to a sum over \( I \) by collapsing the sum over \( \pm \) for \( \iota \in I \setminus I_c \), obtaining

\[ G = 2 \sum_{\iota \in I} \int_0^{2\pi} |v'_{\iota}(\cos \theta)|d\theta. \]

We then make the two-to-one change of variable \( x = \cos \theta \) to find that

\[ G = 4 \sum_{\iota \in I} \int_{-1}^{1} |v'_{\iota}(x)| \frac{1}{\sqrt{1-x^2}} dx \]

\[ \leq 4 \sum_{\iota \in I} \int_{-1}^{1} \frac{(1 + 2C_1)|O_{\iota}|}{2\sqrt{1-x^2}} dx = 4\pi(1 + 2C_1) < \infty, \]

where the first inequality is a result of Lemma 5(b).

This concludes the proof of Theorem 4. □ □

**Remark 4.** Elements of Fourier and Chebyshev transfer operator matrices are uniformly bounded, with

\[ |L_{jk}| \leq 1 \]

for maps in \( \bar{U}_P \) and

\[ |L_{jk}| \leq (2 - \delta_j)(2 + 4C_1) \]

for maps in \( \bar{U}_{NP} \).

This follows by applying Lemma 7(a) in the proofs of Theorems 3-4.

**Remark 5.** The C-uniform expansion condition (CE) is universally the natural expansion condition for spectral basis functions on non-periodic intervals. Our reasoning is as follows. If one wishes to use oscillatory interval techniques on these basis functions as in Lemma 4, it is best for the wavelength of the basis functions to be approximately spatially constant. However, wavelengths of sufficiently high-order spectral basis functions on non-periodic intervals will always be much smaller towards the endpoints of the interval. Potential theory [22] tells us that the optimal transformation to even out high-order basis functions across the interval is always the cosine transformation.

We now turn to proving the main theorem, Theorem 2 and its corollaries. Our idea is to perturb \( L \) such that the associated coefficient matrix is block-upper-triangular (in the Fourier case, with the ordering of basis elements \( e_0, e_1, e_{-1}, e_2, \ldots \)). This isolates the top block \( E_N \rightarrow E_N \), which then approximates the corresponding \( E_N \times E_N \) block of the full, unperturbed transfer operator, yielding convergence on domain \( E_N \).

We summarise this space-agnostically using the following lemma.
Lemma 3. Let $\mathcal{E}_N = (\text{id} - \mathcal{P}_N)\mathcal{L} |_{\mathcal{E}_N}$. Suppose $\mathcal{L}$ has a spectral gap. Then
\begin{equation}
\|\mathcal{L}_N - \mathcal{L}|_{\mathcal{E}_N}\|_{\mathcal{E}} = \|\mathcal{E}_N\|_{\mathcal{E}}
\end{equation}
and
\begin{equation}
\|\mathcal{S}_N - \mathcal{S}|_{\mathcal{E}_N}\|_{\mathcal{E}} \leq \frac{\|\mathcal{S}\|_{\mathcal{E}_N}\|\mathcal{E}_N\|_{\mathcal{E}}}{1 - \|\mathcal{S}\|_{\mathcal{E}_N}\|\mathcal{E}_N\|_{\mathcal{E}}}.
\end{equation}

Proof. The first equality (23) arises simply because $\mathcal{L}_N - \mathcal{L}|_{\mathcal{E}_N} = (\text{id} - \mathcal{P}_N)\mathcal{L}_N|_{\mathcal{E}_N} = \mathcal{E}_N$.

To show the second bound, we decompose the Banach space $\mathcal{E}_N$ into two closed subspaces $\mathcal{E}_N \oplus \mathcal{F}_N$, where $\mathcal{F}_N = \ker \mathcal{P}_N$. The transfer operator $\mathcal{L} : \mathcal{E}_N \rightarrow \mathcal{E}_N$ decomposes accordingly:
\begin{equation}
\mathcal{L} = \begin{pmatrix} \mathcal{L}_N & \mathcal{B}_N \\ \mathcal{E}_N & \mathcal{C}_N \end{pmatrix}.
\end{equation}

Now let $\tilde{\mathcal{L}}_N := \mathcal{L} - \mathcal{E}_N$. This operator has block decomposition
\begin{equation}
\tilde{\mathcal{L}}_N = \begin{pmatrix} \mathcal{L}_N & \mathcal{B}_N \\ \mathcal{E}_N & \mathcal{C}_N \end{pmatrix}.
\end{equation}

Recalling that we defined $\mathcal{S}$ to be total Lebesgue integral and $u$ an element of $\mathcal{E}_N$ with $\mathcal{S}u = 1$, we can define $\tilde{\mathcal{S}}_N := (\text{id} + \mathcal{S}\mathcal{E}_N)^{-1} \mathcal{S}$ and thus
\begin{equation}
\|\tilde{\mathcal{S}}_N - \mathcal{S}\| \leq \frac{\|\mathcal{S}\|\|\mathcal{E}_N\|}{1 - \|\mathcal{S}\|\|\mathcal{E}_N\|}.
\end{equation}

In block form, we have that
\begin{equation}
\tilde{\mathcal{S}}_N = \begin{pmatrix} (\text{id} - \mathcal{L}_N + u\mathcal{S})^{-1} - (\text{id} - \mathcal{L}_N + u\mathcal{S})^{-1}\mathcal{B}_N(\text{id} - \mathcal{C}_N)^{-1} \\ 0 \end{pmatrix},
\end{equation}
and it is clear that the upper-left block is just $\mathcal{S}_N$ (see (11)). Thus, $\tilde{\mathcal{S}}_N|_{\mathcal{E}_N} = \mathcal{S}_N$ and so $\|\mathcal{S}_N - \mathcal{S}|_{\mathcal{E}_N}\| = \|(\tilde{\mathcal{S}}_N - \mathcal{S})|_{\mathcal{E}_N}\| \leq \|\tilde{\mathcal{S}}_N - \mathcal{S}\|$. Combining this with (27) proves (24). □

The following lemma is then required to connect $\|\mathcal{E}_N\|_{BV}$ to spectral matrix coefficients.

Lemma 4. Suppose $\mathcal{F} : BV(\Lambda) \rightarrow BV(\Lambda)$ is an operator for $\Lambda$ either $[0, 2\pi)$ or $[-1, 1]$. Let the matrix $D = (k\delta_{jk})_{j,k\in\mathbb{Z}}$ and $\tilde{D} = (k\delta_{(j-1)k})_{j,k\in\mathbb{N}}$.

If $\mathcal{F}$ has Fourier coefficient matrix $F$, then
\begin{equation}
\|\mathcal{F}\|_{BV} \leq 2\pi(\|DF\|_{\ell^2} + \|F\|_{\ell^2}).
\end{equation}

Similarly, if $\mathcal{F}$ has Chebyshev coefficient matrix $F$, then
\begin{equation}
\|\mathcal{F}\|_{BV} \leq 2\pi(\|\tilde{D}\tilde{F}\tilde{C}^{-1}\|_{\ell^2} + \|\tilde{F}\tilde{C}^{-1}\|_{\ell^2}),
\end{equation}
where $\tilde{C} = (\tilde{k}^{-1/2}\delta_{jk})_{j,k\in\mathbb{N}}$. 

Proof. Consider first the Fourier case. Then since \( \frac{1}{\sqrt{2\pi}} \| \cdot \|_{L^2} \leq \| \cdot \|_{H^1}, \)
\[
\|F\|_{BV} \leq 2\pi \|F\|_{L^2} \to H^1 = \left( \|DF\|_{L^2} + \|F\|_{L^2} \right).
\]
By the Plancherel equality, \( \|DF\|_{L^2} = \|DF\|_{\ell^2} \) and \( \|F\|_{L^2} = \|F\|_{\ell^2} \). This gives the required bound in (28).

Consider instead the Chebyshev case. Define the Jacobi weight function \( j(x) = \sqrt{1-x^2} \), and the Sobolev spaces \( \hat{H}^k \subset L^2([-1,1], 1/j) \), \( k \geq 0 \) with norm
\[
\|\phi\|_{\hat{H}^k} = \sum_{n=0}^k \int_{-1}^1 j^{2n-1}|\phi^{(n)}|^2 \, dx.
\]
Note that \( \hat{H}^0 = L^2([-1,1], 1/j) \).

If \( G \) is the set of even functions on periodic interval \([0,2\pi]\), simple trigonometric manipulations show that the operator \( C : \phi \mapsto \frac{1}{2} \phi \circ \cos \) is an isometry from \( \hat{H}^k \) to \( G \cap H^k([0,2\pi]) \) and similarly from \( \hat{BV}([-1,1]) \) to \( G \cap BV([0,2\pi]) \). Thus,
\[
\|F\|_{BV([-1,1])} = \|\mathcal{C} \mathcal{F} C^{-1}\|_{G \cap BV([0,2\pi])} \leq 2\pi \left( \|\mathcal{D} \mathcal{C} \mathcal{F} C^{-1}\|_{G \cap L^2} + \|\mathcal{C} \mathcal{F} C^{-1}\|_{G \cap L^2} \right),
\]
where the inequality comes from (30). We can then convert back to \( \hat{H}^0 \) to get the inequality
\[
\|F\|_{BV([-1,1])} \leq 2\pi \left( \|C^{-1} \mathcal{D} \mathcal{C} \mathcal{F}\|_{\hat{H}^0} + \|\mathcal{F}\|_{\hat{H}^0} \right).
\]

We can convert these operator norms into matrix norms as follows. The Chebyshev polynomial basis is an orthogonal basis for \( \hat{H}^0 \) with \( \|T_k\|_{\hat{H}^0} = \sqrt{\pi/t_k} \) and furthermore the functions
\[
C^{-1} \mathcal{D} \mathcal{T}_k = k \sin(k \cos^{-1} x)
\]
are orthogonal in \( \hat{H}^0 \) with norms \( k \sqrt{\pi/t_k} \) respectively. The resulting Plancherel equality results in (29). \( \Box \)

We now have the requisite results to tie together to prove Theorem 2.

**Theorem 2.** Maps in \( U_P \) have a spectral gap in \( BV \) as they are uniformly expanding with bounded distortion. Since maps in \( U_{NP} \) have a forward iterate that is uniformly expanding with bounded distortion by Theorem 6 in Appendix A, they also have a spectral gap in \( BV \).

Suppose \( E_N \) is the Fourier coefficient matrix of \( \mathcal{E}_N \) and the expansion coefficient of the associated map \( f \) is \( \lambda > 1 \). Then given \( p \in (\lambda, 1) \), there exists an appropriate spectral decay function \( K \) such that when \( |j| \geq |k| \),
\[
|L_{jk}| \leq K(|j| - p|k|).
\]
Now suppose $f$ satisfies $(\text{DD}_r)$ for some $r \geq 2$. Then $K(M) = CM^{-r}$ for some $C > 0$, and so

$$\|E_N\|_{L^2}^2 \leq \sum_{k=-N}^{N} \sum_{j=N+1}^{\infty} (|L_{jk}|^2 + |L_{-jk}|^2)$$

$$\leq \sum_{k=-N}^{N} \sum_{j=N+1}^{\infty} 2C^2(j - p|k|)^{-2r}$$

$$\leq \frac{2C^2}{2r - 1} \sum_{k=-N}^{N} (N - p|k|)^{1-2r},$$

by converting to an integral. We can then take the supremum of the summands to obtain

$$\sum_{k=-N}^{N} (N - p|k|)^{1-2r} \leq (2N + 1)(N - pN)^{1-2r} \leq \frac{3}{(1-p)^{2r-1}} N^{2-2r}$$

and thus

$$\|E_N\|_{L^2}^2 \leq \frac{6}{(2r - 1)(1-p)^{2r-1}} N^2 K(N)^2.$$  

Similarly,

$$\|DE_N\|_{L^2}^2 \leq \frac{6}{(2r - 2)(1-p)^{2r-2}} N^3 K(N)^2,$$

where $D$ is as in Lemma 3.

Hence as a result of Lemma 4 there exists a function $K' \in \kappa(\text{DD}_r)$ such that $\|E_N\| \leq N \sqrt{N} K'(N)$.

Suppose $f$ instead satisfies $(\text{AD}_\delta)$. Then for some $\zeta \in (0, \delta]$ there exists $p > 1$ such that for all $|j| \geq |k|$, $L_{jk} \leq Ce^{-\zeta(|j| - p|k|)}$. Consequently,

$$\|E_N\|_{L^2}^2 \leq \sum_{k=-N}^{N} \sum_{j=N+1}^{\infty} 2e^{-2\zeta(j - p|k|)} \leq \frac{4N}{\zeta^2} e^{-2\zeta(1-p)N}$$

with a comparable result for $DE_N$. Thus, there exists a function $K' \in \kappa(\text{AD}_\delta)$ such that

$$\|E_N\| \leq NK'(N) \leq N \sqrt{N} K'(N).$$

Similarly, we get the same results up to constants in the Chebyshev case: the $C$ matrices are unproblematic as $\|C\|_{L^2} = 1$ and $\|C^{-1}\|_{L^2} = 2$.

We therefore have by Lemma 3 that if $f$ satisfies some distortion condition (D) then there exists $K' \in \kappa(\text{D})$ such that for any $N$ and $\phi$ in $L_N$,

$$\|L_N \phi - \phi\|_{BV} \leq N \sqrt{N} K'(N)$$

and if $N$ is sufficiently large,

$$\|S_N \phi - S \phi\|_{BV} \leq \frac{\|S\|_{BV} N \sqrt{N} K'(N)}{1 - \|S\|_{BV} N \sqrt{N} K'(N)} \leq 2N \sqrt{N} \|S\|_{BV} K'(N) \|\phi\|_{BV},$$
which is what was required for Theorem 2.

Corollary 1 is a direct result of this convergence and Theorem 1:

**Corollary 1.** We know from Theorem 1 that
\[ \rho = Su. \]
We have also defined
\[ \rho_N = S_Nu, \]
recalling that \( u \) lies in \( E_N \). As a result, by Theorem 2,
\[ \| \rho_N - \rho \|_{BV} = \| S_Nu - Su \|_{BV} \leq N \sqrt{\tilde{N} K(N)} \| u \|_{BV}, \]
as required.

Note that here, unlike in Theorem 2, we actually have that estimates converge in norm to the true values.

Corollary 2 also follows directly from Theorems 1 and 2:

**Corollary 2.** We know from Theorem 1 that on \( V \), the space of zero integral functions, the solution operator \( S \) is identical to \( \sum_{n=0}^{\infty} L^n \). We then need only apply the second part of Theorem 2 to get the required inequality.

**Remark 6.** When a map satisfies (ADA), one might be interested in the best rate of decay one can get for \( \| \mathcal{E}_N \|_{BV} \), which controls the convergence of estimates. In the non-periodic case one can show that

\[
\lim_{N \to \infty} \frac{1}{N} \log \| \mathcal{E}_N \|_{BV} = \sup_{\iota \in I, x \in [0, 2\pi]} | \Im \nu_{\iota} (x + i\zeta) | - \zeta.
\]
The value of \( z \) where the supremum in (32) is maximised will have \( \Im \nu'_{\iota}(z) = 0 \); if this value of \( z \) varies continuously with \( \zeta \), then it will have a maximum when \( |\Re \nu'_{\iota}(z)| \) is 1 or -1. Thus, one expects the right-hand side of (32) to be maximised for

\[
\zeta = \min \{ \inf \{ \| z \| \mid z \in (\nu'_{\iota})^{-1}(\{ \pm 1 \}), \iota \in I \}, \delta \}.
\]
The result is the same in the periodic case but with \( v_{\iota} \) substituted for \( \nu_{\iota} \).

5. Numerical results

In Section 5.1 we will prove some rigorous bounds on basic statistical properties of the Lanford map using the rigorous Algorithm 1; we will then demonstrate the adaptive Algorithm 2 using the Lanford map and a non-smooth circle map, assessing the adaptive algorithm’s accuracy and the spectral method’s rate of convergence.

5.1. Rigorous bounds on statistical quantities: the Lanford map.

The Lanford map, \( f : [0, 1] \to [0, 1] \)
\[ f(x) = 2x + \frac{1}{2} x(1 - x) \mod 1 \]
is a common test case for rigorous estimation of statistical quantities of maps [11, 17, 2]. By linearly rescaling of \([0, 1]\) onto \([-1, 1]\) we can apply our spectral method to it.
The Lanford map's uniform expansion parameter is $\lambda = \frac{3}{2}$ and its distortion bound (on $[0, 1]$) is $C_1 = \frac{4}{9}$. Applying (44), we find that $\|S\|_{BV} \leq 9235$.

By considering the explicit bounds obtained in Lemma 1, we chose $\zeta = \text{cosh}^{-1} \frac{7}{4}$, as it is close to the optimal value for $\zeta$ given in Remark 6. We then used a symbolic mathematics package to show that as a result of Remark 6,

$$|L_{jk}| \leq t_j \sqrt{7 + \frac{\sqrt{33}}{2} e^{\text{cosh}^{-1}(4 - \sqrt{6})k - \text{cosh}^{-1}\frac{7}{4} j}}.$$  \hspace{1cm} (34)

To calculate an estimate of the acim of this map, we implemented Algorithm 1 with $N = 2048$. We found the truncation error was $\|\mathcal{E}_N\|_{BV} \leq 6.75 \times 10^{-133}$, and chose the floating-point precision to be 512 significand bits.

Consequently, we obtained an acim estimate $\tilde{\rho}$ with the rigorously validated error bound

$$\|\tilde{\rho} - \rho\|_{BV} \leq 6.3 \times 10^{-129}.$$  

This estimate is plotted in Figure 3. The Chebyshev coefficients of $\tilde{\rho}$ are available in [Lanford-acim.zip].

We then used this estimate to calculate the Lyapunov exponent of the Lanford map

$$L_{exp} = \int_{\Lambda} \log |f'(x)| \rho(x) dx.$$  

using Clenshaw-Curtis quadrature on $\tilde{\rho} \log |f'| = \tilde{\rho} \log(2 - 3x)$ [22]. This provided the rigorous estimate given in Theorem 5(a).

We then calculated the diffusion coefficient of the observable $\phi(x) = x^2$ by evaluating the natural finite-order approximation of formula (9), using
Clenshaw-Curtis quadrature. obtaining the rigorous bound given in Theo-
rem 5(b).

The results together were obtained in 9 hours over 15 hyper-threaded cores
of a research server running 2 E5-2667v3 CPUs with 128GB of memory. The
most time-consuming operation was inverting the solution operator inverse
matrix $\mathbf{K}^{-1}_{2048}$: this process took up 94% of the runtime, which may stem
partly from using an unoptimised routine. Once $\mathbf{K}^{-1}_{2048}$, i.e. the solu-
tion operator matrix, was supplied, all the statistical quantities were calculated
on a personal computer in seconds.

5.2. Adaptive algorithms. We now present results from the adaptive Al-
gorithm 2 and illustrate the algorithm’s convergence by comparison with a
fixed-order version of Algorithm 2.

We have implemented Algorithm 2 in Julia, an open-source dynamic sci-
entific computing language. This implementation is publically available in
the package Poltergeist [23]. Poltergeist is integrated with ApproxFun, a
comprehensive function approximation package written in Julia [19]; thus,
standard manipulations of functions and operators may readily be applied
to invariant measures, transfer operators and so on.

Using Poltergeist, we present empirical convergence results for the Lan-
ford map (for comparison with rigorous methods), and a circle map which
is $C^4$ but not analytic.

5.2.1. The Lanford map. The Lanford map experiment in Section 5.1 can
be repeated in Poltergeist in a few lines of Julia code:

```julia
using Poltergeist, ApproxFun
f_lift(x) = 5x/2 - x^2/2; d = 0..1
f = modulomap(f_lift,d);
K = SolutionInv(f);
rho = acim(K);
L_exp = lyapunov(f,rho)
sigmasq_A = birkhoffvar(K,Fun(x->x^2,d))
```

This code instantiates a `MarkovMap` object $f$ and creates a `QROperator`
object $K$, which stands in for the corresponding solution operator inverse $\mathbf{K}$
(recalling the definition of the solution operator inverse (5)). The `acim` func-
tion carries out Algorithm 2 by calling ApproxFun’s adaptive QR solver [15]
on the equation $K\rho = u$. The output is an ApproxFun `Fun` object containing
$\rho_N$, the Chebyshev coefficients of the adaptive acim estimate. The Lyapunov
exponent and diffusion coefficient are calculated using special commands de-
fined in the package that call appropriate ApproxFun integration and QR
solving routines, in the latter case via [9]. Once this the relevant functions
have compiled using Julia’s just-in-time compiler, the last five lines of the
code will run in less than 0.12 seconds on a personal computer.

By applying Algorithm 2 with fixed orders $N$, the exponential conver-
ience of $\rho_N$ with $N$ predicted in Theorem 2 was seen to hold in practice.
Indeed, only $N_{\text{opt}} = 24$ columns of the transfer operator were required for convergence using Algorithm 2 (see Figure 4).

The Algorithm 2 estimate for $\rho$ is in fact remarkably accurate: the $\ell^\infty$ error on Chebyshev coefficients is less than $8 \times 10^{-15}$ (40 times the floating point precision) and the BV error on the acim estimate is $3 \times 10^{-13}$ (around 1300 times the floating point precision). The Lyapunov exponent estimate was correct almost to within the floating point precision, with the error compared to the rigorous estimate being $2.2 \times 10^{-16}$: this level of accuracy appears fortuitous rather than representative. More realistically, the estimate for $\sigma^2_f(A)$ was accurate to about 25 times floating point precision ($1.4 \times 10^{-15}$).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure4.png}
\caption{Exponential convergence with $N$ of floating-point estimates of $\rho_N$ for the Lanford circle map. The error of the adaptive estimate for $\rho_{N_{\text{opt}}}$ from Algorithm 2 is shown as a cross for comparison.}
\end{figure}

5.2.2. A non-analytic circle map. We now consider a circle map which does not satisfy an analytic distortion condition (AD$_3$) but rather a differentiable distortion condition (DD$_3$).

Define the uniformly expanding, triple-covering circle map $g : [0, 2\pi) \to [0, 2\pi)$ via the inverse of its lift:

$v_g(x) = \frac{x}{3} + \sum_{m=0}^{\infty} 2^{-\frac{3m}{\pi}} \cos \left(2^m \left(1 - \cos \frac{x}{3}\right)\right)$.

The map $g$ is $C^{4.125-\epsilon}$ and thus satisfies distortion condition (DD$_3$) but not (DD$_4$).

We implement the acim-finding process in a similar fashion to the Lanford map, although to optimise for speed we also supply CircleMap with the derivative for the lift:
g = CircleMap(v_g,0..2pi,diff=v_g_dash,dir=Reverse)
Lg = Transfer(g);
rho_g = acim(Lg);
This routine took approximately 9 minutes to run on a personal computer
and required the evaluation of $N_{\text{opt}} = 2747$ columns of the transfer op-
erator. It produced an acim estimate (plotted in Figure 5) whose BV error
we estimate to be approximately $4.8 \times 10^{-10}$, by comparison with an estimate
obtained using high-precision floating-point arithmetic and $N = 6144$
columns.

\begin{center}
\includegraphics[width=.5\textwidth]{figure5.png}
\end{center}

\textbf{Figure 5.} Invariant measure estimate for $g$ using Algorithm 2.

The convergence of $\rho_N$ is illustrated in Figure 6. The BV error on $\rho_N$ is
estimated to be $O(N^{-2.125})$, which is better than the Theorem 2 estimate
of $O(N^{-1.5})$. We conjecture that acim estimates of $C^{r+\alpha}$ circle maps (i.e.
those satisfying “$(DD^{r-1+\alpha})$”) converge in BV as $O(N^{2-r-\alpha \log N})$.

\textbf{Remark 7.} Numerical experiments demonstrate that eigenvalues and eigen-
functions of $L_N$ converge in norm to those of $L$, as proved in the periodic
case by [2]. The observed rates of convergence are the standard spectral rates.

\section*{Acknowledgements}

The author wishes to thank Georg Gottwald for his support, advice and
comments on the manuscript. The author would also like to thank Sheehan
Olver for discussions on spectral methods and ApproxFun, and Ian
Melbourne, Viviane Baladi, Maria Jose Pacifico and Peter Koltai for their
advice, references and illuminating discussions.

The author was supported by a Research Training Program scholarship,
and by the Erwin Schrödinger Institute, Vienna during a visit in May 2016.
Figure 6. The convergence with $N$ of floating-point estimates of $\rho_N$ for $g$. The error of the adaptive estimate for $\rho_{N_{\text{opt}}}$ using Algorithm 2 is plotted with a cross. The slope of a function $K(N) = CN^{-2.125}$ is plotted with a dashed line. Error estimates are by comparison with an $N = 6144$ high-precision floating point acim estimate.

Appendix A. The relationship between uniform expansion and uniform C-expansion

In this paper we have stipulated that maps on non-periodic domains satisfy a so-called uniform C-expansion condition rather than the usual uniform expansion condition. Neither of these conditions imply the other: in fact it is not hard to construct non-pathological examples of uniformly-expanding maps which are not uniformly C-expanding (see Figure 7).

However in practice, maps in $U_{NP}^u$ are generally also in $U_{NP}$. For example, all piecewise linear maps in $U_{NP}^k$ lie in $U_{NP}$. (In particular, if $f$ is the $k$-tupling map, the uniform C-expansion parameter for $f$ is $\lambda = \sqrt{k}$. A map in $U_{NP}^k$ typically fails to be in $U_{NP}$ if its graph “aims” towards the walls of the domain. To preserve the Markov structure, such a heading must be facilitated by a kink in the map. This is illustrated in Figure 7.

However, if we now consider only maps that are Markov with bounded distortion, we find close connections between C-expansion and classical expansion. In fact, a positive lower bound on one implies a positive lower bound on the other, which may be seen by an adaptation of the proof of Theorem 6 below.

Importantly, uniformly C-expanding maps eventually become uniformly expanding under iteration and vice versa, according to the following theorem.
Figure 7. In black, an example of a map $f \in U_{NP}^u (\lambda = 0.98^{-1})$ which is not uniformly C-expanding ($\lambda \approx 0.763$). Non-C-expanding parts of $f$ marked in mid grey. In light grey, lines of unit C-expansion (i.e. curves $\psi(x)$ for which $(\cos^{-1} \circ \psi \circ \cos)' = \pm 1$).

**Theorem 6.** Suppose $f \in U_{NP}^u$ (resp. $f \in U_{NP}$). Then there exists $n_* \in \mathbb{N}$ such that $f^n \in U_{NP}$ (resp. $U_{NP}^u$) for all $n \geq n_*$. Each $f^n$ satisfies the same distortion conditions as $f$, with possibly different constants.

Thus, maps in $U_{NP}^u$ and in $U_{NP}$ have the same dynamical properties and can additionally be converted from one class to the other. We emphasise that the crucial assumption here is bounded distortion.

We now prove the results stated above, beginning with Theorem 6.

**Theorem 6.** Suppose $f \in U_{NP}^u$ with $|f'| > \lambda$ and distortion constant $C_1$. Then $f^n \in U_{NP}^u$ with $|(f^n)'| > \lambda^n$ and distortion constant bounded by $C_1 \frac{\lambda^{n-1}}{1-\lambda}$. Let us use the notation $f^n = g$ with branches $P_{\iota, \iota} \in I^n$.

Suppose $x \in P_\iota$ for some $\iota \in I^n$. Then

$$1 - |x| |sgn(xg'(x)) - g(x)| \geq \frac{v_{\iota x}(sgn(xg'(x))) - x}{|g(v_{\iota x}(sgn(xg'(x)))) - g(x)|},$$

and by the intermediate value theorem there exists $w \in P_\iota$ such that

$$\frac{v_{\iota x}(sgn(xg'(x))) - x}{|g(v_{\iota x}(sgn(xg'(x)))) - g(x)|} = \frac{1}{|g'(w)|}.$$

Using Lemma 5(a) we find that

$$\frac{1}{|g'(w)|} > e^{-2C_1} \frac{e^{-2C_1}}{|g'(x)|}.$$
and so for all $x \in \bigcup_{i \in I} \mathcal{O}_i$,

$$\sqrt{\frac{1 - x^2}{1 - (g(x))^2}} |g'(x)| \geq \sqrt{\frac{1 + |x|}{\text{sgn}(xg'(x)) + g(x)}} e^{-2C_1} |g'(x)| \geq \sqrt{\frac{1}{2} e^{-2C_1} \lambda^n} > 1$$

for $n$ sufficiently large.

The map $f^n$ is full-branch Markov with bounded distortion, by the above satisfies (CE), and by Lemma 6 satisfies (P). Consequently, $f^n \in U_{NP}$.

Now, suppose $f \in U_{NP}$ and let $n \in \mathbb{N}^+$. For the remainder of this proof, we will use subscript notation for forward iterates: $x_n = f^n(x)$. We will additionally call two points $x, y \in \Lambda$ $n$-companions if there exists a sequence $(t_j)_{j=1,...,n}$ such that $x_{j-1}, y_{j-1} \in \mathcal{O}_{t_j}$ for $j \leq n$.

Given $x \in \Lambda$, choose $y, z$ such that $x, y$ and $z$ are $n$-companions, $y_n = -1$ and $z_n = 1$. Then by the mean value theorem, there exists $w$ between $y$ and $z$ such that

$$(35) \quad |(f^n)'(w)| = \frac{|z_n - y_n|}{|z - y|} \geq \frac{2}{\pi} \dot{\lambda}^n.$$  

Now, since $w$ lies between $y$ and $z$, it is an $n$-companion of $x, y$ and $z$. We will therefore relate $|(f^n)'(w)|$ to $|(f^n)'(x)|$ using bounded distortion.

We expand their quotient out using the chain rule and rewrite:

$$\frac{|(f^n)'(x)|}{|(f^n)'(w)|} = \prod_{j=1}^n \left| \frac{f'(x_{j-1})}{f'(w_{j-1})} \right|$$

$$= \prod_{j=1}^n \left| \frac{v_j'(x_{j})^{-1}}{v_j'(w_{j})^{-1}} \right|$$

$$= e^{-\sum_{j=1}^n \left( \log |v_j'(w_{j})| - \log |v_j'(x_{j})| \right)}$$

$$\geq e^{-\sum_{j=1}^n \left( \log |v_j'(w_{j})| - \log |v_j'(x_{j})| \right)}.$$  

$$(36)$$

We then bound the summands using (DD$_1$) and the fact that $v''_j/v'_j = (\log |v'_j|)''$:

$$\left| \log |v_j'(w_{j})| - \log |v_j'(x_{j})| \right| \leq C_1 |w_j - x_j| \leq C_1 \dot{\lambda}^j \pi.$$  

The sum in (36) can thus be collapsed to give

$$\frac{|(f^n)'(x)|}{|(f^n)'(w)|} \geq e^{-\sum_{j=1}^n C_1 \dot{\lambda}^j \pi} > e^{-C_1 \pi(1 - \dot{\lambda})^{-1}}.$$  

Combining this with (35) gives us that

$$|(f^n)'(x)| \geq e^{-C_1 \pi(1 - \dot{\lambda})^{-1}} \frac{2}{\pi} \dot{\lambda}^n,$$

which implies that $f^n$ is uniformly expanding for sufficiently large $n$. Since as before $f^n$ satisfies all the non-expansion conditions to be in $U_{NP}^u$, we have that $f^n \in U_{NP}^u$.  

$\square$  

$\square$
Appendix B. Results on conditions \((DD_1)\) and \((P)\)

In this appendix, we prove some properties possessed by maps in \(U_{NP}\) used through the rest of the paper. We first give some “non-local” properties of the bounded distortion condition \((DD_1)\), and then prove that \((P)\) is preserved under iteration.

We first prove a lemma relating bounded distortion constants to bounds on derivatives of the map. The properties summarised in Lemma 5 are mostly standard, but we improve the upper bound in part (b) from the exponentially large \(e^{2C_1}\) to a computationally more useful \(1 + 2C_1\).

**Lemma 5.** Suppose \(f : [-1, 1] \to [-1, 1]\) is full-branch Markov with bounded distortion. Suppose the distortion constant of \(f\) is \(C_1\). Then for all \(i \in I\):

(a) For all \(x, w \in [-1, 1]\),
\[
 e^{-2C_1} \leq \frac{|v_i'(x)|}{|v_i'(w)|} \leq e^{2C_1};
\]
(b) For all \(x \in [-1, 1]\),
\[
 e^{-2C_1} \frac{|O_i|}{2} \leq |v_i'(x)| \leq (1 + 2C_1) \frac{|O_i|}{2}.
\]

**Proof of Lemma 5.** Part (a) is a standard result \([13, 18]\).

To prove (b), we have that as a result of the intermediate value theorem there exists some \(w \in [-1, 1]\) such that
\[
 v_i'(w) = v_i(1) - v_i(-1) = \frac{|O_i|}{2}.
\]

By part (a), \(e^{-2C_1} \leq \frac{|v_i'(x)|}{|v_i'(w)|}\). Additionally, the fundamental theorem of calculus gives that
\[
 v_i'(x) = v_i'(w) + \int_w^x v_i''(\xi)d\xi;
\]
and consequently
\[
 |v_i'(x)| \leq |v_i'(w)| + \int_w^x C_1 |v_i'(\xi)|d\xi \\
 \leq \frac{|O_i|}{2} + C_1 \int_{-1}^1 |v_i'(\xi)|d\xi = \left(\frac{1}{2} + C_1\right) |O_i|,
\]
as required.

**Remark 8.** Similarly, suppose that a map \(f \in \bar{U}_{NP}\) obeys analytic distortion condition \((AD_{\delta})\) with constant \(C_{1,\delta}\). Then for all \(x, w \in \Lambda_{\delta}\),
\[
 e^{-2C_{1,\delta} \cosh \delta} \leq \frac{|v_i'(x)|}{|v_i'(w)|} \leq e^{2C_{1,\delta} \cosh \delta}.
\]

We now prove that the partition spacing condition \((P)\) is preserved under composition. Consequently, \(U_{NP}\) and \(\bar{U}_{NP}\) are closed under composition.
Lemma 6. Suppose \( f \) and \( g \) are Markov maps on \([-1, 1]\) satisfying (7), and that in addition \( f \) has bounded distortion with parameter \( C^{(f)}_1 \) and \( g \) has uniform expansion parameter \( \lambda^{(g)} > 0 \).

Then \( g \circ f \) satisfies (7).

Proof. Let \( O_{\phi \gamma} = v_\phi^{(f)} \left( v_\gamma^{(g)}(\Lambda) \right) \) be a branch set of \( g \circ f \). Let \( p \in \partial \Lambda \), i.e. \( p = \pm 1 \).

Since \( O_{\phi \gamma} = v_\phi^{(f)} \left( O_\gamma^{(g)} \right) \), by Lemma 3(c) we have

\[
\frac{|O_{\phi \gamma}|}{|O_\gamma^{(g)}|} \leq \left( 1 + 2C^{(f)}_1 \right) \frac{|O_\phi^{(f)}|}{2},
\]

and thus a preliminary bound on our ratio of interest:

\[
\frac{|O_{\phi \gamma}|}{d(O_{\phi \gamma}, p)} \leq \left( 1 + 2C^{(f)}_1 \right) \frac{\frac{1}{2} |O_\gamma^{(g)}|}{d(O_{\phi \gamma}, p)}|O_\phi^{(f)}|.
\]

We are interested in intervals for which \( p \notin O_{\phi \gamma} \). If \( p \in O_{\phi \gamma} \), then we need \( p \in O_\phi^{(f)} \) and \( f_\phi(p) =: \tau \in O_\gamma^{(g)} \). Note that since \( p \in \partial O_\phi^{(f)} \), then \( \tau \in \partial \Lambda \). Therefore, intervals \( O_{\phi \gamma} \) which do not contain \( p \) either have \( p \notin O_{\phi}^{(f)} \) or \( \tau \notin O_\gamma^{(f)} \).

We split into cases accordingly. In the first case where \( p \notin O_{\phi}^{(f)} \), we have that since \( O_\gamma^{(g)} = v_\gamma^{(g)}([-1, 1]) \), its length must be less than \( 2/\lambda^{(g)} \). Since \( O_{\phi \gamma} \subseteq O_\phi^{(f)} \), we must have \( d(O_{\phi \gamma}, p) \geq d(O_{\phi \gamma}, p) \). Therefore from (37),

\[
\frac{|O_{\phi \gamma}|}{d(O_{\phi \gamma}, p)} \leq (1 + 2C^{(f)}_1) \frac{1}{\lambda^{(g)}} \Xi^{(f)},
\]

where we used that \( |O_\gamma^{(g)}| < |\Lambda|/\lambda^{(g)} \) from the expansion assumption.

For the second case, let \( q \in \partial \Lambda \) be such that \( v_\phi^{(f)}(q) \) lies in between \( O_{\phi \gamma} \) and \( p \) and let \( r \in \partial O_\gamma^{(f)} \) such that \( v_\phi^{(f)}(r) \) is the nearest point in \( O_{\phi \gamma} \) to \( p \) (and thus, \( q \)). Then \( d(O_{\phi \gamma}, p) \) is the length of the interval \( [v_\phi^{(f)}(r), p] \), which is bigger than the length of the interval \( [v_\phi^{(f)}(r), v_\phi^{(f)}(q)] = v_\phi^{(f)}([r, q]) \). Using Lemma 5 the length of this last interval can be bounded:

\[
\frac{|v_\phi^{(f)}([r, q])|}{|r, q|} = \int_r^q \left| v_\phi^{(f)}(x) \right| dx \geq e^{-2C^{(f)}_1} \frac{1}{2} |O_{\phi}^{(f)}|.
\]

Furthermore, the distance between \( r \) and \( q \) is precisely \( d(O_\gamma^{(g)}, q) \).

Combining these results with (37), we find that

\[
\frac{|O_{\phi \gamma}|}{d(O_{\phi \gamma}, p)} \leq \left( 1 + 2C^{(f)}_1 \right) e^{2C^{(f)}_1} \frac{O_\gamma^{(g)}}{d(O_\gamma^{(g)}, q)} \leq \left( 1 + 2C^{(f)}_1 \right) e^{2C^{(f)}_1} \Xi^{(g)}.
\]
Combining the two cases, then, we find that
\[ \Xi^{(g \circ f)} \leq \left( 1 + 2C_1^{(f)} \right) \max \left\{ e^{2C_1^{(f)}} \Xi(g), \frac{1}{\lambda(g)} \Xi(f) \right\}, \]
as required. \hfill \Box

**Appendix C. Proof of Lemma 2**

In this appendix we will prove Lemma 2, which states that standard properties of \( f \) (e.g. differentiability of the distortion) imply the properties of \( \cos^{-1} \circ f \circ \cos \) required to apply Lemma 1 in the proof of Theorem 2.

We remark that while the partition position condition (P) is crucial for the proof in general, it is not necessary if one restricts to maps that only satisfy (DD_1).

We also emphasise that this lemma gives very loose bounds for the \( \Upsilon_n \) and \( \Upsilon_1, \delta \), and that in practice one is best served by calculating these constants directly from the \( \nu_i \).

We will first state and prove two lemmas upon which Lemma 2 relies, and then prove the latter.

**Lemma 7.** Suppose the map \( f \in \bar{U}_{NP} \) is piecewise \( C^{n+1} \), choose \( \sigma \in \partial \Lambda = \{ \pm 1 \} \) and let \( \tau = v_i(\sigma) \). Define the gradient of the chord
\[ \hat{S}_{i,\sigma}(x) = \frac{\tau - v_i(x)}{\sigma - x}. \]
Then
\[ \hat{S}_{i,\sigma}^{(n)}(w) = \frac{1}{n+1} v^{(n+1)}(w) \]
for some \( w \) directly between \( \sigma \) and \( x \).

**Proof.** We can show by induction that for all \( n \geq 0 \)
\[ \hat{S}_{i,\sigma}^{(n)}(x) = n! \frac{\tau - \sum_{m=0}^{n} \frac{v_i^{(m)}(x)(\sigma - x)^m}{(\sigma - x)^{n+1}}}{(\sigma - x)^{n+1}}. \]
Since \( v_i(\sigma) = \tau \), the lemma follows by Taylor’s theorem. \hfill \Box \hfill \Box

**Lemma 8.** Suppose \( f \in \bar{U}_{NP}, i \in I \) and \( \sigma \in \partial \Lambda = \{ \pm 1 \} \) such that \( v_i(\sigma) \notin \partial \Lambda \). Let \( \tau_{i,\sigma} = \sigma \text{sgn } v'_i(0) \) and
\[ T_{i,\sigma}(x) = 1 - v_i(x)/\tau_{i,\sigma}. \]
Then for \( x \in [-1, 1] \), \( T_{i,\sigma}(x) \geq \Xi^{-1} |O_i| \). If \( f \) satisfies analytic distortion condition (AD_3) then there exists \( \zeta \in (0, \delta] \) and \( R_\zeta > 0 \) such that for \( z \in \Lambda_\zeta \)
\[ |T_{i,\sigma}(z)| \geq R_\zeta |O_i|. \]

**Proof.** Recalling that \( \tau_{i,\sigma}^2 = 1 \) we can write
\[ T_{i,\sigma}(x) = \tau_{i,\sigma}(\tau_{i,\sigma} - v_i(\sigma)) + \tau_{i,\sigma}(v_i(\sigma) - v_i(x)). \]
Since \( \tau_{i,\sigma} - v_i(\sigma) \) has the same sign as \( \tau_{i,\sigma} \), the first term can be written as a positive quantity \( |\tau_{i,\sigma} - v_i(\sigma)| \) which is equal to \( d(\tau_{i,\sigma}, O_i) \). By the partition spacing condition (P), we have \( d(\tau_{i,\sigma}, O_i) \geq \Xi^{-1} |O_i| \).
Furthermore, one can apply Taylor’s theorem to the second term to get
that $\tau_{i,\sigma}(v_i(\sigma) - v_i(x)) = \tau_{i,\sigma}(\sigma - x)v_i'(w)$ for some $w$ between $x$ and $\sigma$. Since $v_i'$ keeps its sign on $[-1, 1]$ and the sign of $\sigma - x$ is simply the sign of $\sigma$, the definition of $\tau_{i,\sigma}$ means that the second term is positive on $[-1, 1]$. Thus, for $x \in [-1, 1]$, $T_{i,\sigma}(x) \geq \Xi^{-1}|O_i|$.

On the analytic domain $\Lambda_\zeta$ the situation is more complicated. We write that

$$RT_{i,\sigma}(x) = d(\tau_{i,\sigma}, O_i) - \tau_{i,\sigma}(v_i(x) - v_i(\Re x))$$

and bound terms from below.

Set $C_\zeta = e^{2C_{1,\delta} \cosh \zeta}$. We have that for any point $x$ in $\Lambda_\zeta$, $|v'(x)| \leq C_\zeta |O_i|$ as a result of Remark 8. We will use this fact in the following discussion.

The Bernstein ellipse $\Lambda_\zeta$ has major axis $\cosh \zeta \cdot [-1, 1]$ and minor axis $i \sinh \zeta \cdot [-1, 1]$. As a consequence every point $w$ in $\Lambda_\zeta$ has $\Im w \leq \sinh \zeta$.

We have by Taylor’s theorem that

$$(38) \quad RT_{i,\sigma}(x) = d(\tau_{i,\sigma}, O_i) - \tau_{i,\sigma}(v_i(x) - v_i(\Re x)) - \tau_{i,\sigma} \left( v_i(\Re x) - v_i(\sigma) \right)$$

for $w$ between $x$ and $\Re x$ (i.e. in $\Lambda_\zeta$). Thus,

$$|\tau_{i,\sigma}(v_i(x) - v_i(\Re x))| \leq C_{1,\delta} C_\zeta |O_i| \cosh^2 \zeta.$$

Furthermore,

$$\tau_{i,\sigma}(v_i(\Re x) - v_i(\sigma)) = \sigma^{-1} \sgn v_i'(0)(\Re x - \sigma)v_i'(w)$$

for $w$ between $\Re x$ and $\sigma$, i.e. in $\Lambda_\zeta \cap \Re$. Since $v_i' \neq 0$ on $\Lambda_\zeta$ because of the bounded distortion condition (AD$_\zeta$), and $v_i'$ must be real on $\Lambda_\zeta \cap \Re$ as it is real on $[-1, 1]$ and analytic on the whole interval, we have $\sgn v_i'(w) = \sgn v_i'(0)$ and so

$$\tau_{i,\sigma}(v_i(\Re x) - v_i(\sigma)) = (\Re x / \sigma - 1)|v_i'(w)| \leq (\cosh \zeta - 1)|v_i'(w)|$$

$$\leq (\cosh \zeta - 1)C_\zeta |O_i|.$$

As a result we have from (38)

$$|T_{i,\sigma}(x)| \geq RT_{i,\sigma}(x) \geq \left( \Xi^{-1} - \frac{C_\zeta}{4} \right) \left( C_{1,\delta} \cosh^2 \zeta + 2 \cosh \zeta - 2 \right) |O_i|.$$

When $\zeta$ is small enough, the term multiplying $|O_i|$ is positive. □

With these lemmas in hand, we can now prove Lemma 2.

Lemma 2. We begin with the first part of part (a), bounding derivatives of the $v_i$. We will do this by first proving a formula for the derivatives of $v_i$ and then bounding terms in this formula to get overall bounds.
where in the last line we removed the $(1 - \pi)$ term by using that the last term is zero unless $v$ and $\text{sgn} v'$.

All we need to show is that

$$ \nu_i^{(n+1)}(\cos^{-1} x) = \sum_{q+r+s \leq n} a_{q,r,s,n}(x) Y_{i,1}^{q,n}(x) Y_{i,-1}^{r,n}(x) v^{(s+1)}(x), $$

where $a_{q,r,s,n}$ are polynomials in $x$ with coefficients independent of $f$, and

$$ Y_{i,\sigma}^{m,n}(x) = \begin{cases} (1 - x\sigma^{-1})^{\frac{m}{2}} (S_{i,\sigma})^{-(1/2)}(m) , & v_{i}(\sigma) \in \{-1,1\}, \\ (1 - x\sigma^{-1})^{\frac{m}{2}} (T_{i,\sigma})^{-(1/2)}(m) , & v_{i}(\sigma) \notin \{-1,1\}. \end{cases} $$

We prove this claim by induction. Suppose without loss of generality that $\text{sgn} v' = 1$.

When $n = 0$, we have that

$$ \nu_i'(\cos^{-1} x) = \sqrt{1 - x} \sqrt{1 + x} v_i'(x). $$

From (40), we find that

$$ Y_{i,\sigma}^{0,0}(x) = \frac{1 - x\sigma^{-1}}{1 - v_i(x)\sigma^{-1}}, $$

and thus (39) follows for $n = 0$.

Suppose, then, that (39) is true for some $n$. Then

$$ \nu_i^{(n+2)}(\cos^{-1}(x)) = \sqrt{1 - x^2} (\nu_i^{(n+1)} \circ \cos^{-1})(x). $$

All we need to show is that $\sqrt{1 - x\sigma^{-1}} Y_{i,\sigma}^{m,n}(x)$ and $\sqrt{1 - x\sigma^{-1}} (Y_{i,\sigma}^{m,n})'(x)$ can be written as a product of $Y_{i,\sigma}^{m,n+1}(x)$ (and for the derivative possibly also $Y_{i,\sigma}^{m,n+1}(x)$), and polynomials in $x$. In the case where $v_{i}(\sigma) \in \{-1,1\}$, we have

$$ \sqrt{1 - x\sigma^{-1}} Y_{i,\sigma}^{m,n}(x) = (1 - x\sigma^{-1})^{\frac{m+n+2}{2}} (S_{i,\sigma})^{-(1/2)}(m) $$

and

$$ \sqrt{1 - x\sigma^{-1}} (Y_{i,\sigma}^{m,n})'(x) = (1 - x\sigma^{-1})^{\frac{m+n+1}{2}} (S_{i,\sigma})^{-(1/2)}(m+1) $$

$$ - \pi_n \sigma^{-1} (1 - x\sigma^{-1})^{\frac{m+n}{2}} (S_{i,\sigma})^{-(1/2)}(m+1) $$

$$ = (1 - x\sigma^{-1})^{\pi_n} Y_{i,\sigma}^{m+1,n+1}(x) - \pi_n \sigma^{-1} (1 - x\sigma^{-1})^{\pi_n-1} Y_{i,\sigma}^{m,n+1}(x) $$

$$ = (1 - x\sigma^{-1})^{\pi_n} Y_{i,\sigma}^{m+1,n+1}(x) - \pi_n \sigma^{-1} Y_{i,\sigma}^{m,n+1}(x), $$

where in the last line we removed the $(1 - x\sigma^{-1})^{\pi_n-1}$ element from the last term by using that the last term is zero unless $\pi_n = 1$. The relation when $v_{i}(\sigma) \notin \{-1,1\}$ is clearly analogous, from which the claim falls.
We now attempt to bound the expression in (39). To bound the \( Y_{\iota,\sigma}^{m,n} \), we need to bound derivatives of \( S_{\iota,\sigma}^{-1/2} \) and \( T_{\iota,\sigma}^{-1/2} \). One may show by induction that for \( n \geq 1 \) there exist multivariate polynomials \( q_n \) such that for any function \( U \),

\[
(U^{-1/2})^{(n)} = U^{-1/2}q_n \left( \frac{U'}{U}, \ldots, \frac{U^{(n)}}{U} \right).
\]

By Lemma 7, we have that when \( v_i(\sigma) \in \{-1, 1\} \)

\[
|S_{\iota,\sigma}^{(n)}(x)| = \frac{1}{n+1}|v^{(n+1)}(w)|
\]

for some \( w \in [-1, 1] \). Using distortion bound (DD) and Lemma 5 we can bound this again to get that

\[
|S_{\iota,\sigma}^{(n)}(x)| \leq \frac{C_n e^{2C_1}}{n+1} |v'(x)|.
\]

We also have that \( |S_{\iota,\sigma}(x)| = |v'(w)| \geq e^{-2C_1} |v'(x)| \) for some \( w \in [-1, 1] \).

Substituting these bounds into (39) we find that

\[
\left( S_{\iota,\sigma}^{-1/2} \right)^{\infty} \leq e^{C_1} |v'|^{-1/2} |q_n| \left( \frac{C_1 e^{4C_1}}{2}, \ldots, \frac{C_n e^{4C_1}}{n+1} \right)
\]

when \( v_i(\sigma) \in \{-1, 1\} \).

Similarly, we have that \( |T_{\iota,\sigma}(x)| = |v^{(n)}(x)| \leq C_n |v'(x)| \) and, by Lemma 8 when \( v_i(\sigma) \notin \{-1, 1\} \) that \( |T_{\iota,\sigma}(x)| \geq (2\Xi^{-1}) |O_\iota| \geq 2\Xi^{-1} e^{-2C_1} |v'(x)| \). These bounds can be substituted into (41) similarly to give

\[
\left( T_{\iota,\sigma}^{-1/2} \right)^{\infty} \leq \sqrt{2} e^{C_1} |v'|^{-1/2} |q_n| \left( \frac{C_1 e^{2C_1}}{2}, \ldots, \frac{C_n e^{2C_1}}{2} \right)
\]

when \( v_i(\sigma) \notin \{-1, 1\} \).

Thus, there exist constants \( p^{m,n} \) depending on the distortion constants and partition spacing constant such that for all \( \iota \in I \) and \( \sigma \in \{-1, 1\} \), we have \( |Y_{\iota,\sigma}^{m,n}(x)| \leq p^{m,n} |v^{(n)}(x)|^{-1/2} \).

Returning to (39), we have that since \( |v^{(n+1)}(x)| \leq C_n |v'(x)| \),

\[
|\nu^{(n+1)}(\cos^{-1} x)| \leq \sum_{a_{q,r,s,n}(1) \neq \emptyset} p^{q,n} p^{r,n} C_n,
\]

for \( x \in [-1, 1] \), and thus \( |\nu^{(n+1)}(\theta)| \) is bounded by the same constant for \( \theta \in [0, 2\pi \beta'] \).

The proof of the first part of part (b) is essentially the same as the above with \( n = 1 \). The major difference is that we apply Remark 8 and the second bound in Lemma 8 instead of Lemma 5 and the first bound, respectively. We also use that \( \cos^{-1} \Lambda = \Lambda_{\iota,\sigma} \) so bounds on \( \nu^{(n)}(\theta) \) transfer directly to bounds on \( \nu^{(n)}(\cos^{-1}(x)) \).
The second parts of (a) and (b) are much more straightforward. In both cases we seek to bound
\begin{equation}
\frac{h_i^{(n)}}{h_i} = \frac{(v_i' \circ \cos)^{(n)}}{|v_i' \circ \cos|}
\end{equation}
on appropriate domains. The $n$th derivative of $v_i' \circ \cos$ can be written as a linear combination of $v_i^{(m+1)} \circ \cos, m \leq n$ with coefficients of trigonometric polynomials. Trigonometric polynomials are bounded on $[0, 2\pi \lambda]$ and $\Lambda_\lambda$; on these respective domains, the $|v_i^{(m+1)} \circ \cos|$ are bounded by $C_m|v_i' \circ \cos|$ and by $C_{1,\beta}|v_i' \circ \cos|$ for $m = 1$ respectively. Thus, we find that (42) are bounded by constants depending on $C_m, m \leq n - 1$, and in the analytic case on $\zeta$ (which parameterised $\tilde{\Lambda}_\zeta$) and $C_{1,\zeta}$. □ □

**Appendix D. Explicit bounds on the norm of the solution operator in BV**

In [18], explicit a priori bounds on decay of correlations were stated in the Lipschitz norm. Specifically, if a map on $[0, 1]$ has expansion coefficient $\lambda$ and (DD1) distortion constant $C_1$, then with $V$ the space of zero-integral functions on $[0, 1]$, the following bound holds:

\begin{align*}
R &= \frac{2C_1}{1 - \lambda^{-1}} \\
D &= 4e^{R(1 + R)}, \\
\xi &= \frac{1}{2}e^{-R(1 - \lambda^{-1})}, \\
\|L^n V\|_{\text{Lip}} &\leq De^{-\xi n}.
\end{align*}

In this appendix we sketch how these explicit bounds work through to bound $\|S\|_{BV}$.

Let $\text{Lip}([0, 1])$ be the space of Lipschitz functions on the interval $[0, 1]$ with the usual norm.

Suppose that $\|L^n V\|_{\text{Lip}} \leq K_n < 1/2$. Suppose that $g \in BV([0, 1]) \cap V$ with $\|g\|_{BV} = 1$. Let $\hat{g}_n$ be the piecewise linear interpolant to $g$ at the points $0, \frac{1}{n}, \frac{2}{n}, \ldots, 1$. It can be seen that $\text{Lip} \hat{g}_n \leq n$ and $\|\hat{g}_n - g\|_1 \leq \frac{1}{2n}$.

Consequently,

\begin{align*}
\|L^n g\|_1 &\leq \|L^n \hat{g}_n\|_1 + \|L^n (g - \hat{g}_n)\|_1 \\
&\leq \frac{1}{5}\|L^n \hat{g}_n\|_{\text{Lip}} + \|g - \hat{g}_n\|_1 \\
&\leq \frac{K_n}{5}(\text{Lip} \hat{g}_n + \|\hat{g}_n\|_\infty) + \|g - \hat{g}_n\|_1 \\
&\leq \frac{K_n}{5}(n + 1) + \frac{1}{2n},
\end{align*}

where we used that $\|h\|_{\text{Lip}} \geq 5\|h\|_1$ and $\|h\|_{BV} \geq \|h\|_\infty$ for $h \in V$. 

Setting \( n = \lceil K_n^{-1/2} \rceil \), we have
\[
\| L^n g \|_1 \leq \sqrt{K_n(7 + 4\sqrt{K_n})} \leq \sqrt{K_n}.
\]

Hence, as a result of the standard \( BV \) Lasota-Yorke inequality \([11]\) we find that
\[
\| L^{m+n} g \|_{BV} \leq \frac{5}{4} \| L^{m+n} g \|_{BV} \leq \frac{5}{4} (\lambda^{-m} C_1 \sqrt{K_n}).
\]

Using that \( \| L^n |_V \|_{Lip} \leq De^{-\xi n} \) from (43), and choosing
\[
n = \left\lceil \frac{4 + 2 \log(\max\{C_1, 1\} \sqrt{D})}{\xi} \right\rceil,
\]
\[
m = \left\lceil \frac{2}{\log \lambda} \right\rceil,
\]
we have
\[
\| L^n |_V \| \leq \frac{e^{-4}}{\min\{1, C_1^{-2}\} D^{-1}} =: K_n < 1/2.
\]

Consequently from (44) we have that \( \| L^{m+n} \|_{BV} \leq \frac{5}{2} e^{-2} \leq \frac{2}{5} \).

As a result,
\[
\| S \|_{BV} \leq 1 + 5\frac{3}{3}(m + n)C',
\]
where \( C' := 1 + \frac{1}{3} \frac{C_1}{1-\lambda} \geq \sup_{n \in \mathbb{N}} \| L^n \|_{BV} \leq\) This bounding property of \( C' \) is the result of the Lasota-Yorke inequality and the fact that \( \| g \|_{BV} \geq 3 \| g \|_1 \) for \( g \in BV \cap V \).

As a result of (44), we finally obtain the a priori bound on the solution operator
\[
\| S \|_{BV} \leq 1 + \frac{5}{3}(m + n)C'(3 + C').
\]

**References**

[1] J. L. Aurentz and L. N. Trefethen, *Chopping a Chebyshev series*, ACM Transactions on Mathematical Software (TOMS), 43 (2017), p. 33.

[2] W. Bahsoun, S. Galatolo, I. Nisoli, and X. Niu, *Rigorous approximation of diffusion coefficients for expanding maps*, Journal of Statistical Physics, 163 (2016), pp. 1486–1503.

[3] V. Baladi, *Positive transfer operators and decay of correlations*, vol. 16, World Scientific, Shanghai, 2000.

[4] V. Baladi and M. Holschneider, *Approximation of nonessential spectrum of transfer operators*, Nonlinearity, 12 (1999), p. 525.

[5] C. Bose and R. Murray, *The exact rate of approximation in Ulam’s
method, Discrete and continuous dynamical systems, 7 (2001), pp. 219–235.

[6] J. P. Boyd, *Chebyshev and Fourier spectral methods*, Courier Corporation, Mineola, NY, 2001.

[7] P. Collet and J.-P. Eckmann, *Concepts and results in chaotic dynamics: a short course*, Springer Science & Business Media, Berlin, 2007.

[8] W. de Melo and S. van Strien, *One-dimensional dynamics*, vol. 25, Springer Science & Business Media, Berlin, 2012.

[9] M. Dellnitz, G. Froyland, and O. Junge, *The algorithms behind GAIO — Set oriented numerical methods for dynamical systems*, in Ergodic theory, analysis, and efficient simulation of dynamical systems, Springer, 2001, pp. 145–174.

[10] G. Froyland, O. Junge, and P. Koltai, *Estimating long-term behavior of flows without trajectory integration: The infinitesimal generator approach*, SIAM Journal on Numerical Analysis, 51 (2013), pp. 223–247.

[11] S. Galatolo and I. Nisoli, *An elementary approach to rigorous approximation of invariant measures*, SIAM Journal on Applied Dynamical Systems, 13 (2014), pp. 958–985.

[12] G. A. Gottwald, J. P. Wormell, and J. Wouters, *On spurious detection of linear response and misuse of the fluctuation–dissipation theorem in finite time series*, Physica D: Nonlinear Phenomena, 331 (2016), pp. 89–101.

[13] S. Gouëzel, *Central limit theorem and stable laws for intermittent maps*, Probability Theory and Related Fields, 128 (2004), pp. 82–122.

[14] S. Gouëzel, I. Melbourne, et al., *Moment bounds and concentration inequalities for slowly mixing dynamical systems*, Electron. J. Probab, 19 (2014), pp. 1–30.

[15] A. C. Hansen, *Infinite-dimensional numerical linear algebra: theory and applications*, in Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, The Royal Society, 2010, p. rspa20090617.

[16] M. Holschneider, *Wavelet analysis of transfer operators acting on n-dimensional Hölder, Zygmund Triebel spaces*, (1996).

[17] O. Jenkinson, M. Pollicott, and P. Vytnova, *Rigorous computation of diffusion coefficients for expanding maps*, Journal of Statistical Physics, 170 (2018), pp. 221–253.

[18] A. Korepanov, Z. Kosloff, and I. Melbourne, *Explicit coupling argument for uniformly expanding maps*, arXiv preprint arXiv:1602.03795, (2016).

[19] S. Olver, *ApproxFun*. Available at [http://github.com/JuliaApproximation/ApproxFun.jl](http://github.com/JuliaApproximation/ApproxFun.jl) and in the Julia package repository.

[20] S. Olver and A. Townsend, *A fast and well-conditioned spectral
method, SIAM Review, 55 (2013), pp. 462–489.

[21] J. Revels, M. Lubin, and T. Papamarkou, Forward-mode automatic differentiation in Julia, arXiv:1607.07892 [cs.MS], (2016).

[22] L. N. Trefethen, Approximation theory and approximation practice, Siam, Philadelphia, PA, 2013.

[23] C. Wormell, Poltergeist. Available at http://github.com/wormell/Poltergeist.jl and in the Julia package repository.

[24] L.-S. Young, What Are SRB Measures, and Which Dynamical Systems Have Them?, Journal of Statistical Physics, 108 (2002), pp. 733–754.