Stability and approximation of statistical limit laws for multidimensional piecewise expanding maps

Harry Crimmins* and Gary Froyland†

School of Mathematics and Statistics
University of New South Wales
Sydney NSW 2052, Australia

March 13, 2018

Abstract

The unpredictability of chaotic nonlinear dynamics leads naturally to statistical descriptions, including probabilistic limit laws such as the central limit theorem and large deviation principle. A key tool in the Nagaev-Guivarc’h spectral method for establishing statistical limit theorems is a “twisted” transfer operator. In the abstract setting of Keller-Liverani [27] we prove that derivatives of all orders of the leading eigenvalues and eigenprojections of the twisted transfer operators with respect to the twist parameter are stable when subjected to a broad class of perturbations. As a result, we demonstrate stability of the variance in the central limit theorem and the rate function from a large deviation principle with respect to deterministic and stochastic perturbations of the dynamics and perturbations induced by numerical schemes. We apply these results to piecewise expanding maps in one and multiple dimensions, including new convergence results for Ulam projections on quasi-Hölder spaces.

1 Introduction

Transfer operators have proved to be powerful tools for the analysis of dynamical systems possessing some expanding properties. If \((X, m)\) is a probability space and \(T : X \to X\) is a non-singular transformation, one aims to select a Banach space \((B, \|\cdot\|), B \subset L^1(m)\) that is compatible with the dynamics of \(T\) in the sense that the transfer operator \(L : B \to B\) is quasi-compact. Quasi-compactness implies a variety of desirable phenomena, including a finite number of absolutely continuous invariant probability measures (ACIMs) with densities lying in \(B\), and under a mixing condition, a unique ACIM with exponential decay of correlations [21, 29].

* h.crimmins@unsw.edu.au
† g.froyland@unsw.edu.au
A natural question is how the spectral data corresponding to the isolated eigenvalues of \( \mathcal{L} \) behave when either the map \( T \) or the operator \( \mathcal{L} \) is perturbed. It is well known that many interesting perturbations are not close to \( \mathcal{L} \) in the operator norm induced by the norm of \( B \), and therefore standard operator perturbation theory, e.g. \([24]\) cannot be employed. For piecewise monotonic maps of the interval, with \( \mathcal{L} \) acting on the space of functions of bounded variation \( \text{BV} \), Keller \([25]\) introduced a “triple norm” \( \|A\| = \sup_{\text{var}(f)+|f|_{L^1} \leq 1} |A f|_{L^1} \) for \( A : \text{BV} \to \text{BV} \). For a variety of deterministic, stochastic, and numerical perturbations, the resulting perturbed transfer operator \( \mathcal{L}_\epsilon \) is close to \( \mathcal{L} \) in \( \|\cdot\| \). These results were abstracted in a seminal paper by Keller and Liverani \([27]\), where \( |\cdot|_{L^1} \) becomes a “weak” norm \( |\cdot| \), and \( \text{var}(\cdot) + |\cdot|_{L^1} \) becomes a “strong” norm \( \|\cdot\| \) on a Banach space \( B \), with the closed unit \( \|\cdot\|\)-ball compact in \( |\cdot| \). Using this abstract setup, with additional conditions on the growth of the norms of iterates of \( \mathcal{L}_\epsilon \) in the weak and strong norms, \([27]\) proved stability of the isolated spectral data.

The use of spectral theory has also been a remarkably successful strategy for establishing statistical laws for dynamical systems such as central limit theorems \([35, 10, 20, 2]\) and large deviation principles \([20, 33]\), as well as local central limit theorems \([20, 14]\), Berry-Esseen theorems \([18, 14]\), and vector-valued almost-sure invariance principles \([30, 15]\). We refer the reader to the excellent survey \([16]\) and the references therein. Assumptions concerning stability of the derivatives of spectral data of twisted quasi-compact operators, \([27]\) to twisted transfer operators, providing access to stability results for statistical laws obtained via spectral theory across a range of perturbations in a general abstract setting. For the specific problem of numerical approximation of statistical quantities, our flexible setup allows a variety of projection methods, enabling the projection to be tailored to the particular class of dynamics to achieve the most efficient numerical scheme.

We outline below our first main abstract result (see Theorem \([2.3]\) for the formal result) concerning stability of the derivatives of spectral data of twisted quasi-compact operators, taken with respect to the twist parameter \( z \).

**Theorem A**

Let \( A_0 \) be a quasi-compact operator with simple leading eigenvalue \( \lambda_0 \) and quasi-compact decomposition \( \lambda_0 \Pi_0 + N_0 \), and let \( \{A_\epsilon\}_{\epsilon \geq 0} \) be a family of perturbations of type \((KL)\). Let \( A_\epsilon(z) = A_\epsilon M(z) \) for a compactly \( |\cdot|\)-bounded twist \( M(z) \) (see Definition \([2.4]\)). For \( z \) in a sufficiently small neighbourhood of the origin and \( \epsilon \) sufficiently small one has:

1. a quasi-compact decomposition \( A_\epsilon(z) = \lambda_\epsilon(z) \Pi_\epsilon(z) + N_\epsilon(z) \), where \( \lambda_\epsilon(z) \), \( \Pi_\epsilon(z) \) and \( N_\epsilon(z) \) depend analytically on \( z \),
2. for each $n \in \mathbb{N}$ the following convergence for $n^{th}$-order derivatives with respect to $z$ as $\epsilon \to 0$:

(a) $\lambda^{(n)}(\cdot)$ converges compactly to $\lambda_0^{(n)}(\cdot),$

(b) $\Pi^{(n)}(\cdot)$ converges compactly to $\Pi_0^{(n)}(\cdot)$ in $\|\cdot\|,$

(c) $N^{(n)}(\cdot)$ converges compactly to $N_0^{(n)}(\cdot)$ in $\|\cdot\|.$

When $A_0$ is the transfer operator $L$ corresponding to a uniformly expanding, piece-wise $C^2$ map $T : X \to X,$ where $X$ may be an interval, or a more complicated higher dimensional domain as in [36], three specific examples of the types of perturbations (KL) we consider include:

NP: Numerical approximations of $L$ by some finite-rank operator, such as in Ulam’s method [28]. Given a partition of $X$ into $n$ connected elements of diameter less than $\epsilon,$ define $E_n$ to be the conditional expectation operator with respect to this partition and let $L_\epsilon = E_n \circ L.$

For untwisted operators, stability of the ACIM with respect to (NP) has been proven in one dimension [28] with $B$ the space of bounded variation $BV$ and in multiple dimensions [13, 31] with $B$ the space of generalised bounded variation. In Section 4.1 we briefly recount the fact that the family of operators $L_\epsilon$ satisfies (KL) for interval maps. In Section 5 we prove the stability of the ACIM when $B$ is the Banach algebra of quasi-Hölder functions and show that the family of operators $L_\epsilon$ satisfies (KL) for multidimensional piecewise expanding maps.

SP: Stochastic perturbations that arise from the convolution of the Perron-Frobenius operator with an appropriate bistochastic, nonnegative kernel $K_\epsilon(x,y): L_\epsilon f(x) = \int (L f)(y) K_\epsilon(y,x) \, dm(y).$

With $X = [0,1]$ and $B$ the space of functions of bounded variation, if for $C \subset X$ the measure $m_\epsilon(C) := \int_{C \times C} K_\epsilon \, dm \times m$ converges weakly to $m$ lifted to the diagonal of $X \times X$ and under mild monotonicity conditions, Corollary 17 [25], proves stability of the ACIM with respect to (SP) for piecewise expanding $T$ with $|T'| > 2.$ Under these conditions, the family of operators $L_\epsilon$ also satisfies (KL). In Section 5 we provide a new proof of stability of the ACIM when $B$ is the space of quasi-Hölder functions and show that the family of operators $L_\epsilon$ satisfies (KL).

DP: Deterministic perturbations of $T$ in an appropriate metric space. For example, in a “Skorohod”-type metric in the case of piecewise expanding maps on the interval [25, Section 3]:

$$d(T, T_\epsilon) := \inf\{\delta > 0: \exists A \subset [0,1], m(A) > 1 - \delta; \exists \text{ a diffeomorphism } h : [0,1] \to [0,1],$$

$$T|_C = T_\epsilon \circ h|_C, \text{ and } \forall x \in C, |h(x) - x| < \delta, |(1/h'(x)) - 1| < \delta\}.$$ (1)

If $\lim_{\epsilon \to 0} d(T, T_\epsilon) = 0$ then the family of operators $\{L_\epsilon\}_{\epsilon \geq 0}$ satisfies (KL) [25, §3].

Given perturbations $\{L_\epsilon\}_{\epsilon \geq 0}$ satisfying (KL), we use the spectral formulae for the variance $\sigma_g^2$ and rate function $r_g$ to define perturbed versions of these quantities. Let
\( \sigma^2_{g,\epsilon} = \lambda^{(2)}_\epsilon(0) \) be the perturbed variance and let the convex conjugate of the map \( z \mapsto \log \lambda_\epsilon(z) \) be the perturbed rate function. Below we outline our second main result (see Theorems 3.4, 3.7 and Proposition 3.9 for formal statements), which guarantees stability of the variance \( \sigma^2_g \) and the rate function \( r_g \) under the perturbations of type (KL). We emphasise that we only need to check the perturbation conditions (KL) for the \textit{untwisted} \((z = 0)\) operators.

**Theorem B**

Under the hypotheses of Theorem A and the Nagaev-Guivarc’h spectral method, with \( A_\epsilon = L_\epsilon, M(z)f = e^{zg}f, \) and \( B \) a Banach algebra, one has

1. Stability of the variance: \( \lim_{\epsilon \to 0} \lambda^{(2)}_\epsilon(0) = \sigma^2_g. \)

2. Stability of the rate function: If each \( L_\epsilon \) is positive then for every compact subset \( C \) of the domain of \( r_g \) there exists a closed interval \( V \) such that \( \lim_{\epsilon \to 0} \sup_{z \in V} (sz - \ln(\lambda_\epsilon(z))) = r_g(s) \) uniformly for \( s \in C. \)

In particular, Theorem B guarantees convergence of numerical estimates of statistical laws produced by a variety of numerical schemes. Furthermore, to our knowledge this is the first rigorous stability and approximation result for the rate function for deterministic dynamical systems. In this work we explore Ulam’s method in detail, including new results on the quasi-Hölder spaces introduced by Saussol [36], which provide a flexible setting for piecewise expanding maps on complicated (possibly fractal) domains \( X \subset \mathbb{R}^d. \) We prove that perturbations of the type (NP) and (SP) above are of type (KL) in the quasi-Hölder spaces and thus Theorems A and B apply to such perturbations. While we have focussed on Ulam’s method, our stability results can also be applied to other numerical methods of projection type (e.g. Galerkin methods projecting onto other locally or globally supported bases).

There has been recent interest in numerical approximation of the variance \( \sigma^2_g. \) Convergence of Ulam-based estimates for piecewise \( C^2 \) expanding interval maps has been rigorously implemented in one dimension [4], with a rigorous computation for the full-branched Lanford map, and extensions of the method for the intermittent Liverani-Saussol-Vaienti (LSV) map. A periodic point algorithm for real analytic expanding interval maps and real analytic observables [22] exploits the analyticity to achieve rigorous estimates of the variance with rapid convergence. For full-branched maps of the interval, Ulam’s method has been replaced with a Galerkin method that utilises a Fourier or Chebyshev basis [40] to capitalise on the smoothness of the dynamics and increase the speed of convergence. There are also results on the stability of the variance under perturbations to the underlying dynamics [5]. In this work we create a general abstract framework for obtaining stability results for the variance. Our results only depend on quasi-compactness and related properties of the Perron-Frobenius operator, and therefore cover a very wide range of perturbations and numerical schemes.

An outline of the paper is as follows. Section 2.1 reviews the stability theory of [27] for (untwisted) quasi-compact operators. Section 2.2 introduces abstract compactly-bounded twists and proves the abstract stability of derivatives of the twisted spectral data (Theorem 2.5). Section 3 reviews the Nagaev-Guivarc’h spectral method and verifies the applicability of the results of Section 2.2 in the setting of the method. Using the stability of the derivatives of twisted spectral data, we prove the stability of the variance
in Section 3.1 (Theorem 3.4), and stability of the rate function in Section 3.2 (Theorem 3.7 and Proposition 3.9). In Section 3.1 we also obtain a derivative-free expression for the approximate variance $\lambda^{(2)}(0)$ for use in computations (Proposition 3.5). In Section 4 we provide detailed numerical experiments for a piecewise expanding interval map and four different observables, illustrating the approximation of the variance and rate function, and interpreting the results in terms of the interplay of dynamics and observable. In Section 5 we treat the multidimensional piecewise expanding maps of Saussol [36] and prove that perturbations of the type (NP) and (SP) are compatible with the Banach algebra of quasi-Hölder functions, yielding stability and approximation results of statistical properties in a multidimensional setting.

2 Spectral stability

In Section 2.1 we discuss known results concerning the stability of the spectrum of quasi-compact operators and in Section 2.2 we state our main results on spectral stability of “twisted” quasi-compact operators. Let $(B, \|\cdot\|)$ be a Banach space over $\mathbb{C}$ and denote by $L(B)$ the Banach space of bounded linear operators from $B$ to itself.

**Definition 2.1** (Quasi-compactness). We say that $A \in L(B)$ is quasi-compact if there exists $A_1, A_2 \in L(B)$ such that $A = A_1 + A_2$, rank($A_1$) < $\infty$, $A_1A_2 = A_2A_1 = 0$ and $\rho(A_2) < \rho(A_1) = \rho(A)$. We say that $A$ is quasi-compact of diagonal type if $A_1$ is diagonalisable, and that $A$ is a simple quasi-compact operator if it is of diagonal type and rank($A_1$) = 1.

If $A$ is quasi-compact of diagonal type then there exists a decomposition of $A$ of the form

$$A = \sum_{i=1}^{s} \lambda_i \Pi_i + N,$$

where each $\lambda_i$ is distinct and satisfies $|\lambda_i| = \rho(A)$, each $\Pi_i$ is a finite-rank projection such that $\Pi_i \Pi_j = 0$ whenever $i \neq j$, and $N$ is a bounded operator such that $\rho(N) < \rho(A)$ and $N \Pi_i = \Pi_i N = 0$ for each $i$. We refer to such a decomposition as the quasi-compact decomposition of $A$.

2.1 Spectral stability of quasi-compact operators

We review the spectral stability theory for quasi-compact operators from [27]. Let $|\cdot|$ be another norm on $B$ such that $|\cdot| \leq \|\cdot\|$ and the closed, unit ball in $(B, |\cdot|)$ is compact in the topology induced by $|\cdot|$. We denote the norm on $L((B, |\cdot|), (B, |\cdot|))$ by $\|\cdot\|$; that is

$$\|A\| = \sup_{\|f\|=1} |Af|.$$

The relevant families of operators are those satisfying the following condition.

**Definition 2.2.** We say that a family of operators $\{A_\epsilon\}_{\epsilon \geq 0} \subseteq L(B)$ satisfies the Keller-Liverani (KL) condition if it verifies each of the following conditions.
There exists a monotone upper-semicontinuous function \( \tau : [0, \infty) \to [0, \infty) \) such that \( \tau(\epsilon) > 0 \) whenever \( \epsilon > 0 \), \( \| A_\epsilon - A_0 \| \leq \tau(\epsilon) \), and \( \tau(\epsilon) \to 0 \) as \( \epsilon \to 0 \).

There exists \( C_1, K_1 > 0 \) such that \( |A_\epsilon^n| \leq C_1 K_1^n \) for every \( \epsilon \geq 0 \) and \( n \in \mathbb{N} \).

There exists \( C_2, C_3, K_2 > 0 \) and \( \alpha \in (0, 1) \) such that
\[
\| A_\epsilon^n f \| \leq C_2 \alpha^n \| f \| + C_3 K_2^n |f|
\]
for every \( \epsilon \geq 0 \), \( f \in B \) and \( n \in \mathbb{N} \).

**Remark.** Denote the essential spectrum of \( A \) by
\[
\sigma_{\text{ess}}(A) = \{ \lambda \in \sigma(A) : \lambda \text{ is not an eigenvalue of } A \text{ with finite algebraic multiplicity} \}.
\]

If an individual operator \( A \) satisfies (KL2) and (KL3), then its essential spectral radius \( \rho_{\text{ess}}(A) \) is bounded by \( \alpha \) [19, 38]. If, in addition, the spectral radius of \( A \) is strictly greater than \( \alpha \) then \( A \) is quasi-compact.

We will now summarise the conclusions of [27], making explicit the \( \epsilon \)-uniform control on the location of the peripheral and essential spectrum. When we speak of the spectral stability of a family of operators, we mean that the following holds.

**Proposition 2.3.** Let \( \{ A_\epsilon \}_{\epsilon \geq 0} \subseteq L(B) \) satisfy (KL), where \( A_0 \) is a simple quasi-compact operator with decomposition \( A_0 = \lambda_0 \Pi_0 + N_0 \) and \( \alpha < |\lambda_0| \). For sufficiently small \( \delta > 0 \) and each \( r \in (\max\{\alpha, \rho(N_0)\}, |\lambda_0|) \) there exists \( \epsilon_{\delta,r} > 0 \) such that \( A_\epsilon \) is a simple quasi-compact operator with decomposition \( A_\epsilon = \lambda_\epsilon \Pi_\epsilon + N_\epsilon \) whenever \( \epsilon \in [0, \epsilon_{\delta,r}] \). Furthermore, for each \( \epsilon \in [0, \epsilon_{\delta,r}] \) the spectral data of \( A_\epsilon \) satisfies
\[
\lambda_\epsilon \in B_\delta(\lambda_0), \text{ and } \rho(N_\epsilon) < r,
\]
and converges as follows
\[
\lim_{\epsilon \to 0} \lambda_\epsilon = \lambda_0, \quad \lim_{\epsilon \to 0} \|\Pi_\epsilon - \Pi_0\| = 0, \quad \text{and} \quad \lim_{\epsilon \to 0} \|N_\epsilon - N_0\| = 0.
\]

**Proof.** Choose \( \delta \) to be sufficiently small so that \( B_\delta(\lambda_0) \cap (B_\epsilon(0) \cup \sigma(A_0)) = \{\lambda_0\} \). Then [27] Theorem 1 and Corollary 1] implies that there exists \( \epsilon_{\delta,r} > 0 \) such that for every \( \epsilon \in [0, \epsilon_{\delta,r}] \) the operator
\[
\Pi_\epsilon = \frac{1}{2\pi i} \int_{\partial B_\delta(\lambda_0)} (\omega - A_\epsilon)^{-1} d\omega
\]
is a well-defined element of \( L(B) \) and that \( \lim_{\epsilon \to 0} \|\Pi_\epsilon - \Pi_0\| = 0 \). After possibly shrinking \( \delta \) further in the above argument, by [27] part 3 of Corollary 1] there exists \( \epsilon_{\delta,r} \in (0, \epsilon_{\delta,r}] \) such that \( \text{rank}(\Pi_\epsilon) = \text{rank}(\Pi_0) = 1 \) for every \( \epsilon \in [0, \epsilon_{\delta,r}] \). Fixing \( \epsilon \in [0, \epsilon_{\delta,r}] \), as \( \text{rank}(\Pi_\epsilon) = 1 \) it follows that \( A_\epsilon \) has a simple eigenvalue \( \lambda_\epsilon \in B_\delta(\lambda_0) \). By reducing \( \delta \), we may ensure that \( \sup_{\epsilon \in [0, \epsilon_{\delta,r}]} |\lambda_\epsilon - \lambda_0| \) is arbitrarily small, and so \( \lim_{\epsilon \to 0} \lambda_\epsilon = \lambda_0 \). As \( r > \rho(N_0) \) we have \( \sigma(A_0) \cap \{\omega \in \mathbb{C} : |\omega| = r\} = \emptyset \), so [27] Corollary 2] implies that the operator
\[
\Gamma_\epsilon = \frac{1}{2\pi i} \int_{\partial B_r(0)} (\omega - A_\epsilon)^{-1} d\omega
\]
is a well-defined element of \( L(B) \) and that \( \lim_{\epsilon \to 0} \| \Gamma_\epsilon - \Gamma_0 \| = 0 \). Let \( N_\epsilon = A_\epsilon \Gamma_\epsilon \). We have
\[
\| N_\epsilon - N_0 \| = \| A_\epsilon \Gamma_\epsilon - A_\epsilon \Gamma_0 + A_\epsilon \Gamma_0 - A_0 \Gamma_0 \| \\
\leq |A_\epsilon| \| \Gamma_\epsilon - \Gamma_0 \| + \| A_\epsilon - A_0 \| \| \Gamma_0 \| ,
\]
so, after recalling (KL1) and (KL2) we have \( \lim_{\epsilon \to 0} \| N_\epsilon - N_0 \| = 0 \). By the definition of \( \Gamma_\epsilon \) we have \( \rho(N_\epsilon) < r \). By [27, Theorem 1 and (10)], we have \( \sigma(A_\epsilon) \subseteq B_r(0) \cup B_\delta(\lambda_0) \). Thus \( A_\epsilon = A_\epsilon \Pi_\epsilon + A_\epsilon \Gamma_\epsilon = \lambda_\epsilon \Pi_\epsilon + N_\epsilon \), and so \( A_\epsilon \) is quasi-compact with the claimed quasicompact decomposition.

### 2.2 Spectral stability of twisted quasi-compact operators

We begin by defining an abstract twist.

**Definition 2.4** (Twist). If \( M : D \to L(B) \) is analytic on an open neighbourhood \( D \subseteq \mathbb{C} \) of \( 0 \) and \( M(0) \) is the identity, then we call \( M \) a twist. If \( A \in L(B) \) then the operators \( A(z) := AM(z) \) are said to be twisted by \( M \). We say that \( M \) is compactly \( |\cdot| \)-bounded if for every compact \( V \subseteq D \) there exists \( C_M(V) > 0 \) such that
\[
\sup_{z \in V} |M(z)| \leq C_M(V).
\]

For each \( r > 0 \) let \( D_r = \{ z \in \mathbb{C} : |z| < r \} \). Our first main result roughly speaking, says that one can ‘uniformly extend’ the spectral stability of a family of operators satisfying (KL) to the corresponding twisted family of operators in some neighbourhood of \( 0 \), provided the twist is compatible with \( |\cdot| \).

**Theorem 2.5.** Let \( \{ A_\epsilon \}_{\epsilon \geq 0} \) satisfy (KL), where \( A_0 \) is a simple quasi-compact operator with leading eigenvalue \( \lambda_0 \) satisfying \( \alpha < |\lambda_0| \), and let \( M : D \to \mathbb{C} \) be a compactly \( |\cdot| \)-bounded twist. Then there exists \( \theta > 0 \) such that for every compact \( V \subseteq D_\theta \) there exists \( \epsilon_V > 0 \) and, for each \( \epsilon \in [0, \epsilon_V] \), analytic functions \( \lambda_\epsilon(\cdot) : V \to \mathbb{C} \), \( \Pi_\epsilon(\cdot) : V \to L(B) \), and \( N_\epsilon(\cdot) : V \to L(B) \) such that \( A_\epsilon(z) \) is a simple quasi-compact operator with decomposition \( A_\epsilon(z) = \lambda_\epsilon(z) \Pi_\epsilon(z) + N_\epsilon(z) \) whenever \( z \in V \). Additionally, for each \( n \in \mathbb{N} \) we have the following convergence as \( \epsilon \to 0 \) on \( V \):

1. \( \lambda_\epsilon^{(n)}(\cdot) \) converges uniformly to \( \lambda_0^{(n)}(\cdot) \).
2. \( \Pi_\epsilon^{(n)}(\cdot) \) converges uniformly to \( \Pi_0^{(n)}(\cdot) \) in \( \|\cdot\| \).
3. \( N_\epsilon^{(n)}(\cdot) \) converges uniformly to \( N_0^{(n)}(\cdot) \) in \( \|\cdot\| \).

**Remark.** When we say a map is analytic on an arbitrary compact subset \( V \) of \( \mathbb{C} \), we mean that it may be extended to an analytic map on some larger open subset of \( \mathbb{C} \).

**Remark.** We have only considered the case of stability of a leading simple ‘twisted’ eigenvalue, as it is a substantial simplification and all that we require in applications. We anticipate that analogous results would hold, mutatis mutandis, for any eigenvalue outside the essential spectrum of finite algebraic multiplicity.
Let us describe the strategy for proving Theorem 2.5. Firstly, using the fact that \( \{A_\epsilon\}_{\epsilon \geq 0} \) satisfies (KL), we show that there exists \( \psi > 0 \) such that \( \{A_\epsilon(z)\}_{\epsilon \geq 0} \) satisfies (KL) uniformly in \( z \) on compact subsets of \( D_\psi \). In our setting standard arguments \([35, 32, 24, 20]\) imply that \( A_\theta(z) \) is a simple quasi-compact operator on some \( D_\theta \), where we may also assume that \( \theta \in (0, \psi) \). Using a technical lemma concerning the boundedness of the resolvents of \( A_\theta(z) \) on compact subsets of \( D_\theta \), we then apply theory of \([27]\) to obtain a uniform version of Proposition 2.3 for the family of operators \( \{A_\epsilon(z)\}_{\epsilon \geq 0} \) when \( z \in V \). Theorem 2.5 immediately follows.

2.2.1 Proof of Theorem 2.5

Step 1: Verification of (KL1)

Lemma 2.6. If \( \{A_\epsilon\}_{\epsilon \geq 0} \) satisfies (KL), \( M : D \to L(B) \) is a twist and \( V \subseteq D \) is compact, then there exists an upper-semicontinuous function \( \tau_V : [0, \infty) \to [0, \infty) \) such that \( (KL) \) holds for \( \{A_\epsilon(z)\}_{\epsilon \geq 0} \) for every \( z \in V \) with \( \tau_V \) in place of \( \tau \).

Proof. Let \( \tau_V : [0, \infty) \to [0, \infty) \) be defined by \( \tau_V(\epsilon) = (\sup_{z \in V} ||M(z)||)\tau(\epsilon) \). Note that \( \tau_V \) is finite as \( V \) is compact and \( M \) is continuous on \( V \). For each \( \epsilon > 0 \) and \( z \in V \) the definition of \( \|\| \) implies that \( ||A_\epsilon(z) - A_0(z)|| \leq ||A_\epsilon - A_0|| ||M(z)|| \), and so using (KL1) we find that

\[
\sup_{z \in V} ||A_\epsilon(z) - A_0(z)|| \leq ||A_\epsilon - A_0|| \left( \sup_{z \in V} ||M(z)|| \right) \leq \tau_V(\epsilon),
\]

as required. \( \square \)

Step 2: Verification of (KL2)

Lemma 2.7. If \( \{A_\epsilon\}_{\epsilon \geq 0} \) satisfies (KL), \( M : D \to L(B) \) is a compactly \( \cdot \)-bounded twist, and \( V \subseteq D \) is compact, then there exists \( K_1(V) > 0 \) such that for every \( \epsilon \geq 0 \) and \( n \in \mathbb{N} \) we have

\[
\sup_{z \in V} |A_\epsilon(z)^n| \leq K_1(V)^n.
\]

In particular, \( (KL2) \) holds for \( \{A_\epsilon(z)\}_{\epsilon \geq 0} \) for every \( z \in V \).

Proof. As \( M \) is compactly \( \cdot \)-bounded there exists \( C_M(V) > 0 \) such that

\[
\sup_{z \in V} |M(z)| \leq C_M(V).
\]

Set \( K_1(V) = C_1K_1C_M(V) \). Then for each \( z \in V \), \( n \in \mathbb{N} \) and \( \epsilon \geq 0 \) we have

\[
|A_\epsilon(z)^n| \leq |A_\epsilon|^n |M(z)|^n \leq (C_1K_1)^n C_M(V)^n = K_1(V)^n.
\]

\( \square \)

Step 3: Verification of (KL3)

Lemma 2.8. Under the hypotheses of Theorem 2.5 there exists \( \psi > 0 \), \( \alpha_\psi \in (0, 1) \) and
\[ C_2(\psi), C_3(\psi), K_2(\psi) > 0 \text{ such that} \]
\[ \| A_{\varepsilon}(z)^n f \| \leq C_2(\psi)\alpha^n_{\psi} \| f \| + C_3(\psi)K_2(\psi)^n |f| \]
for every \( z \in D_{\psi}, f \in B, \) and \( n \in \mathbb{N}. \)

**Proof.** Fix \( m \) sufficiently large so that \( C_2\alpha^m < 1/2. \) Using (KL3) for \( \{A_{\varepsilon}\}_{\varepsilon \geq 0} \) yields
\[ \| A_{\varepsilon}(z)^m f \| \leq \| A_{\varepsilon}^m f \| + \| A_{\varepsilon}(z)^m - A_{\varepsilon}^m \| \| f \| \]
\[ \leq (1/2 + \| A_{\varepsilon}(z)^m - A_{\varepsilon}^m \|) \| f \| + C_3K^m |f|. \]

As
\[ A_{\varepsilon}(z)^m - A_{\varepsilon}^m = \sum_{k=0}^{m-1} A_{\varepsilon}(z)^k (A_{\varepsilon}(z) - A_{\varepsilon}) A_{\varepsilon}(z)^{m-1-k}, \]
we may apply (KL3) again to obtain
\[ \| A_{\varepsilon}(z)^m - A_{\varepsilon}^m \| \leq \sum_{k=0}^{m-1} \| A_{\varepsilon} \|^k \| M(z) - M(0) \| \| A_{\varepsilon}(z) \|^{m-1-k} \]
\[ \leq \| M(z) - M(0) \| \sum_{k=0}^{m-1} (C_2\alpha + C_3K)^m \| M(z) \|^{m-1-k}. \]

Since the right-hand side of (3) is continuous in \( z \) and vanishes at \( z = 0, \) there exists \( \psi > 0 \) such that
\[ \sup_{z \in D_{\psi}} \sup_{\varepsilon \geq 0} \| A_{\varepsilon}(z)^m - A_{\varepsilon}^m \| \leq 1/4. \]
Applying this to (2), for each \( \varepsilon \geq 0 \) and \( z \in D_{\psi} \) we have
\[ \| A_{\varepsilon}(z)^m f \| \leq (3/4) \| f \| + C_3K^m |f|. \]
We can use Lemma 2.7 to iterate this inequality, obtaining (KL3) for \( \{A_{\varepsilon}(z)^m\}_{\varepsilon \geq 0} \) with uniform coefficients. Standard arguments imply that (KL3) also holds for \( \{A_{\varepsilon}(z)\}_{\varepsilon \geq 0} \) with suitable modified coefficients.

We have now verified that under the hypotheses of Theorem 2.5, the families of operators \( \{A_{\varepsilon}(z)\}_{\varepsilon \geq 0} \) satisfy (KL) uniformly in \( z \) on every compact \( V \subseteq D_{\psi}. \)

**Step 4: Quasi-compactness of \( A_0(z). \)** The following result is standard in the theory of analytic perturbations of linear operators \[35, 32, 24, 20. \] We provide an outline of the proof in our setting.

**Lemma 2.9.** Assume the hypotheses of Theorem 2.5, and recall \( \psi \) and \( \alpha_{\psi} \) from Lemma 2.8. There exists \( \theta \in (0, \psi] \) and \( \lambda_0(\cdot) : D_\theta \to \mathbb{C}, \Pi_0(\cdot) : D_\theta \to L(B) \) and \( N_0(\cdot) : D_\theta \to L(B) \) such that for each \( z \in D_\theta \) the operator \( A_0(z) \) has quasi-compact decomposition \( A_0(z) = \lambda_0(z) \Pi_0(z) + N_0(z) \) and
\[ \max_{z \in D_\theta} \left\{ \alpha_{\psi}, \sup \rho(N_0(z)) \right\} < \inf_{z \in D_\theta} |\lambda_0(z)|. \]
Proof. As $z \mapsto A_0(z)$ is analytic and $A_0$ is quasi-compact with decomposition $A_0 = \lambda_0 \Pi_0 + N_0$, it is standard that there exists $\theta > 0$, and analytic maps $\lambda_0(z) : D_\theta \to \mathbb{C}$, $\Pi_0(z) : D_\theta \to L(B)$ and $N_0(z) : D_\theta \to L(B)$ such that $A_0(z)$ is a simple quasi-compact operator with decomposition $A_0(z) = \lambda_0(z)\Pi_0(z) + N_0(z)$ for each $z \in D_\theta$. By possibly shrinking $\theta$ we may assume that $\theta \leq \psi$. Since $\lambda_0(\cdot)$ is analytic, we may shrink $\theta$ to guarantee that $\alpha_\psi < \inf_{z \in D_\theta} |\lambda_0(z)|$. Furthermore, as the spectral radius is upper-semicontinuous as a function of the operator \cite[IV.3.1]{24}, we may shrink $\theta$ more so that $\sup_{z \in D_\theta} \rho(N_0(z)) < \inf_{z \in D_\theta} |\lambda_0(z)|$. \hfill \Box

Step 5: Uniform bounds on the norm of the resolvent of $A_0(z)$. In order to apply the theory in \cite{27} ‘uniformly’ to obtain Theorem 2.5, we need a uniform bound for the norms of the resolvents of the twisted operators $A_0(z)$. For $A \in L(B)$, $\delta > 0$ and $r > \rho_{ess}(A)$ define

$$V_{r,\delta}(A) = \{ \omega \in \mathbb{C} : |\omega| \leq r \text{ or } \text{dist}(\omega, \sigma(A)) \leq \delta\}.$$ 

Noting that $\sigma(A) \subseteq V_{r,\delta}(A)$, let

$$J_{r,\delta}(A) = \sup \{ \| (\omega - A)^{-1}\| : \omega \in \mathbb{C} \setminus V_{r,\delta}(A) \}.$$

Lemma 2.10. Assume the hypotheses of Theorem 2.5 and recall $\theta$ from Lemma 2.9. For every compact $V \subseteq D_\theta$, $\delta > 0$ and $r > \sup_{z \in D_\theta} \rho(N_0(z))$ we have

$$\sup_{z \in V} J_{r,\delta}(A_0(z)) < \infty.$$ 

Proof. For every $z \in D_\theta$ and $\omega \in \mathbb{C} \setminus \sigma(A_0(z))$ let $R(\omega, z) = (\omega - A_0(z))^{-1}$ denote the resolvent of $A_0(z)$ at $\omega$. Fix $z \in V$. Recall from Lemma 2.9 that $A_0(z)$ is a simple quasi-compact operator with decomposition $A_0(z) = \lambda_0(z)\Pi_0(z) + N_0(z)$. As $\lambda_0(z)$ is an isolated simple eigenvalue of $A_0(z)$, the partial-fraction decomposition of the resolvent \cite[III-(6.32)]{24} yields

$$R(\omega, z) = (\omega - \lambda_0(z))^{-1}\Pi_0(z) + S(\omega, z),$$

where for each $\omega \in \{\lambda_0(z)\} \cup \mathbb{C} \setminus \sigma(A_0(z))$ the operator $S(\omega, z) = \lim_{\omega' \rightarrow \omega} R(\omega', z)(\text{Id} - \Pi_0(z))$ is the reduced resolvent of $A_0(z)$ with respect to $\lambda_0(z)$ at $\omega$ \cite[III-(6.30), III-(6.31)]{24}. We have

$$\sup_{\omega \in \mathbb{C} \setminus V_{r,\delta}(A_0(z))} \| R(\omega, z) \| \leq \delta^{-1} \sup_{\omega \in \mathbb{C} \setminus V_{r,\delta}(A_0(z))} \| \Pi_0(z) \| + \sup_{\omega \in \mathbb{C} \setminus V_{r,\delta}(A_0(z))} \| S(\omega, z) \|.$$ 

Since $\Pi_0(\cdot)$ is analytic on $V$, which is compact, $\sup_{\omega \in V} \| \Pi_0(\omega) \| < \infty$. Hence, to complete the Lemma it suffices to bound $\sup_{\omega \in \mathbb{C} \setminus V_{r,\delta}(A_0(z))} \| S(\omega, z) \|$ uniformly in $z \in V$. As noted immediately after \cite[III-(6.25)]{24}, when restricted to $(\text{Id} - \Pi_0(z))B$ the operator $S(\omega, z)$ coincides with the resolvent of $A_0(z)(\text{Id} - \Pi_0(z)) = N_0(z)$ at $\omega$. Since $S(\omega, z)$ vanishes on $\Pi_0(z)B$, this implies that $S(\omega, z) = (\omega - N_0(z))^{-1}(\text{Id} - \Pi_0(z))$. As $r > \sup_{z \in D_\theta} \rho(N_0(z))$, for $\omega \in \mathbb{C} \setminus B_r(0)$ the standard Neumann series\footnote{While the formula in \cite[I-(5.10)]{24} is only given for the finite-dimensional case, it also holds in the current setting by the same arguments.} for the resolvent \cite[I-(5.10)]{24} yields

$$S(\omega, z) = \left( \sum_{k=0}^{\infty} \omega^{-k-1}N_0(z)^k \right)(\text{Id} - \Pi_0(z)).$$
By the spectral radius formula, there exists some \( H > 0 \) such that for every \( n \in \mathbb{N} \) we have \( \|N_0(z)^n\| \leq H \sup_{z \in D_\theta} \rho(N_0(z)) \). Using (6) we therefore have

\[
\sup_{\omega \in C \setminus \mathcal{V}_{\delta,r}(A_0(z))} \|S(\omega,z)\| \leq \sup_{\omega \in C \setminus B_\theta(0)} \|S(\omega,z)\|
\]

\[
\leq \sup_{\omega \in C \setminus B_\theta(0)} \left( \sum_{k=0}^{\infty} \|\omega^{-k-1} \cdot \|N_0(z)^k\| \cdot \|\text{Id} - \Pi_0(z)\| \right)
\]

\[
\leq \left( \sum_{k=0}^{\infty} Hr^{-k-1} \left( \sup_{z \in D_\theta} \rho(N_0(z)) \right)^k \right) \left( \sup_{z \in V} \|\text{Id} - \Pi_0(z)\| \right),
\]

which is finite as \( \sup_{z \in D_\theta} \rho(N_0(z)) < r \) and \( \text{Id} - \Pi_0(z) \) is analytic on \( V \). Recalling (5) and the definition of \( J_{\delta,r}(A_0(z)) \), we complete the proof. \( \Box \)

**Step 6: Combining the previous steps.** Recall \( \theta \) from Lemma 2.9 and let \( V \subseteq D_\theta \) be compact. Out of necessity, we construct \( \lambda_\gamma(z), \Pi_\gamma(z) \) and \( N_\gamma(z) \) for \( z \) in a larger compact set whose interior contains \( V \). As \( V \) is compact there exists some \( \gamma \in (0, \theta) \) such that \( V \) is contained in \( D_\gamma \). By Lemmas 2.6, 2.7 and 2.8, for each \( z \in D_\gamma \) the family of operators \( \{A_\gamma(z)\}_{\gamma \geq 0} \) satisfies (KL) with data

\[
\tau_\gamma, C_1(\gamma), C_2(\gamma), C_3(\gamma), K_1(\gamma), K_2(\gamma) \text{ and } \alpha_\gamma \in [0, 1),
\]

only depending on \( D_\gamma \).

By Lemma 2.9 there exists \( r > \max\{\alpha_\psi, \sup_{z \in D_\theta} \rho(N_0(z))\} \) and \( \delta > 0 \) such that \( r + \delta < \inf_{z \in D_\gamma} |\lambda_0(z)| \), which implies that \( B_\delta(\lambda_0(z)) \cap (B_r(0) \cup \sigma(A_0(z))) = \{\lambda_0(z)\} \) for every \( z \in D_\gamma \). Hence, by [27] Theorem 1 and the inequality (10) for each \( z \in D_\gamma \) there exists \( \epsilon_{\delta,r,z} > 0 \) such that for \( \epsilon \in [0, \epsilon_{\delta,r,z}] \) and \( \omega \in \partial B_{\delta}(\lambda_0(z)) \) the operator \( (\omega - A_\gamma(z))^{-1} \) is bounded and so the spectral projection

\[
\Pi_\epsilon(z) = \frac{1}{2\pi i} \int_{\partial B_\delta(\lambda_0(z))} (\omega - A_\gamma(z))^{-1} d\omega
\]

is a well-defined element of \( L(B) \). From the definitions of \( \epsilon_0 \) and \( \epsilon_1 \) in the proof of [27] Corollary 1, and the definition [27] (13)] we see that \( \epsilon_{\delta,r,z} \) may be chosen independently of \( z \in D_\gamma \) as (KL) is satisfied for each \( \{A_\gamma(z)\}_{\gamma \geq 0} \) with data (7) independent of \( z \in D_\gamma \) and as

\[
\sup_{z \in D_\gamma} J_{\delta,r}(A_0(z)) < \infty
\]

by Lemma 2.10. The same argument applied to [27] part (3) of Corollary 1 implies that there exists \( \epsilon_{\delta,z} \in (0, \epsilon_{\delta,r,z}] \) and \( \gamma_{\delta} > 0 \) such that if \( \gamma \in (0, \gamma_{\delta}) \) then \( \text{rank}(\Pi_\gamma(z)) = \text{rank}(\Pi_0(z)) = 1 \) for every \( \epsilon \in [0, \epsilon_{\delta,z}] \) and \( z \in \overline{D_\gamma} \). Since \( \text{rank}(\Pi_\gamma(z)) = 1 \), each \( A_\gamma(z) \) has a simple eigenvalue \( \lambda_\gamma(z) \in B_\delta(\lambda_0(z)) \). By [27] (10) we have \( \sigma(A_\gamma(z)) \subseteq B_{\delta}(\lambda_0(z)) \cup B_r(0) \). Since \( \sigma(A_\gamma(z)) \cap B_{\delta}(\lambda_0(z)) = \{\lambda_\gamma(z)\} \), it follows that \( \sigma(A_\gamma(z)) \setminus \{\lambda_\gamma(z)\} \subseteq B_r(0) \). Hence,

\[
\text{Id} - \Pi_\epsilon(z) = \frac{1}{2\pi i} \int_{\partial B_r(0)} (\omega - A_\gamma(z))^{-1} d\omega.
\]
Defining \( N_\epsilon(z) = A_\epsilon(z)(\text{Id} - \Pi_\epsilon(z)) \), for every \( z \in \bar{D}_\gamma \) and \( \epsilon \in [0, \epsilon_V] \) we therefore have that \( A_\epsilon(z) \) is a simple quasi-compact operator with decomposition \( A_\epsilon(z) = \lambda_\epsilon(z)\Pi_\epsilon(z) + N_\epsilon(z) \).

We will now show that for each \( \epsilon \in [0, \epsilon_V] \) the maps \( z \mapsto \lambda_\epsilon(z) \), \( z \mapsto \Pi_\epsilon(z) \), and \( z \mapsto N_\epsilon(z) \) are analytic on \( D_\gamma \). As the contour in the integral in (10) is fixed, it is well-known that \( z \mapsto \Pi_\epsilon(z) \) is analytic on \( D_\gamma \) [24, VIII.1.3 Theorem 1.7]. Hence \( z \mapsto A_\epsilon(z) \) is analytic and by the definition of \( N_\epsilon(z) \), the map \( z \mapsto N_\epsilon(z) \) is analytic on \( D_\gamma \). Since \( \lambda_\epsilon(z) \) has algebraic multiplicity 1, by the discussion in [24, II.1.8], the map \( z \mapsto \lambda_\epsilon(z) \) is analytic on \( D_\gamma \). Hence the maps \( \lambda_\epsilon(\cdot) \), \( \Pi_\epsilon(\cdot) \) and \( N_\epsilon(\cdot) \) are analytic on \( V \) as they may be extended to analytic maps on an open subset of \( \mathbb{C} \) that contains \( V \), namely \( D_\gamma \).

We now confirm that the required convergence of spectral data holds as \( \epsilon \to 0 \). By [27, Corollary 1], there exists \( H_{\delta,r,z} > 0 \) and \( \eta > 0 \) such that \( \|\Pi_\epsilon(z) - \Pi_0(z)\| \leq H_{\delta,r,z}\tau_V^\eta(\epsilon)^\eta \) for every \( \epsilon \in [0, \epsilon_V] \) and \( z \in \bar{D}_\gamma \). Recalling the bound \( (5) \) and that \( (7) \) is independent of \( \bar{D}_\gamma \), we conclude from the proof of [27, Corollary 1] that \( H_{\delta,r,z} \) can be chosen independently of \( z \in \bar{D}_\gamma \). In particular \( \lim_{\epsilon \to 0} \sup_{z \in \bar{D}_\gamma} \|\Pi_\epsilon(z) - \Pi_0(z)\| = 0 \). As \( \lambda_\epsilon(z) \in B_\delta(\lambda_0(z)) \) and \( \delta \) can be made arbitrarily small, we have \( \lim_{\epsilon \to 0} \sup_{z \in \bar{D}_\gamma} |\lambda_\epsilon(z) - \lambda_0(z)| = 0 \). Examining the proof of [27, Corollary 2] and using the same arguments as before, we have \( \lim_{\epsilon \to 0} \sup_{z \in \bar{D}_\gamma} \|N_\epsilon(z) - N_0(z)\| = 0 \). Since \( V \subseteq \bar{D}_\gamma \), the required uniform convergence on \( V \) holds. By a standard application of Cauchy’s integral formula, it follows that the derivatives of all orders of \( \lambda_\epsilon(\cdot) \), \( \Pi_\epsilon(\cdot) \) and \( N_\epsilon(\cdot) \) converge uniformly on every compact subset of \( D_\gamma \). In particular, they converge on \( V \). This concludes the proof of Theorem 25.

Remark. Note that \( 0 < \inf_{z \in \bar{V}, \epsilon \in [0, \epsilon_V]} |\lambda_\epsilon(z)| \). This bound will be important when defining approximate rate functions in Section 3.2.

3 Stability of statistical limit laws

Let \((X, m)\) be a probability space, \(T : X \to X\) a non-singular transformation, and \(\mathcal{L}\) the Perron-Frobenius operator associated with \(T\). For a comprehensive overview of the Nagaev-Guivarc’h spectral method for obtaining statistical laws for dynamical systems we refer the reader to [16], while for a broader description of both the method and its history we refer to [20, II, 35, 18, 32]. In Section 3.1 we show that if \(T\) satisfies a central limit theorem (CLT) due to the Nagaev-Guivarc’h method then the variance of the CLT is stable with respect to perturbations of \(\mathcal{L}\) satisfying (KL), and in Section 3.2 we obtain the analogous result for the stability of the rate function when \(T\) satisfies a large deviation principle (LDP). We now detail a version of the method [20, 1], incorporating some conditions from [27], which we require for our stability results.

Definition 3.1. A Banach space \((B, \|\cdot\|)\) contained in \(L^1(m)\) satisfies condition \((S)\) if each of the following conditions holds.

\[(S1) \ |\cdot|_{L^1} \leq \|\cdot\|, \text{ where } |\cdot|_{L^1} \text{ is the } L^1(m) \text{ norm.}\]

\[(S2) \ The \ closed \ unit \ ball \ of \ (B, \|\cdot\|) \ is \ compact \ in \ (L^1(m), |\cdot|_{L^1}).\]

\[(S3) \ The \ constant \ functions \ are \ in \ B.\]
(S4) $B$ is a Banach algebra i.e. there exists $C_B > 0$ such that $\|fg\| \leq C_B \|f\| \|g\|$ for all $f, g \in B$.

(S5) $B$ is a vector sublattice of $L^1(m)$ i.e. $\overline{f}, |f| \in B$ for every $f \in B$.

Under condition (S) and some conditions on the Perron-Frobenius operator, such as quasi-compactness, the Nagaev-Guivarc’h method yields a CLT and LDP for $T$.

**Theorem 3.2** ([1, Theorems 1 and 2]). For each $n \in \mathbb{Z}^+$ and $g \in B$ we denote the partial Birkhoff sums $\sum_{k=0}^{n-1} g \circ T^k$ by $S_n(g)$. If $(B, \|\|)$ and the Perron-Frobenius operator $\mathcal{L} : B \to B$ is simple and quasi-compact with $\rho(\mathcal{L}) = 1$, then there exists a unique $T$-invariant measure $\mu$ such that $d\mu/dm \in B$ and for any real-valued $g \in B \cap L^\infty(m)$ satisfying $\int g \, d\mu = 0$ we have

1. The variance $\lim_{n \to \infty} \frac{1}{n} \int \frac{S_n(g)^2}{n} \, d\mu =: \sigma_g^2$ exists.

2. The sequence of random variables $\frac{S_n(g)}{\sqrt{n}}$ converges in distribution as $n \to \infty$ to a $N(0, \sigma_g^2)$ random variable in the probability space $(X, \nu)$ for every probability measure $\nu$ such that $d\nu/dm \in B$.

3. If $\sigma_g^2 > 0$, there exists an open interval $I \subseteq \mathbb{R}$ containing 0 and a rate function $r_g : I \to \mathbb{R}$, which is non-negative, continuous, strictly convex, and with unique minimum $r_g(0) = 0$, such that

$$\lim_{n \to \infty} \frac{1}{n} \ln \nu(S_n(g) > ns) = -r_g(s)$$

for every $s \in I \cap (0, \infty)$ and probability measure $\nu$ such that $d\nu/dm \in B$.

Before stating our stability results we describe how the variance and rate function are determined. Assume the setting of Theorem 3.2. Let $M_g : \mathbb{C} \to L(B)$ be defined by $M_g(z)(f) = e^{zg}f$, where $e^{zg}$ is defined by the usual power series. From this power-series representation it follows that $M_g$ is a $|\cdot|_{L^1}$-bounded twist. The twisted Perron-Frobenius operator is then defined by $\mathcal{L}(z) := \mathcal{L}M_g(z)$. As $\mathcal{L}$ is a simple quasi-compact operator, Lemma 2.9 implies that $\mathcal{L}(z)$ has quasi-compact decomposition

$$\mathcal{L}(z) = \lambda(z)\Pi(z) + N(z),$$

where $\lambda(\cdot), \Pi(\cdot)$ and $N(\cdot)$ are analytic functions on some neighbourhood of 0, say $D_\theta$.

The variance of the CLT [20, Lemma IV.3] is given by

$$\sigma_g^2 = \lambda^{(2)}(0), \quad (11)$$

and the rate function of the LDP [20, Proposition VIII.3] is

$$r_g(s) = \sup_{z \in (-\theta, \theta)} (sz - \ln(\lambda(z))), \quad (12)$$

or, equivalently, the convex conjugate (Definition 3.6) of the function $\Lambda : (-\theta, \theta) \to \mathbb{R}$ defined by $\Lambda(z) = \log(\lambda(z))$. As both the variance and the rate function depend on the derivatives of the spectral data of the twisted Perron-Frobenius operator, we can use Theorem 2.5 to prove their stability.
Corollary 3.3. Let \( (B, \|\cdot\|) \), \( \mathcal{L} \), and \( g \) satisfy the hypotheses of Theorem 3.2, and define the \( |\cdot|_{L_1} \)-compactly bounded twist \( M_g: \mathbb{C} \to L(B) \) by \( M_g(z)(f) = e^{gz}f \). Let \( \{\mathcal{L}_\epsilon\}_{\epsilon \geq 0} \subseteq L(B) \) be a family of operators satisfying (KL), with \( \mathcal{L}_0 = \mathcal{L} \). There exists \( \theta > 0 \), for each compact \( V \subseteq D_\theta \), an \( \epsilon_V > 0 \) and, for each \( \epsilon \in (0, \epsilon_V) \), analytic maps \( \lambda_\epsilon : V \to \mathbb{C} \), \( \Pi_\epsilon : V \to L(B) \), and \( N_\epsilon : V \to L(B) \), such that for each \( z \in V \) the operator \( \mathcal{L}_\epsilon(z) \) is quasi-compact with decomposition \( \mathcal{L}_\epsilon(z) = \lambda_\epsilon(z)\Pi_\epsilon(z) + N_\epsilon(z) \). Moreover, as \( \epsilon \to 0 \) the spectral data \( \lambda_\epsilon(\cdot), \Pi_\epsilon(\cdot) \) and \( N_\epsilon(\cdot) \) converge in the manner described in Theorem 2.5.

Proof. Note that condition (S) implies that \( (B, \|\cdot\|) \) satisfies the requirements on the Banach space in Section 2 with \( |\cdot| \) taken to be \( |\cdot|_{L_1} \). As noted, \( M_g \) is a compactly \( |\cdot|_{L_1} \)-bounded twist. Since \( \mathcal{L} \) is a simple quasi-compact operator with \( \rho(\mathcal{L}) = 1 \) it has leading eigenvalue 1, which is strictly greater than the constant \( \alpha \) in (KL3). Thus we may apply Theorem 2.5 which yields the required conclusion. \( \square \)

3.1 Stability and estimation of the variance

In view of (11), Corollary 3.3 immediately yields the stability of the variance.

Theorem 3.4. Under the hypotheses of Corollary 3.3, \( \lambda_\epsilon^{(2)}(0) \) is well defined for sufficiently small \( \epsilon \) and \( \lim_{\epsilon \to 0} \lambda_\epsilon^{(2)}(0) = \sigma_g^2 \).

Proof. Let \( V \) be a compact neighbourhood of 0 with non-empty interior that is contained in \( D_\theta \), where \( \theta \) is from Theorem 2.5. By Theorem 2.5 there exists \( \epsilon_V > 0 \) and, for each \( \epsilon \in (0, \epsilon_V) \), analytic maps \( \lambda_\epsilon(\cdot) : V \to \mathbb{C} \) such that \( \lambda_\epsilon(z) \) is a simple leading eigenvalue of \( \mathcal{L}_\epsilon(z) \) and so that \( \lim_{\epsilon \to 0} \lambda_\epsilon^{(2)}(0) = \lambda^{(2)}(0) = \sigma_g^2 \). \( \square \)

Using the analyticity of \( M_g(z) \), and the fact that each \( \mathcal{L}_\epsilon(z) \) is a simple quasi-compact operator, we can obtain an explicit (derivative-free) expression for \( \lambda^{(2)}(z) \). In particular, in Section 4.2 we specialise this formula to the case \( z = 0 \) to numerically estimate the ‘approximate variance’ \( \lambda_\epsilon^{(2)}(0) \) using Ulam’s method without having to approximate any derivatives of spectral data.

Proposition 3.5. Under the hypotheses of Theorem 3.3, if for any \( z \in D_\theta \) and sufficiently small \( \epsilon \geq 0 \) we choose \( v_{\epsilon,z} \in B \) and \( \varphi_{\epsilon,z} \in B^* \) so that \( \Pi_\epsilon(z)f = \varphi_{\epsilon,z}(f)v_{\epsilon,z} \), then

\[
\lambda_\epsilon^{(2)}(z) = \lambda_\epsilon(z)\varphi_{\epsilon,z}(g^2v_{\epsilon,z}) + 2\lambda_\epsilon(z)\varphi_{\epsilon,z}(g(\lambda_\epsilon(z) - \mathcal{L}_\epsilon(z)))^{-1}\mathcal{L}_\epsilon(z)(Id - \Pi_\epsilon(z))(gv_{\epsilon,z})).
\] (13)

Proof. For any \( z \in D_\theta \) there exists a compact set \( V \) such that \( z \) is in the interior of \( V \). Assume \( \epsilon \in (0, \epsilon_V) \). Differentiating the identity \( (\lambda_\epsilon(z) - \mathcal{L}_\epsilon(z))\Pi_\epsilon(z) = 0 \) once with respect to \( z \) yields

\[
\lambda_\epsilon^{(1)}(z)\Pi_\epsilon(z) = \mathcal{L}_\epsilon^{(1)}(z)\Pi_\epsilon(z) - (\lambda_\epsilon(z) - \mathcal{L}_\epsilon(z))\Pi_\epsilon^{(1)}(z),
\] (14)

while differentiating a second time yields

\[
\lambda_\epsilon^{(2)}(z)\Pi_\epsilon(z) = \mathcal{L}_\epsilon^{(2)}(z)\Pi_\epsilon(z) - 2(\lambda_\epsilon^{(1)}(z) - \mathcal{L}_\epsilon^{(1)}(z))\Pi_\epsilon^{(1)}(z) - (\lambda_\epsilon(z) - \mathcal{L}_\epsilon(z))\Pi_\epsilon^{(2)}(z).
\] (15)

As \( \Pi_\epsilon(z)(\lambda_\epsilon(z) - \mathcal{L}_\epsilon(z)) = 0 \), by applying \( \Pi_\epsilon(z) \) on the left of (14), we obtain

\[
\lambda_\epsilon^{(1)}(z)\Pi_\epsilon(z) = \Pi_\epsilon(z)\mathcal{L}_\epsilon^{(1)}(z)\Pi_\epsilon(z).
\] (16)
Similarly,
\[
\lambda_{\epsilon}^{(2)}(z)\Pi_\epsilon(z) = \Pi_\epsilon(z)\mathcal{L}_{\epsilon}^{(2)}(z)\Pi_\epsilon(z) - 2\Pi_\epsilon(z)(\lambda_{\epsilon}^{(1)}(z) - \mathcal{L}_{\epsilon}^{(1)}(z))\Pi_\epsilon^{(1)}(z).  \tag{17}
\]

As \(\lambda_{\epsilon}(z)\) is an isolated simple eigenvalue, by \[24\] II-(2.14) we have\(^2\)
\[
\Pi_{\epsilon}^{(1)}(z) = \Pi_{\epsilon}(z)\mathcal{L}_{\epsilon}^{(1)}(z)S_{\epsilon}(z) + S_{\epsilon}(z)\mathcal{L}_{\epsilon}^{(1)}(z)\Pi_{\epsilon}(z), \tag{18}
\]
where \(S_{\epsilon}(z) = (\lambda_{\epsilon}(z) - \mathcal{L}_{\epsilon}(z))^{-1}\)\((\text{Id} - \Pi_{\epsilon}(z))\). Note that \(\Pi_{\epsilon}(z)S_{\epsilon}(z) = S_{\epsilon}(z)\Pi_{\epsilon}(z) = 0\). Applying (18) to (17), we find that
\[
\lambda_{\epsilon}^{(2)}(z)\Pi_{\epsilon}(z) = \Pi_{\epsilon}(z)\mathcal{L}_{\epsilon}^{(2)}(z)\Pi_{\epsilon}(z) - 2\lambda_{\epsilon}^{(1)}(z)\Pi_{\epsilon}(z)\Pi_{\epsilon}^{(1)}(z) + 2\Pi_{\epsilon}(z)\mathcal{L}_{\epsilon}^{(1)}(z)\Pi_{\epsilon}^{(1)}(z)
\]
\[
= \Pi_{\epsilon}(z)\mathcal{L}_{\epsilon}^{(2)}(z)\Pi_{\epsilon}(z) - 2\lambda_{\epsilon}^{(1)}(z)\Pi_{\epsilon}(z)\mathcal{L}_{\epsilon}^{(1)}(z)S_{\epsilon}(z)
\]
\[
+ 2\Pi_{\epsilon}(z)\mathcal{L}_{\epsilon}^{(1)}(z)(\Pi_{\epsilon}(z)\mathcal{L}_{\epsilon}^{(1)}(z)S_{\epsilon}(z) + S_{\epsilon}(z)\mathcal{L}_{\epsilon}^{(1)}(z)\Pi_{\epsilon}(z)).
\]

Applying \(\Pi_{\epsilon}(z)\) on the right then yields
\[
\lambda_{\epsilon}^{(2)}(z)\Pi_{\epsilon}(z) = \Pi_{\epsilon}(z)\mathcal{L}_{\epsilon}^{(2)}(z)\Pi_{\epsilon}(z) + 2\Pi_{\epsilon}(z)\mathcal{L}_{\epsilon}^{(1)}(z)S_{\epsilon}(z)\mathcal{L}_{\epsilon}^{(1)}(z)\Pi_{\epsilon}(z). \tag{19}
\]

Recall that the \(\mathcal{L}_{\epsilon}(z) = \mathcal{L}_{\epsilon}M_{g}(z), \) where \(M_{g}(z)(f) = e^{zf}.\) For each \(n \in \mathbb{N} \) and \(f \in \mathbb{B}\) we therefore have
\[
\mathcal{L}_{\epsilon}^{(n)}(z)(f) = \mathcal{L}_{\epsilon}M_{g}^{(n)}(z)(f) = \mathcal{L}_{\epsilon}(g_{n}e^{zf}) = \mathcal{L}_{\epsilon}(z)(g_{n}f). \tag{20}
\]
As \(\Pi_{\epsilon}(z)(f) = \varphi_{\epsilon,z}(f)v_{\epsilon,z},\) the left and right eigenvectors of \(\mathcal{L}_{\epsilon}(z)\) for the eigenvalue \(\lambda_{\epsilon}(z)\) are \(v_{\epsilon,z}\) and \(\varphi_{\epsilon,z},\) respectively. Moreover, as \(\Pi_{\epsilon}(z)\) is a projection, we have \(\varphi_{\epsilon,z}\Pi_{\epsilon}(z) = \varphi_{\epsilon,z}\) and \(\Pi_{\epsilon}(z)v_{\epsilon,z} = v_{\epsilon,z}.\) Using (20), and then applying \(\varphi_{\epsilon,z}\) on the left and \(v_{\epsilon,z}\) on the right to (19), we obtain
\[
\lambda_{\epsilon}^{(2)}(z) = \lambda_{\epsilon}(z)\varphi_{\epsilon,z}(g_{n}v_{\epsilon,z}) + 2\lambda_{\epsilon}(z)\varphi_{\epsilon,z}(g_{n}(\lambda_{\epsilon}(z) - \mathcal{L}_{\epsilon}(z)))^{-1}(\text{Id} - \Pi_{\epsilon}(z))\mathcal{L}_{\epsilon}(z)(g_{n}v_{\epsilon,z})).
\]

We obtain the required statement upon noting that \(\text{Id} - \Pi_{\epsilon}(z)\) and \(\mathcal{L}_{\epsilon}(z)\) commute. \(\square\)

### 3.2 Stability and estimation of the rate function

We begin with the definition of the convex conjugation, which is also known as the Legendre-Fenchel transform.

**Definition 3.6.** The convex conjugate of a function \(f : I \to \mathbb{R},\) where \(I \subseteq \mathbb{R}\) is an interval, is the function \(f^{\ast} : \mathbb{R} \to \mathbb{R}\) defined by
\[
f^{\ast}(y) = \sup_{x \in I}(xy - f(x)).
\]

Recall \(\theta\) from Corollary 3.3 and let \(V\) be a closed interval contained in \((-\theta, \theta)\). For the purpose of proving stability of the rate function we assume that each \(\mathcal{L}_{\epsilon}\) is a positive operator. Fix \(\epsilon \in [0, \epsilon_{V}]\) and \(z \in V.\) As \(g\) is real valued, \(\mathcal{L}_{\epsilon}(z)\) is therefore also a positive

\(^2\)We note that the sign discrepancy between \[24\] II-(2.14) and (18) is due to an additional factor of \(-1\) in the definition of the resolvent in \[24\].
operator. It follows that \( \rho(\mathcal{L}_c(z)) = \lambda_c(z) \) is non-negative \([37\) Proposition V.4.1\]. In view of the remark at the end of Section 2, we have \( |\lambda_c(z)| > 0 \) and so \( \lambda_c(z) \) is positive.

For each \( \epsilon \in [0, \epsilon_V] \) let \( \Lambda_{\epsilon,V} : \mathbb{R} \to \mathbb{R} \) be defined by \( \Lambda_{\epsilon,V}(z) = \ln(\lambda_c(z)) \), and note that

\[
\Lambda^*_{\epsilon,V}(s) = \sup_{z \in V} (sz - \ln(\lambda_c(z))).
\]

Compare this with the rate function \( r_g \) for the LDP (Theorem \([12\]), which is defined in \([12\]) to be the convex conjugate of the function \( \ln(\lambda_0(\cdot)) : (-\theta, \theta) \to \mathbb{R} \). Our main result for this section is the following.

**Theorem 3.7.** Suppose that we are in the setting of Corollary \([3,3\] with the additional requirement that each \( \mathcal{L}_c \) is a positive operator. For every closed interval \( V \subseteq (-\theta, \theta) \) the sequence of functions \( \Lambda^*_{\epsilon,V} \) converges uniformly to \( \Lambda^*_0 \).

**Proof.** Fix \( \eta > 0 \). By Corollary \([3,3\] there exists \( \epsilon_\eta > 0 \) such that for \( \epsilon \in [0, \epsilon_\eta] \) and \( z \in V \) the functions \( \Lambda_{\epsilon,V}(z) \) and \( \Lambda_{0,V}(z) \) are well defined and satisfy \( |\Lambda_{\epsilon,V}(z) - \Lambda_{0,V}(z)| < \eta \). Let \( y \in \mathbb{R} \). Since \( V \) is closed and the function \( z \mapsto zy - \Lambda_{0,V}(z) \) is continuous on \( V \), there exists \( z_y \in V \) such that

\[
\Lambda_{0,V}(y) = z_y y - \Lambda_{0,V}(z_y).
\]

Hence,

\[
|\Lambda_{\epsilon,V}(z_y) - (z_y y - \Lambda^*_0(y))| = |\Lambda_{\epsilon,V}(z) - \Lambda_{0,V}(z)| < \eta,
\]

and so

\[
\Lambda^*_{0,V}(y) < \eta + z_y y - \Lambda_{\epsilon,V}(z_y) < \eta + \sup_{z \in V} (zy - \Lambda_{\epsilon,V}(z)) = \eta + \Lambda^*_{\epsilon,V}(y).
\]

Applying the same argument but reversing the roles of \( \epsilon \) and 0 yields

\[
|\Lambda^*_{0,V}(y) - \Lambda^*_{\epsilon,V}(y)| < \eta.
\]

We conclude by noting that \( \eta \) is independent of \( y \), and so the previous inequality holds for all \( y \in \mathbb{R} \). \( \square \)

While Theorem \([3,7\] confirms that \( \Lambda^*_{\epsilon,V} \) and \( \Lambda^*_0 \) are close, we don’t know whether \( \Lambda^*_0 \) and \( r_g \) are necessarily close. We finish this section with a Proposition that clarifies this relationship, but first we must prove that \( z \mapsto \ln(\lambda_0(z)) \) is convex on \((-\theta, \theta)\). While it is known that \( z \mapsto \ln(\lambda_0(z)) \) is convex in a small real neighbourhood of 0 \([20\), Proposition VIII.3], we are not aware of a proof for convexity on its entire domain, which in this case is \((-\theta, \theta)\).

**Lemma 3.8.** The function \( z \mapsto \ln(\lambda_0(z)) \) is convex on \((-\theta, \theta)\).

**Proof.** Adapting arguments from \([11\) Proposition 5, Proposition 6\], for \( z \in (-\theta, \theta) \) we have

\[
\ln(\lambda_0(z)) = \lim_{n \to \infty} \frac{1}{n} \ln \int e^{zS_n(g)} \, dm.
\]

(21)

Let \( a, b \in (-\theta, \theta) \) and \( t \in [0, 1] \). Then \((21)\) yields

\[
\ln(\lambda_0(at + (1 - t)b)) = \lim_{n \to \infty} \frac{1}{n} \ln \int e^{S_n(g)(at + (1 - t)b)} \, dm.
\]

16
Recalling that $e^{S_n(g)z}$ is positive for $z \in \mathbb{R}$ and applying Hölder’s inequality with conjugate exponents $1/t$ and $1/(1-t)$ yields

$$\frac{1}{n} \ln \int e^{S_n(g)(at+(1-t)b)} \, dm = \frac{1}{n} \ln \left( \int e^{S_n(g)a} \, dm \right)^{t} \left( \int e^{S_n(g)b} \, dm \right)^{1-t} \leq \frac{1}{n} \ln \left( \int e^{S_n(g)a} \, dm \right)^{t} + \frac{1-t}{n} \ln \left( \int e^{S_n(g)b} \, dm \right).$$

Letting $n \to \infty$ yields the required inequality.

**Proposition 3.9.** Suppose that we are in the setting of Corollary 3.3. If $V \subseteq (-\theta, \theta)$ is a closed interval then $r_g$ and $\Lambda_{0,V}^*$ are equal on $\Lambda_0(V)$. Consequently, for every compact set $C \subseteq \Lambda_0^*(-\theta, \theta)$ there exists a closed interval $V_C \subseteq (-\theta, \theta)$ such that $r_g$ and $\Lambda_{0,V_C}^*$ are equal on $C$.

**Proof.** For brevity let $\Lambda_0(z) = \ln(\lambda_0(z))$. By the definition of the convex conjugate we have

$$\Lambda_{0,V}^*(y) = \sup_{z \in V} (yz - \Lambda_0(z)). \tag{22}$$

Note that for each $y \in \mathbb{R}$ the function $z \mapsto yz - \Lambda_{0,V}(z)$ that is defined on $(-\theta, \theta)$ is concave by Lemma 3.8. By differentiating, we therefore see that if $y \in \Lambda_0'(V)$ then the supremum in (22) is attained by some $z$ satisfying $\Lambda_0(z) = y$. Hence, for $y \in \Lambda_0(V)$ we have

$$\Lambda_{0,V}^*(y) = y(\Lambda_0')^{-1}(y) - \Lambda_0(\Lambda_0')^{-1}(y)).$$

The same argument shows that the same formula holds for $r_g$ on $\Lambda_0'(V)$, and so the two functions are equal on $\Lambda_0'(V)$, which is a closed interval as $\Lambda_0'$ is monotonic. For the second conclusion simply take $V_C$ to be the inverse image under $\Lambda_0'$ of the convex hull of $C$, which is clearly a closed interval.

4 Stability and approximation of statistical laws for piecewise expanding interval maps

In this section we will demonstrate the utility of our theory by applying the results of Sections 2 and 3 to piecewise expanding interval maps. This setting is classical, so we refer the reader to [9, 6] for further details. We obtain stability of the variance and rate function under standard classes of perturbations, including perturbations arising from the Ulam numerical scheme [28].

Denote by $(BV, \|\cdot\|)$ the Banach space of functions of bounded variation on $[0,1]$, equipped with the norm $\|\cdot\| = \text{Var}(\cdot) + \|\cdot\|_{L^1}$. We call $T : [0,1] \to [0,1]$ a Lasota-Yorke map if there are $0 = a_0 < a_1 < \cdots < a_r = 1$ and $\gamma > 1$ such that for each $i = 1, \ldots, r$, $T$ is $C^1$ on $(a_{i-1}, a_i)$ with a $C^1$ extension to $[a_{i-1}, a_i]$, $\left| T'_{[a_{i-1}, a_i]} \right| \geq \gamma$ and $1/ \left| T'_{[a_{i-1}, a_i]} \right| \in BV$.

It is classical that BV satisfies condition (S) (Definition 3.1) and that for a topologically mixing Lasota-Yorke map $T$ the associated Perron-Frobenius operator is simple, quasi-
compact and has spectral radius 1. Thus $T$ satisfies a CLT and LDP by Theorem 3.2 for appropriate observables. In the following corollary we confirm the stability of statistical parameters in this setting, as a consequence of Theorems 3.3, 3.4, and 3.7 and Proposition 3.9.

Corollary 4.1. Let $T$ be a topologically mixing Lasota-Yorke map, $\mu$ its invariant measure, $\{L_\epsilon\}_{\epsilon \geq 0}$ a family of operators satisfying (KL) where $L_0 = L$ is the Perron-Frobenius operator associated with $T$, and $g \in BV$ be a real-valued observable such that $\int g \, d\mu = 0$. Then each of the following holds as $\epsilon \to 0$:

1. Theorem 3.4: the approximate variance $\lambda_\epsilon^{(2)}(0)$ converges to the true variance $\sigma_g^2$.

2. Theorem 3.7: if each $L_\epsilon$ is positive then for every compact subset $C$ of the domain of $r_g$ there exists a closed interval $V$ such that $\lim_{\epsilon \to 0} \sup_{z \in V} (sz - \ln(\lambda_\epsilon(z))) = r_g(s)$ uniformly for $s \in C$.

In particular, the numerical, stochastic, and deterministic perturbations (NP), (SP), and (DP) detailed in the introduction satisfy (KL) and we obtain the corresponding approximation and stability of the variance and rate function under these perturbations.

The map we consider for the numerics in the remainder of this section is the non-Markov piecewise affine map with $a = 2.1$.

$$T_a(x) = \begin{cases} 
ax, & 0 \leq x < 1/4; \\
-a(x - 1/2), & 1/4 \leq x < 1/2; \\
-a(x - 1/2) + 1, & 1/2 \leq x < 3/4; \\
a(x - 1) + 1, & 3/4 \leq x \leq 1.
\end{cases} \quad (23)$$

By standard arguments, we obtain $\alpha = 1/1.05 < 1$ in the Lasota-Yorke inequality for $L$ and thus $L$ is quasi-compact. Moreover, it is clear from the graph of $T$ (Figure 1, upper left) that forward images of any interval $I \subset [0, 1]$ eventually cover all of $[0, 1]$; thus, the eigenvalue 1 of $L$ is simple.

Note that the dynamics of $T_a$ for $a \gtrsim 2$ has infrequent transitions between the left and right halves of the unit interval; in such a situation, these sets are sometimes called almost-invariant sets [12, 11]. We select observables $g$ (taking values approximately in the range $[-1, 1]$) that emphasise this structure to varying extents, and illustrate the combined effects of the dynamics and the observable on variances and rate functions; see Figure 2 for graphs of the various $g$. For example, the Birkhoff sums of $g(x) = 1_{[0,1/2]} - 1_{(1/2,1]}$ will typically take longer to converge because of frequent long sequences of similar $g$ values (in this case either 1 or $-1$). On the other hand, the observable $g(x) = \cos(2\pi x) - 0.0614$ is not strongly correlated with the almost-invariant dynamics and one expects a more rapid convergence of Birkhoff sums. These arguments are reflected in the table of variances, Table 4.2 and the graph of rate functions, Figure 2 (right).

### 4.1 Numerical schemes

Our choice of numerical scheme is dictated by the class of map. Because we are considering general (non-Markov) Lasota-Yorke maps, the natural choice of Banach space is BV,
Figure 1: Graph of $T_\alpha$ (upper left), an approximation of the invariant density with $n = 25000$ (upper right), a zoom of an apparent “flat” section of the invariant density showing fine structure (lower left), and an approximation of the second eigenfunction with eigenvalue $\lambda_2 \approx 0.8079$ (lower right).

with the weak and strong norms being the $L^1$ and BV norms, respectively. Since the eigenfunctions of $\mathcal{L}$ can be discontinuous (see Figure 1 upper right and lower right), we use locally supported functions for our approximation space, and in particular, locally constant functions, leading to the well-known Ulam scheme [39]. If, on the other hand, we restricted ourselves to globally differentiable, full-branched maps (a smaller and better-behaved class), then it would be natural to work with $C^r$ functions and use a globally supported basis consisting of Chebyshev polynomials (if the phase space is an interval) or trigonometric polynomials / Fourier modes (if the phase space is a circle); these bases will exploit the smoothness of the map to produce commensurately more accurate estimates.

For a partition of $[0, 1]$ into subintervals $I_1, \ldots, I_n$, setting $\mathcal{B}_n = \text{span}\{1_{I_i} : 1 \leq i \leq n\}$, we define the conditional expectation operator $\mathbb{E}_n : L^1([0, 1]) \rightarrow \mathcal{B}_n$ by

$$\mathbb{E}_nf = \sum_{i=1}^{n} \int_{I_i} f \, dm \cdot 1_{I_i}.$$  \hspace{1cm} (24)

It is well known (e.g. [28]) that the matrix representation of $\mathbb{E}_n \mathcal{L}$ on $\mathcal{B}_n$ is

$$P_{ij} = \frac{m(I_i \cap T^{-1}I_j)}{m(I_j)},$$  \hspace{1cm} (25)
4.2 Estimating the variance

We specialise the formulate for $\lambda^{(2)}(z)$ in Proposition 3.5 to $\lambda^{(2)}(0)$. Evaluating (13) at $z = 0$ yields

$$\lambda^{(2)}(0) = \lambda_0(0)\varphi_{x,0}(g^2v_{x,0}) + 2\lambda_0(0)\varphi_{x,0}(g(\lambda_0(0) - \mathcal{L}_x)^{-1}\mathcal{L}_x(\text{Id} - \Pi_x)(gv_{x,0})).$$

(26)

As each $\mathcal{L}_x$ is a Markov operator, we have $\lambda_0(0) = 1$. Additionally, we may take $\varphi_{x,0}$ to be the map $f \mapsto \int f \, dm$, which implies that $\int v_{x,0} \, dm = \varphi_{x,0}(v_{x,0}) = 1$. When applied to (26) this yields

$$\lambda^{(2)}(0) = \int g^2v_{x,0} + 2g(\text{Id} - \mathcal{L}_x)^{-1}\mathcal{L}_x(\text{Id} - \Pi_x)(gv_{x,0}) \, dm.$$

We may replace $(\text{Id} - \Pi_x)(gv_{x,0})$ with $gv_{x,0}$, which is equivalent to setting $\int gv_{x,0} \, dm = 0$, to obtain

$$\lambda^{(2)}(0) = \int g^2v_{x,0} + 2g(\text{Id} - \mathcal{L}_x)^{-1}\mathcal{L}_x(gv_{x,0}) \, dm,$$

(27)

which is the approximation used in the computation of the variance. We note that for $\epsilon = 0$ the expression (27) is equal to the expression for $\lambda^{(2)}(0)$ from [20, Corollary III.11].

The MATLAB function for computing the variance is given in Appendix B.

| Ulam subintervals | Variance estimates for observable $g(x)$ |
|-------------------|------------------------------------------|
|                   | $\cos(2\pi x) - 0.0614$ | $2x - 1$ | $\sin(2\pi x)$ | $1_{[0,1/2]} - 1_{(1/2,1]}$ |
| 200               | 0.51357                  | 4.3355  | 6.5368         | 17.006  |
| 1000              | 0.50496                  | 4.2886  | 6.4959         | 16.859  |
| 5000              | 0.50430                  | 4.2871  | 6.4950         | 16.855  |
| 25000             | 0.50396                  | 4.2860  | 6.4936         | 16.851  |

Table 1: Computed variances for four different observables and at four different Ulam resolutions.

For a transformation $T : [0, 1] \to [0, 1]$ let $\mathbb{P}$ denote a row-stochastic Ulam matrix constructed on an equi-partition of $[0, 1]$ in MATLAB’s sparse format. To compute an estimate of the variance $v$ for the observable $g(x) = \sin(2\pi x)$, use:

3Strictly speaking, Ulam’s method for twisted transfer operators will involve integrals of $g$, which can be numerically evaluated. We have chosen the above midpoint approximation of $g$ for computational convenience; note that the midpoint rule is the same order of accuracy as the trapezoidal method of numerical quadrature and often slightly more accurate (errors are about a factor 1/2 smaller). We additionally computed the values in Table 1.2 with an “exact” implementation of Ulam and the errors due to the midpoint estimate of $g$ were several orders of magnitude smaller than the errors due to the overall Ulam discretisation.
obs= @(x) sin(2*pi*x);
[v,~,~,~,~] = variance(P,obs);

There have been a number of prior rigorous numerical estimates of variance for interval maps. Bahsoun et al. [4], Pollicott et al [22], and Wormell [40] develop algorithms that output an interval in which the variance is guaranteed to lie. Bahsoun et al. applies to general Lasota-Yorke maps, uses Ulam’s method, and employs a “brute force” approach of taking high powers of $L_{1/n}$ to achieve convergence. The method of [22] applies to real analytic expanding (full-branch) maps with real analytic observables, and is based on evaluations on all periodic orbits up to a certain order. Wormell [40] applies to full-branched, $C^3$ expanding maps and uses an approach most similar to ours, with computations in Chebyshev/Fourier bases. In each of these papers, an interval containing the variance of the Lanford map $T(x) = 2x + x(1-x)/2$ (mod 1) for the observable $g(x) = x^2$ is obtained. The latter two papers, exploiting the analyticity of the map $T$ and observable $g$ can achieve more accurate estimates for the same computational effort.

In comparison to [4] we can avoid raising the very sparse matrix $L_{1/n}$ to high powers (in the full-branch Lanford map studied in [4] $L_{1/12}^{112}$ is computed). We exploit the differentiability properties of the spectral data with respect to the twist parameter (which exist even for general Lasota-Yorke maps) and preserve the high degree of sparseness of $L_{1/n}$, which is quickly destroyed by taking powers. We only need to solve a single sparse linear equation to obtain an estimate for $\lambda^{(2)}(0)$, which is related to the equation solved in [40]. In comparison to [22] and [40] we can treat general Lasota-Yorke maps, via the flexible choice of a locally supported basis, however, as explained above, for smoother classes of maps as in [22, 40], one should adapt the basis accordingly as the Ulam basis will not be competitive with specialised approaches. Our variance estimates rigorously converge to the true value as $n \to \infty$; and while it is likely possible to provide an “interval of guarantee”, as in the above methods, we have not pursued this here.

4.3 Estimating the rate function

For a fixed value of $s$, we estimate $r_g(s) = -\min_z (\log \lambda(z) - zs)$ by applying MATLAB’s built-in unconstrained function minimising routine fminunc to the function $f(z) = \log \lambda(z) - sz$. We use the default quasi-newton algorithm option for fminunc (we found the trust-region algorithm used slightly more iterates) and supply an expression for the first derivative of $f(z)$ with respect to $z$, namely $\phi(z)(g\lambda(z)v(z)) - s$ (here $\phi(z)$ and $v(z)$ are the leading left and right, respectively, eigenvectors of $L(z)$); all other settings are the defaults. Each evaluation of $f(z)$ requires the computation of $\lambda(z)$ (we obtain $v(z)$ at the same time) and each evaluation of $f'(z)$ requires an additional computation of $\phi(z)$. These two eigencomputations are made by simply repeatedly iterating $v(0)$ and $\phi(0)$ with $L(z)$ and $L(z)^*$ (and normalising), respectively until the change in the estimated eigenvalue is below a tolerance (we used $5 \times 10^{-12}$). This is relatively efficient because the Ulam matrix approximation of $L(z)$ is very sparse, and we found this is also faster than using MATLAB’s built-in eigs routine to find the single leading eigenvalue. We estimate $r_g(s)$ on a grid of $s$ values (in our experiments $s$ ranges from 0 to 0.8 in steps of 0.01), stepping from one grid point to the next. We use the previous optimal $z$ as the initial seed for the quasi-newton algorithm to find the optimal $z$ for the next $s$ grid point.
and found this choice results in slightly fewer quasi-Newton steps than choosing a fixed initialisation.

For a transformation $T : [0, 1] \to [0, 1]$ let $P$ denote a row-stochastic Ulam matrix constructed on an equipartition of $[0, 1]$ in MATLAB’s sparse format. To compute estimates of the rate function $r_g(s)$, for the observable $g(x) = \sin(2\pi x)$, at $s \in [0, 0.8]$ spaced 0.01 apart, and store these estimates in a vector $r$, use:

```matlab
s = 0:.01:.8;
obsv = @(x)sin(2*pi*x);
[r, ~] = rate_function(s, P, obs);
```

The necessary MATLAB functions are given in Appendix B; to run the above code to compute these 81 values of the rate function takes approximately 1, 4, and 12 seconds for Ulam matrices of sizes 1000, 5000, and 25000, respectively. We use the same set of four observables $g$ as in the variance computations (Figure 2 left). The corresponding rate functions are shown in Figure 2 right). Note that the four observables $g$ yield rate functions of increasingly lower value (higher likelihood of large deviations occurring). This corresponds to the correlation between the value of the observable and the almost-invariant sets $[0, 1/2], [1/2, 1]$. The observable $g(x) = \cos(2\pi x) - 0.0614$ is not particularly correlated with the almost-invariant sets and thus large deviations have low probability. On the other hand, the observables $2x - 1$ and $\sin(2\pi x)$ have moderate correlation with the almost-invariant sets and large deviations have an increased probability of occurring (interestingly, there is a crossover of these two rate functions around $s = 0.8$). The observable $g(x) = 1_{[0,1/2]} - 1_{(1/2,1]}$ is very strongly correlated with the almost-invariant sets and we see a correspondingly small rate function. Figure 3 shows the decrease in errors relative to $n = 25000$ for the calculations using $n = 200, 1000$, and 5000, typically with somewhat larger errors for larger thresholds $s$, as expected. We are not aware of prior rigorous numerical methods for estimating rate functions for deterministic dynamics. Prior

![Figure 2: Graphs of four different observables (left) and corresponding rate functions (right), computed with $n = 1000$.](image)

---

4 on a 7th-generation intel core i5 processor.
work on estimating rate functions includes [23, 34], which use the Legendre transform but not the spectral approach we use here.

Finally, we note that the rate of escape from the interval $[0, 1/2]$ can be estimated via the observable $g(x) = 1_{[0,1/2]} - 1_{[1/2,1]}$ by computing the rate function for a threshold very close to 1. With $n = 1000$, and an $s$ threshold of $1 - 10^{-15}$, one obtains a rate function value of $0.04879016416945$ (in this experiment, the optimality tolerance in `rate_function.m` was decreased to $10^{-13}$). Alternatively, computing the negative logarithm of the leading eigenvalue of the transfer operator restricted to the interval $[0, 1/2]$ (this is particularly straightforward with an Ulam basis, see e.g. [3, 8]), one obtains $0.04879016416943$. Thus, the rate function calculation and the escape rate calculation are consistent up to 13 decimal places for an Ulam matrix of size $n = 1000$ (note we are not claiming accuracy of the true values up to this precision).

5 Stability of statistical laws for piecewise expanding maps in multiple dimensions

In this section we consider the class of multi-dimensional piecewise expanding maps studied in [36] by Saussol. Using the fact that the space of quasi-Hölder functions satisfies condition (S) – in particular, it is a Banach algebra – it is shown in [1] that such maps obey a CLT and LDP for quasi-Hölder observables via the Nagaev-Guivarc’h method. Using the theory developed in Section 3, we prove that the invariant density, variance and rate function of these maps are stable with respect to perturbations satisfying (KL). We then give two examples of classes of perturbations satisfying (KL) in this setting: stochastic perturbations and perturbations arising via Ulam’s method.

Let $d \geq 2$ and denote by $m$ the Lebesgue measure on $\mathbb{R}^d$. Let $\Omega \subseteq \mathbb{R}^d$ be compact and satisfy $\text{int}(\Omega) = \Omega$. Without loss of generality we may assume that $m(\Omega) = 1$. Let
$T: \Omega \rightarrow \Omega$ be a piecewise expanding map in the sense of [36]. To prevent clashes with our notation, we denote the constants $\alpha$ and $\epsilon_0$ associated with $T$ in [36] by $\beta$ and $\eta_0$, respectively. We note that $\beta \in (0, 1]$ and $\eta_0 > 0$.

We now define the space of quasi-Hölder functions, reproducing some of [1, Section 4] for the reader’s convenience. Let $A \subseteq \mathbb{R}^d$ be an arbitrary Borel set. For each $f \in L^1(\mathbb{R}^d)$ the oscillation of $f$ over $A$ is defined to be

$$\text{osc}(f, A) = \text{ess sup}_{(y_1, y_2) \in A \times A} |f(y_1) - f(y_2)|,$$

where the essential supremum is taken with respect to the product measure $m \times m$ on $A \times A$. For every $f \in L^1(\mathbb{R}^d)$ and $\eta > 0$ the function $x \mapsto \text{osc}(f, B_\eta(x))$ is well defined and lower-semicontinuous, and therefore also measurable. Let

$$|f|_\beta = \sup_{0 < \eta \leq \eta_0} \eta^{-\beta} \int_{\mathbb{R}^d} \text{osc}(f, B_\eta(x)) \, dx,$$

and is a Banach space when endowed with the norm $||| \cdot |||_\beta = |\cdot|_{L^1} + |\cdot|_\beta$ [36, 26].

**Proposition 5.1** ([1, Section 4], [36, Theorem 5.1]). The space $\mathbb{V}_\beta(\Omega)$ satisfies condition (S) and the Perron-Frobenius operator $\mathcal{L}$ associated with the map $T$ is a quasi-compact operator on $L(\mathbb{V}_\beta(\Omega))$ with $\rho(\mathcal{L}) = 1$.

Hence, if $\mathcal{L}$ a simple quasi-compact operator then $T$ obeys both a CLT and LDP by Theorem 3.2. Consequently, due to the theory developed in Section 3 we obtain the stability of the statistical parameters associated with these laws under perturbations satisfying (KL).

**Theorem 5.2.** If the Perron-Frobenius operator $\mathcal{L}$ is simple and quasi-compact, and $\{\mathcal{L}_\epsilon\}_{\epsilon \geq 0}$ satisfies (KL), with $\mathcal{L}_0 = \mathcal{L}$, then each of the following holds as $\epsilon \rightarrow 0$:

1. Proposition 2.3: the leading eigenvector $v_\epsilon$ of $\mathcal{L}_\epsilon$ converges to the unique $T$-invariant density $\frac{d\mu}{dm}$ in $|\cdot|_{L^1}$.
2. Theorem 3.4: for each real-valued $g \in \mathbb{V}_\beta(\Omega)$ such that $\int g \, d\mu = 0$ the approximate variance $\lambda_\epsilon^{(2)}(0)$ converges to the true variance $\sigma^2_g$.
3. Theorem 3.7: if each $\mathcal{L}_\epsilon$ is positive then for every real-valued $g \in \mathbb{V}_\beta(\Omega)$ such that $\int g \, d\mu = 0$ and compact subset $C$ of the domain of $r_g$ there exists a closed interval $V$ such that $\lim_{\epsilon \rightarrow 0} \sup_{z \in V} (sz - \ln(\lambda_\epsilon(z))) = r_g(s)$ uniformly for $s \in C$.

**Proof.** As noted, the space $\mathbb{V}_\beta(\Omega)$ and Perron-Frobenius $\mathcal{L}$ satisfy the conditions of Theorem 3.2. We note that $\mathbb{V}_\beta(\Omega) \subseteq L^\infty(m)$ by [36, Proposition 3.4] and so every $g \in \mathbb{V}_\beta(\Omega)$ is also in $\mathbb{V}_\beta(\Omega) \cap L^\infty(m)$. Hence, Corollary 3.3 holds, which immediately yields the stability of the variance by Theorem 3.4. In the case where each $\mathcal{L}_\epsilon$ is a positive operator the stability of the rate function follows by Theorem 3.7 and Proposition 3.9. Proposition 2.3 is an immediate consequence of (KL) holding for $\{\mathcal{L}_\epsilon\}_{\epsilon \geq 0}$. \hfill ☐
We will now gather some important properties of $\mathbb{V}_\beta(\Omega)$ for later use. The first result concerns the continuity of inclusion of $\mathbb{V}_\beta(\Omega)$ into $L^\infty(m)$.

**Proposition 5.3** ([36, Proposition 3.4]). Let $\nu_d$ denote the measure of the $d$-dimensional unit ball i.e. $\nu_d = m(B_1(0))$. For every $f \in \mathbb{V}_\beta(\Omega)$ we have

$$\|f\|_{L^\infty} \leq \frac{1}{\nu_d \eta_0^\beta} (\eta_0^\beta |f|_\beta + |f|_{L^1}) \leq \max\{1, \eta_0^\beta\} \frac{\nu_d}{\eta_0} \|f\|_\beta.$$ 

In later sections we will prove that (KL3) holds for certain perturbations of the Perron-Frobenius operator by making use of the analogous inequality for the unperturbed operator.

**Proposition 5.4** ([36, Lemma 4.1]). Provided that $\eta_0$ is small enough, there exists $\gamma < 1$ and $D < \infty$ such that for each $f \in \mathbb{V}_\beta(\Omega)$ we have

$$|\mathcal{L}f|_\beta \leq \gamma |f|_\beta + D |f|_{L^1}.$$ 

**Remark.** We assume that $\eta_0$ is sufficiently small so that the conclusion of Proposition 5.4 holds.

**Remark.** In [36] the space $\mathbb{V}_\beta(\Omega)$ consists of real-valued functions only and so the proof of [36, Lemma 4.1] only applies to real-valued $f \in \mathbb{V}_\beta(\Omega)$. Examining the proof of [36, Lemma 4.1], we note that the same conclusion holds for complex-valued $f$ after minor modifications to the arguments. In particular, the essential infimum in [36, Proposition 3.2 (iii)] must be replaced by an essential supremum and consequently the resulting essential supremum term that appears when bounding $R^{(1)}_x$ must be bounded by $|f(y_i)| + \text{osc} (f, B(y_i))$. The rest of the argument holds mutatis mutandis.

### 5.1 Ulam perturbations

We will now prove that Ulam approximations of the Perron-Frobenius operator satisfy (KL) on $\mathbb{V}_\beta(\Omega)$. As a consequence, the theory developed in Section 3 yields the stability of the rate function and variance with respect to numerical approximation by Ulam’s method. While (KL2) is standard, (KL1) and (KL3) require significantly more work. In particular, the proof of (KL3) is quite long and depends critically on the geometry of the partitions inducing the Ulam approximations, so we defer its proof to Appendix A. The class of partitions we consider is the following.

**Definition 5.5.** For $\kappa \geq 1$ let $\mathcal{P}(\kappa)$ be the collection of finite measurable partitions $Q$ of $\Omega$ satisfying the following conditions:

1. The elements of $Q$ are compact and convex polytopes with non-empty interiors.
2. For every $I \in Q$ we have $\text{diam}(Q) \leq \kappa \text{diam}(B_I)$, where $B_I$ is a ball of maximal volume inscribed in $I$ and $\text{diam}(Q) = \max_{J \in Q} \text{diam}(J)$.

Each $Q \in \mathcal{P}(\kappa)$ induces a conditional expectation operator $\mathbb{E}_Q$ that is given by

$$\mathbb{E}_Q f = \sum_{I \in Q} \hat{f}_I 1_I,$$
where \( \hat{f}_K \) denotes the expected value of \( f \) on some measurable set \( K \) i.e.

\[
\hat{f}_K = \frac{1}{m(K)} \int_K f \, dm.
\]

We adopt the convention that \( \hat{f}_K = 0 \) if \( m(K) = \infty \). As in Section 4.1, \( \mathcal{L}_Q = \mathbb{E}_Q \mathcal{L} \) is the Ulam approximation of \( \mathcal{L} \) induced by \( Q \). If \( \{Q_{k}\}_{k>0} \subseteq \mathcal{P}(\kappa) \) is a sequence of partitions then we define the corresponding sequence of perturbed Perron-Frobenius operators by \( \mathcal{L}_{k} = \mathbb{E}_{Q_k} \mathcal{L} \), with \( \mathcal{L}_0 = \mathcal{L} \) as usual.

**Theorem 5.6 (Stability of statistical parameters under Ulam approximations).** Suppose that \( \mathcal{L} \) is a simple quasi-compact operator. Let \( \{Q_{k}\}_{k>0} \subseteq \mathcal{P}(\kappa) \) be such that \( \lim_{k \to 0} \text{diam}(Q_k) = 0 \), and let \( \{\mathcal{L}_{k}\}_{k>0} \) be the corresponding Ulam approximations of the Perron-Frobenius operator. If

\[
2d \left( 1 + \frac{\kappa}{\sqrt{d/2} - 1} \right)^{\beta} < \frac{1}{\gamma},
\]

where the constant \( \gamma \) is from Proposition 5.4, then there exists \( \epsilon_0 > 0 \) such that \( \{\mathcal{L}_{k}\}_{0 \leq k \leq \epsilon_0} \) satisfies (KL). Consequently, the conclusion of Theorem 5.2 holds.

**Example 5.7.** Let \( \Omega = [0, 1]^d \). For each \( n \in \mathbb{Z}^+ \) the set

\[
Q_n = \left\{ \frac{\Omega}{n} + b : b \in \mathbb{Z}^d \right\}
\]

is a measurable partition of \( \Omega \) consisting of cubes congruent to \([0, 1/n]^d\). It is straightforward to show that \( Q_n \in \mathcal{P} \left( \sqrt{d} \right) \) for every \( n \in \mathbb{Z}^+ \). Thus, Theorem 5.6 applies to any piecewise expanding map (in the sense of [36]) \( T : \Omega \to \Omega \) such that

\[
2d \left( 1 + \frac{\sqrt{d}}{\sqrt{d/2} - 1} \right)^{\beta} < \frac{1}{\gamma},
\]

where \( \gamma \) is from Proposition 5.4 and \( \beta \) is equal to the constant \( \alpha \) appearing in [36] Section 2, (PE2)).

**Proposition 5.8 ([KL1] for Ulam approximations).** If \( \{Q_{k}\}_{k>0} \subseteq \mathcal{P}(\kappa) \) is a sequence of partitions with \( \lim_{k \to 0} \text{diam}(Q_k) = 0 \), then there exists \( \epsilon_1 > 0 \) such that for every \( \epsilon \in [0, \epsilon_1] \) we have

\[
\| \mathcal{L}_{k} - \mathcal{L}_0 \| \leq 2 \text{diam}(Q_k)^{\beta} \| \mathcal{L} \|_{\beta}.
\]

In particular \( \{\mathcal{L}_{k}\}_{0 \leq k \leq \epsilon_1} \) satisfies (KL1).

**Proof.** The statement is clearly true for \( \epsilon = 0 \), so we may assume that \( \epsilon > 0 \). One has

\[
\| \mathbb{E}_{Q_k} - \text{Id} \| = \sup_{\| f \|_{\alpha} = 1} \sum_{I \in Q_k} \left| \int_I \hat{f}_I - f \right| \, dm.
\]

(28)
Fix \( I \in Q_\epsilon \). Let \( f_r \) and \( f_i \) be the real and imaginary parts of \( f \in \mathbb{V}_\beta(\Omega) \), respectively. By linearity of integration and the triangle inequality we have

\[
\int_I |\hat{f}_I - f| \, dm \leq \int_I |\hat{f}_r| \, dm + \int_I |\hat{f}_i| \, dm.
\]

Applying this to (28) yields

\[
\|E_{Q_\epsilon} - \text{Id}\| \leq 2 \sup \left\{ \sum_{I \in Q_\epsilon} \int_I |\hat{f}_I - f| \, dm : \|f\|_\alpha = 1, f \text{ is real valued} \right\}.
\] (29)

Let \( f \in \mathbb{V}_\beta(\Omega) \) be real valued and fix \( x \in I \). Then for almost every \( y_1, y_2 \in I \) we have

\[
|f(y_1) - f(y_2)| \leq \text{osc} \left( f, B_{\text{diam}(Q_\epsilon)(x)} \right).
\]

Taking the expectation with respect to \( y_1 \) over \( I \) we find that

\[
\int_I |\hat{f}_I - f(y_2)| \, dm \leq \int_I \text{osc} \left( f, B_{\text{diam}(Q_\epsilon)(x)} \right) \, dx.
\]

As \( \lim_{\epsilon \to 0} \text{diam}(Q_\epsilon) = 0 \) there exists \( \epsilon_1 > 0 \) such that \( \text{diam}(Q_\epsilon) \leq \eta_0 \) for all \( \epsilon \in (0, \epsilon_1] \). Recalling the definition of \( \|\cdot\|_\beta \), for each \( \epsilon \in (0, \epsilon_1] \) we have

\[
\sum_{I \in Q_\epsilon} \int_I |\hat{f}_I - f| \, dm \leq \int_{\mathbb{R}^d} \text{osc} \left( f, B_{\text{diam}(Q_\epsilon)(x)} \right) \, dx \leq \text{diam}(Q_\epsilon) \beta \|f\|_\beta,
\]

which, when applied to (29), then yields

\[
\|E_{Q_\epsilon} - \text{Id}\| \leq 2 \text{diam}(Q_\epsilon) \beta.
\]

We conclude the proof by noting that \( \|L_\epsilon - \mathcal{L}\| \leq \|E_{Q_\epsilon} - \text{Id}\| \|\mathcal{L}\|_\beta \).

**Proposition 5.9** ((KL2) for Ulam approximations). If \( \{Q_\epsilon\}_{\epsilon > 0} \subseteq \mathcal{P}(\kappa) \) is a sequence of partitions, then for each \( \epsilon \geq 0 \) and \( n \in \mathbb{Z}^+ \) we have

\[
|L_\epsilon^n|_{L^1} \leq 1.
\]

In particular, \( \{L_\epsilon\}_{\epsilon > 0} \) satisfies (KL2).

**Proof.** Note that \( |L|_{L^1} = 1 \) and that, for each \( \epsilon > 0 \), we have \( \|E_{Q_\epsilon}\|_{L^1} = 1 \). The conclusion follows immediately.

We will now verify (KL3) for the Ulam approximations. The main technical requirements are the following two lemmas, which we prove in Appendix A.
Lemma 5.10. If \( \{Q_\epsilon\}_{\epsilon > 0} \subseteq \mathcal{P}(\kappa) \) is a sequence of partitions with \( \lim_{\epsilon \to 0} \text{diam}(Q_\epsilon) = 0 \) then there exists \( \epsilon_2 > 0 \) such that

\[
\sup_{0 < \epsilon \leq \epsilon_2} |\mathbb{E}_{Q_\epsilon}|_\beta \leq 2d \left( 1 + \frac{\kappa}{\sqrt[3]{\frac{3}{2}} - 1} \right)^\beta.
\]

Lemma 5.11. If \( Q \in \mathcal{P}(\kappa) \) then \( \mathbb{E}_Q \in L(\mathbb{V}_\beta(\Omega)) \).

Proposition 5.12. (KL3) for Ulam approximations. Under the hypotheses of Theorem 5.6 there exists \( \alpha \in (0, 1) \) and \( C_3 > 0 \) such that for all \( f \in \mathbb{V}_\beta(\Omega) \), \( n \in \mathbb{Z}^+ \) and \( \epsilon \in [0, \epsilon_2] \) we have

\[
\|\mathcal{L}_\epsilon^n f\|_\beta \leq \alpha^n \|f\|_\beta + C_3 |f|_{L^1},
\]

where

\[
\alpha = \gamma 2d \left( 1 + \frac{\kappa}{\sqrt[3]{\frac{3}{2}} - 1} \right)^\beta \quad \text{and} \quad C = 2dD \left( 1 + \frac{\kappa}{\sqrt[3]{\frac{3}{2}} - 1} \right)^\beta.
\]

Proof. By Proposition 5.4 we have

\[
|\mathcal{L} f|_\beta \leq \gamma |f|_\beta + D |f|_{L^1},
\]

where \( \gamma < 1 \) and \( D < \infty \). Let \( \epsilon_2 \) be as in Lemma 5.10. If \( \epsilon \in [0, \epsilon_2] \) then

\[
|\mathcal{L}_\epsilon f|_\beta \leq \left( \sup_{0 < \epsilon \leq \epsilon_2} |\mathbb{E}_{Q_\epsilon}|_\beta \right) \|\mathcal{L} f\|_\beta \leq \alpha |f|_\beta + C |f|_{L^1},
\]

where

\[
\alpha = \gamma 2d \left( 1 + \frac{\kappa}{\sqrt[3]{\frac{3}{2}} - 1} \right)^\beta \quad \text{and} \quad C = 2dD \left( 1 + \frac{\kappa}{\sqrt[3]{\frac{3}{2}} - 1} \right)^\beta.
\]

Note that \( \alpha < 1 \) by the hypotheses of Theorem 5.6. By iterating (30) and applying Proposition 5.9 for each \( n \in \mathbb{Z}^+ \) we have

\[
|\mathcal{L}_\epsilon^n f|_\beta \leq \alpha^n |f|_\beta + \frac{C}{1 - \alpha} |f|_{L^1}.
\]

Finally, recalling the definition of \( \|\cdot\|_\beta \) and applying Proposition 5.9 again yields

\[
\|\mathcal{L}_\epsilon^n f\|_\beta \leq \alpha^n \|f\|_\beta + \left( 1 + \frac{C}{1 - \alpha} \right) |f|_{L^1}.
\]

Proof of Theorem 5.6. With \( \epsilon_1 \) as in Proposition 5.8 and \( \epsilon_2 \) as in Proposition 5.12 set \( \epsilon_0 = \min\{\epsilon_1, \epsilon_2\} \). By Lemma 5.11 and Propositions 5.8, 5.9 and 5.12 the family of operators \( \{\mathcal{L}_\epsilon\}_{0 \leq \epsilon \leq \epsilon_0} \) satisfies (KL). Upon noting that each \( \mathcal{L}_\epsilon \) is positive the required result follows from Theorem 5.2.

\[\Box\]
5.2 Stochastic perturbations

We now consider the case where the Perron-Frobenius operator is perturbed by convolution with a stochastic kernel.

Definition 5.13. We say that \( \{q_\epsilon\}_{\epsilon > 0} \subseteq L^1(\mathbb{R}^d) \) approximates the identity if

1. Each \( q_\epsilon \) is non-negative and satisfies \( |q_\epsilon|_{L^1} = 1 \).

2. For every \( \delta > 0 \) we have
   \[
   \lim_{\epsilon \to 0} \int_{|x| \geq \delta} q_\epsilon(x) \, dx = 0.
   \]

Let \( \{q_\epsilon\}_{\epsilon > 0} \) approximate the identity. We define the corresponding stochastically perturbed Perron-Frobenius operators by

\( \mathcal{L}_\epsilon = (q_\epsilon \ast \mathcal{L})1_\Omega \). As usual we set \( \mathcal{L}_0 = \mathcal{L} \). Our main result for this section is the following.

Theorem 5.14 (Stability of statistical parameters under stochastic perturbations). Suppose that \( \mathcal{L} \) is a simple quasi-compact operator. Let \( \{q_\epsilon\}_{\epsilon > 0} \) approximate the identity, and let \( \{\mathcal{L}_\epsilon\}_{\epsilon \geq 0} \) be the corresponding stochastically perturbed Perron-Frobenius operators. If

\[
\left( 1 + \frac{1}{\nu_0^{\frac{1}{\gamma}} \eta^\beta} \sup_{0 < \eta \leq \eta_0} \eta^{-\beta} m(B_\eta(\partial \Omega)) \right) < \frac{1}{\gamma},
\]

where the constant \( \gamma \) is from Proposition 5.4, then \( \{\mathcal{L}_\epsilon\}_{\epsilon \geq 0} \) satisfies (KL). Consequently, the conclusion of Theorem 5.2 holds.

Remark. It is not obvious when \( \sup_{0 < \eta \leq \eta_0} \eta^{-\beta} m(B_\eta(\partial \Omega)) < \infty \). This is the case if, for example, \( \Omega \) is convex [17, Theorem 6.6].

As in the previous section, we prove Theorem 5.14 by showing that \( \{\mathcal{L}_\epsilon\}_{\epsilon \geq 0} \) satisfies (KL).

Proposition 5.15 (KL1 for stochastic perturbations). If \( \{q_\epsilon\}_{\epsilon > 0} \) approximates the identity and \( \{\mathcal{L}_\epsilon\}_{\epsilon \geq 0} \) denotes the corresponding stochastically perturbed Perron-Frobenius operators, then

\[
\lim_{\epsilon \to 0} \|\mathcal{L}_\epsilon - \mathcal{L}_0\| = 0.
\]

Proof. We have

\[
\|\mathcal{L}_\epsilon - \mathcal{L}_0\| \leq \sup_{\|f\|_{\beta} = 1} \int_{\Omega} \int_{\mathbb{R}^d} |(\mathcal{L} f)(x - y) - (\mathcal{L}_\epsilon f)(x)| q_\epsilon(y) \, dy \, dx.
\]

Let \( \delta \in (0, \eta_0) \). We break the inner integral in (31) into two parts: the component where \( y \in B_\delta(0) \) and the component where \( y \in \mathbb{R}^d \setminus B_\delta(0) \). Dealing with the first component we have

\[
\int_{\Omega} \int_{B_\delta(0)} |(\mathcal{L} f)(x - y) - (\mathcal{L}_\epsilon f)(x)| q_\epsilon(y) \, dy \, dx \leq \int_{\Omega} \text{osc} (\mathcal{L} f, B_\delta(x)) \, dx \leq \delta^\beta \|\mathcal{L} f\|_{\beta}. \quad (32)
\]
Whereas for the second component we find
\[
\int_{\Omega} \int_{\mathbb{R}^d \setminus B_{\delta}(0)} |(\mathcal{L} f)(x - y) - (\mathcal{L} f)(x)| \, q_\epsilon(y) \, dy \, dx \leq 2 \|\mathcal{L} f\|_{L^\infty} \int_{\Omega} \int_{\mathbb{R}^d \setminus B_{\delta}(0)} q_\epsilon(y) \, dy \, dx
\]
\[
\leq 2 \max\{1, \eta_0^\beta\} \|\mathcal{L} f\|_{L^\infty} \int_{\mathbb{R}^d \setminus B_{\delta}(0)} q_\epsilon(y) \, dy,
\]
where the final inequality is obtained from Proposition 5.3. Combining (31), (32) and (33) we obtain
\[
\|\mathcal{L}_\epsilon - \mathcal{L}_0\| \leq \delta^\beta \|\mathcal{L}\|_{L^\infty} + 2 \max\{1, \eta_0^\beta\} \|\mathcal{L}\|_{L^\infty} \int_{\mathbb{R}^d \setminus B_{\delta}(0)} q_\epsilon(y) \, dy.
\]
As \(\{q_\epsilon\}_{\epsilon > 0}\) is an approximation to the identity, taking \(\epsilon \to 0\) yields
\[
\limsup_{\epsilon \to 0} \|\mathcal{L}_\epsilon - \mathcal{L}_0\| \leq \delta^\beta \|\mathcal{L}\|_{L^\infty}.
\]
We conclude the proof upon recalling that \(\delta\) may be chosen to be arbitrarily small. \(\square\)

**Proposition 5.16** ((KL2) for stochastic perturbations). If \(\{q_\epsilon\}_{\epsilon > 0}\) approximates the identity and \(\{\mathcal{L}_\epsilon\}_{\epsilon \geq 0}\) denotes the corresponding stochastically perturbed Perron-Frobenius operators, then for each \(\epsilon \geq 0\) and \(n \in \mathbb{Z}^+\) we have
\[
|\mathcal{L}_n^\epsilon|_{L^1} \leq 1.
\]

**Lemma 5.17.** If \(\{q_\epsilon\}_{\epsilon > 0}\) approximates the identity and \(\{\mathcal{L}_\epsilon\}_{\epsilon \geq 0}\) denotes the corresponding stochastically perturbed Perron-Frobenius operators, then for every \(\epsilon > 0\) and \(f \in \mathbb{V}_\beta(\Omega)\) we have
\[
|\mathcal{L} f|_{\beta} \leq \left(1 + \frac{1}{\nu_d \eta_0^d} \sup_{0 < \eta \leq \eta_0} \eta^{-\beta} m(B_\eta(\partial\Omega))\right) |\mathcal{L} f|_{\beta} + \frac{1}{\nu_d \eta_0^d} \left(\sup_{0 < \eta \leq \eta_0} \eta^{-\beta} m(B_\eta(\partial\Omega))\right) |f|_{L^1}.
\]

**Proof.** Fix \(\epsilon > 0\) and \(\eta \in (0, \eta_0]\). Since
\[
\text{osc} (\mathcal{L}_\epsilon f, B_\eta(x)) = \sup_{y_1, y_2 \in B_\eta(x)} \| (q_\epsilon * \mathcal{L})(y_1) 1_{\Omega}(y_1) - (q_\epsilon * \mathcal{L})(y_2) 1_{\Omega}(y_2) \|,
\]
we may consider three (not necessarily distinct) cases when bounding \(\text{osc} (\mathcal{L}_\epsilon f, B_\eta(x))\). Depending on how many of the characteristic function terms contribute to the essential supremum, we either have \(\text{osc} (\mathcal{L}_\epsilon f, B_\eta(x)) = 0\),
\[
\text{osc} (\mathcal{L}_\epsilon, B_\eta(x)) = \sup_{y_1 \in B_\eta(x)} \| (q_\epsilon * \mathcal{L} f)(y_1) \|,
\]

or
\[
\text{osc} (\mathcal{L}_\epsilon, B_\eta(x)) = \text{osc} (q_\epsilon * \mathcal{L} f, B_\eta(x)).
\]
As the support of \(f\) is a subset of \(\Omega\), if \(x \in \mathbb{R}^d \setminus B_\eta(\Omega)\) then \(\text{osc} (\mathcal{L}_\epsilon f, B_\eta(x)) = 0\). By
a similar argument, if \([34]\) holds, \(\text{osc}(\mathcal{L}_\epsilon f, B_\eta(x)) \neq 0\) and \([35]\) does not hold, then \(x \in B_\eta(\Omega)\). Hence,
\[
\int_{\mathbb{R}^d} \text{osc}(\mathcal{L}_\epsilon f, B_\eta(x)) \, dx \leq \int_{\mathbb{R}^d} \text{osc}(q_\epsilon \ast \mathcal{L} f, B_\eta(x)) \, dx + \int_{B_\eta(\Omega)} \text{ess sup}_{y_1 \in B_\eta(x)} |(q_\epsilon \ast \mathcal{L} f)(y_1)| \, dx.
\]
We now bound the quantity \([34]\). As \(|\mathcal{L}|_{L^1} \leq 1\) and by Proposition 5.3 we have
\[
|(q_\epsilon \ast \mathcal{L} f)(y_1)| = \left| \int_{\mathbb{R}^d} q_\epsilon(y)(\mathcal{L} f)(y_1 - y) \, dy \right| \leq \|\mathcal{L} f\|_{\infty} \leq \frac{1}{\nu d \eta^0}(\eta^0_{\beta}|\mathcal{L} f|_{\beta} + |f|_{L^1}).
\]
Hence,
\[
\int_{B_\eta(\Omega)} \text{ess sup}_{y_1 \in B_\eta(x)} |(q_\epsilon \ast \mathcal{L} f)(y_1)| \, dx \leq \frac{m(B_\eta(\partial \Omega))}{\nu d \eta^0}(\eta^0_{\beta}|\mathcal{L} f|_{\beta} + |f|_{L^1}). \tag{37}
\]
Alternatively, to bound \([35]\) we note that
\[
\text{osc}(q_\epsilon \ast \mathcal{L} f, B_\eta(x)) = \text{ess sup}_{y_1, y_2 \in B_\eta(x)} \left| \int_{\mathbb{R}^d} q_\epsilon(y)((\mathcal{L} f)(y_1 - y) - (\mathcal{L} f)(y_2 - y)) \, dy \right| \leq \int_{\mathbb{R}^d} q_\epsilon(y)\text{osc}(\mathcal{L} f, B_\eta(x - y)) \, dy.
\]
By changing variables and applying Fubini-Tonelli we obtain
\[
\int_{\mathbb{R}^d} \text{osc}(q_\epsilon \ast \mathcal{L} f, B_\eta(x)) \, dx \leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} q_\epsilon(y)\text{osc}(\mathcal{L} f, B_\eta(x - y)) \, dy \, dx \leq \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} q_\epsilon(x - y) \, dx \right) \text{osc}(\mathcal{L} f, B_\eta(y)) \, dy \tag{38}
\]
Applying \([37]\) and \([38]\) to \([36]\) yields
\[
\int_{\mathbb{R}^d} \text{osc}(\mathcal{L}_\epsilon f, B_\eta(x)) \, dx \leq \int_{\mathbb{R}^d} \text{osc}(\mathcal{L} f, B_\eta(x)) \, dx + \frac{m(B_\eta(\partial \Omega))}{\nu d \eta^0}(\eta^0_{\beta}|\mathcal{L} f|_{\beta} + |f|_{L^1}).
\]
Thus,
\[
|\mathcal{L}_\epsilon f|_{\beta} \leq |\mathcal{L} f|_{\beta} + \frac{1}{\nu d \eta^0} \left( \sup_{0 < \eta \leq \eta_0} \eta^{-\beta} m(B_\eta(\partial \Omega)) \right)(\eta^0_{\beta}|\mathcal{L} f|_{\beta} + |f|_{L^1}),
\]
which yields the required bound. \(\square\)

**Proposition 5.18** ([(KL3)] for stochastic perturbations). Under the hypotheses of Theorem 5.14 there exists \(\alpha \in (0, 1)\) and \(C_3 > 0\) such that for every \(\epsilon \geq 0\) and \(f \in \mathbb{V}_\beta(\Omega)\) we have
\[
|\mathcal{L}_\epsilon^n f|_{\beta} \leq \alpha^n |f|_{\beta} + C_3 |f|_{L^1}.
\]

31
Proof. By Lemma 5.17 and Propositions 5.4 and 5.16 for each $\epsilon > 0$ and $f \in V_\beta(\Omega)$ we have

$$|L_\epsilon f|_* \leq \alpha |f|_* + C |f|_{L^1}$$

where

$$\alpha = \gamma \left( 1 + \frac{\eta_0^\beta}{\nu_\beta \eta_0^\nu} \sup_{0 < \eta \leq \eta_0} \eta^{-\beta} m(B_\eta(\partial\Omega)) \right)$$

and

$$C = D + \left( D + 1 \right) \frac{\eta_0^\beta}{\nu_\beta \eta_0^\nu} \left( \sup_{0 < \eta \leq \eta_0} \eta^{-\beta} m(B_\eta(\partial\Omega)) \right).$$

Note that $\alpha < 1$ by the hypotheses of Theorem 5.14. The remainder of the argument is identical to that of Proposition 5.12.

Proof of Theorem 5.14. By Lemma 5.17 each operator $L_\epsilon$ is in $L(V_\beta(\Omega))$. By Propositions 5.15, 5.16 and 5.18 the family of operators $\{L_\epsilon\}_{\epsilon \geq 0}$ satisfies (KL). We also note that each $L_\epsilon$ is a positive operator, as each $q_\epsilon$ is non-negative. The required result then follows from Theorem 5.2.

6 Acknowledgements

HC is supported by an Australian Government Research Training Program Scholarship and the UNSW School of Mathematics and Statistics. GF is partially supported by an Australian Research Council Discovery Project. Both authors thank Davor Dragičević for helpful conversations during the writing of this work.

A The proofs of Lemmas 5.10 and 5.11

Before discussing our strategy for proving Lemmas 5.10 and 5.11 we must discuss the relationship between the space $V_\beta(\Omega)$ and the seminorm $||_{\beta}$. It is noted in [26] that while $V_\beta(\Omega)$ is independent of $\eta_0$, the seminorm $||_{\beta}$ is not. However, changing $\eta_0$ preserves the topology induced by the relevant seminorm, which will be critical to proofs in this appendix. The following lemma gives the relevant bounds.

Lemma A.1. For $\zeta > 0$ and $f \in L^1(\mathbb{R}^d)$ let

$$|f|_{\beta,\zeta} = \sup_{0 < \eta \leq \zeta} \eta^{-\beta} \int_{\mathbb{R}^d} \text{osc} (f, B_\eta(x)) \, dx.$$

If $0 < t \leq s$ then

$$|f|_{\beta,t} \leq |f|_{\beta,s} \leq S(t,s) |f|_{\beta,t},$$

where $S(t,s)$ denotes the minimal number of balls of radius $t$ required to cover (up to a set of measure 0) a ball of radius $s$.

Proof. The inequality $|f|_{\beta,t} \leq |f|_{\beta,s}$ is trivial. Let $f \in L^1(\mathbb{R}^d)$. If

$$\sup_{0 < \eta \leq t} \eta^{-\beta} \int_{\mathbb{R}^d} \text{osc} (f, B_\eta(x)) \, dx = \sup_{0 < \eta \leq s} \eta^{-\beta} \int_{\mathbb{R}^d} \text{osc} (f, B_\eta(x)) \, dx,$$

(39)
then, as $S(t, s) \geq 1$, we clearly have $|f|_{\beta, s} \leq S(t, s) |f|_{\beta, t}$. Alternatively, if (39) does not hold then

$$\sup_{0 < \eta \leq t} \eta^{-\beta} \int_{\mathbb{R}^d} \text{osc} \left( f, B_\eta(x) \right) \, dx < \sup_{t < \eta \leq s} \eta^{-\beta} \int_{\mathbb{R}^d} \text{osc} \left( f, B_\eta(x) \right) \, dx.$$  

By the definition of $S(t, s)$ there exists $\{c_i\}_{i=1}^{S(t,s)} \subseteq \mathbb{R}^d$ and a set $N$ of measure 0 such that

$$B_s(x) \setminus (N + x) \subseteq \bigcup_{i=1}^{S(t,s)} B_t(x + c_i)$$

for every $x \in \mathbb{R}^d$. Hence, for any $\eta \in (t, s]$ and $x \in \mathbb{R}^d$ we have

$$\text{osc} \left( f, B_\eta(x) \right) \leq \text{osc} \left( f, B_s(x) \right) \leq \sum_{i=1}^{S(t,s)} \text{osc} \left( f, B_t(x + c_i) \right).$$

After integrating, taking the supremum and applying the definition of $|\cdot|_{\beta}$ we obtain

$$\sup_{t < \eta \leq s} \eta^{-\beta} \int_{\mathbb{R}^d} \text{osc} \left( f, B_\eta(x) \right) \, dx \leq S(t, s) t^{-\beta} \int_{\mathbb{R}^d} \text{osc} \left( f, B_t(x) \right) \, dx \leq S(t, s) |f|_{\beta,t},$$

completing the proof.

We obtain Lemmas 5.10 and 5.11 as corollaries to the following result.

**Proposition A.2.** If $Q \in \mathcal{P}(\kappa)$ satisfies $\text{diam}(Q) < \eta_0$, then

$$|E_Q|_{\beta} \leq S(\eta_0 - \text{diam}(Q), \eta_0) \left( 1 + \frac{\kappa}{\sqrt{\frac{d}{2}} - 1} \right)^{\beta}.$$

(40)

We prove Proposition A.2 by using Lemma A.1 to extend a bound for $\sup_{|f|_{\beta}=1} |E_Q f|_{\beta, \eta_0 - \text{diam}(Q)}$ to a bound for $|E_Q|_{\beta}$. We bound $\sup_{|f|_{\beta}=1} |E_Q f|_{\beta, \eta_0 - \text{diam}(Q)}$ by combining two bounds for $\eta^{-\beta} \int \text{osc} \left( E_Q f, B_\eta(x) \right) \, dm$ over $\eta \in (0, \eta_0 - \text{diam}(Q))$:

1. We obtain the ‘big’ $\eta$ bound by scaling the $\eta$ balls in osc up to $\eta + \text{diam}(Q)$ balls. This bound is useful for large $\eta$, but grows unboundedly as $\eta$ vanishes. We obtain this bound in Lemma A.4.

2. We obtain the ‘small’ $\eta$ bound by using the geometry of the elements of $Q$ to quantify the decay of the measure of the support of osc $(E_Q f, B_\eta(\cdot))$ as $\eta$ vanishes. This bound is used for $\eta$ arbitrarily close to 0. Obtaining this bound is more complicated and is developed in Lemmas A.5 - A.7.

Before proving either of these bounds we derive an explicit expression for $\int \text{osc} \left( E_Q f, B_\eta(x) \right) \, dm$. 

33
Lemma A.3. Let $Q \in \mathcal{P}(\kappa)$, define $Q' = Q \cup \{\Omega^c\}$, and for each $\eta > 0$ and $x \in \mathbb{R}^d$ let

$$N(x, \eta) = \{J \in Q' : B_\eta(x) \cap J \neq \emptyset\}.$$ 

For each $\eta > 0$, $f \in \mathcal{V}_\beta(\Omega)$ and $S \subseteq Q'$ let

$$M_S(f) = \max_{J, K \in S} |\hat{f}_J - \hat{f}_K| \quad \text{and} \quad A_{S, \eta} = \{x \in \mathbb{R}^d : N(x, \eta) = S\}.$$ 

Then each $A_{S, \eta}$ is measurable, and for every $x \in \mathbb{R}^d$

$$\text{osc} \left( \mathbb{E}_Q f, B_\eta(x) \right) = M_{N(x, \eta)}(f). \quad (41)$$

Hence,

$$\int_{\mathbb{R}^d} \text{osc} \left( \mathbb{E}_Q f, B_\eta(x) \right) \, dx = \sum_{S \subseteq Q'} m(A_{S, \eta}) M_S(f). \quad (42)$$

Proof. For every $J \in Q$ the equality

$$\{x \in \mathbb{R}^d : B_\eta(x) \cap J \neq \emptyset\} = \bigcup_{y \in J} B_\eta(y),$$

implies that both sets are open, and therefore measurable. Recalling from Definition 5.5 that $Q$ is finite and noting the equality

$$A_{S, \eta} = \left( \bigcap_{J \in S} \{x \in \mathbb{R}^d : B_\eta(x) \cap J \neq \emptyset\} \right) \cap \left( \bigcap_{K \in Q' \setminus S} \{x \in \mathbb{R}^d : B_\eta(x) \cap K = \emptyset\} \right),$$

we conclude that each $A_{S, \eta}$ is measurable. Considering the definition of $N(x, \eta)$, we note that the family of sets $\{A_{S, \eta} : S \subseteq Q'\}$ partitions $\mathbb{R}^d$. Hence,

$$\int_{\mathbb{R}^d} \text{osc} \left( \mathbb{E}_Q f, B_\eta(x) \right) \, dx = \sum_{S \subseteq Q'} \int_{A_{S, \eta}} \text{osc} \left( \mathbb{E}_Q f, B_\eta(x) \right) \, dx,$$

where we note that the sum on the right-hand side is well defined as only finitely many terms are ever non-zero. Thus, in order to prove (42) it suffices to prove (41). If $N(x, \eta) = S$ then

$$(\mathbb{E}_Q f)(B_\eta(x)) = \left\{ \hat{f}_J : J \cap B_\eta(x) \neq \emptyset \right\} = \left\{ \hat{f}_J : J \in S \right\},$$

which is finite as $Q'$ is finite. By applying the definition of osc, we find that

$$\text{osc} \left( \mathbb{E}_Q f, B_\eta(x) \right) = \max_{J, K \in S} |\hat{f}_J - \hat{f}_K| = M_S(f),$$

which is exactly (42). \hfill \Box

We may now obtain the ‘big’ $\eta$ bound.
Lemma A.4. If \( Q \in \mathcal{P}(\kappa) \) then for each \( f \in \mathcal{V}_\beta(\Omega) \) and \( \eta > 0 \) we have

\[
\int_{\mathbb{R}^d} \text{osc}(\mathbb{E}_Q f, B_\eta(x)) \, dx \leq \int_{\mathbb{R}^d} \text{osc}(f, B_{\eta + \text{diam}(Q)}(x)) \, dx. \tag{43}
\]

Furthermore, if \( \text{diam}(Q) < \eta_0 \) and \( \eta \in (0, \eta_0 - \text{diam}(Q)] \) then

\[
\eta^{-\beta} \int_{\mathbb{R}^d} \text{osc}(\mathbb{E}_Q f, B_\eta(x)) \, dx \leq \left(1 + \frac{\text{diam}(Q)}{\eta}\right)^\beta |f|_\beta. \tag{44}
\]

Proof. Fix \( x \in \mathbb{R}^d \). By Lemma A.3, we have

\[
\text{osc}(\mathbb{E}_Q f, B_\eta(x)) = \max_{J,K} \left| \hat{f}_J - \hat{f}_K \right|.
\]

Suppose that \( y \in I \) for some \( I \in N(x, \eta) \setminus \{\Omega^c\} \). By the definition of \( N(x, \eta) \) there exists \( z \in I \) such that \( |z - x| < \eta \). Since \( |z - y| \leq \text{diam}(Q) \) we have \( |y - x| < \eta + \text{diam}(Q) \). Hence

\[
\bigcup_{I \in N(x, \eta) \setminus \{\Omega^c\}} I \subseteq B_{\eta + \text{diam}(Q)}(x). \tag{45}
\]

Now suppose that \( J, K \in N(x, \eta) \). In the case where \( J, K \in N(x, \eta) \setminus \{\Omega^c\} \) the inclusion \( \{45\} \) implies that for almost every \( (y_1, y_2) \in J \times K \) we have

\[
|f(y_1) - f(y_2)| \leq \text{osc}(f, B_{\eta + \text{diam}(Q)}(x)).
\]

By taking expectations with respect to \( y_1 \) over \( J \) and \( y_2 \) over \( K \) we obtain

\[
\left| \hat{f}_J - \hat{f}_K \right| \leq \text{osc}(f, B_{\eta + \text{diam}(Q)}(x)). \tag{46}
\]

Alternatively, if either \( J \) or \( K \) is equal to \( \Omega^c \) then

\[
\left| \hat{f}_J - \hat{f}_K \right| = \max \left\{ \left| \hat{f}_J \right|, \left| \hat{f}_K \right| \right\} \leq \max_{I \in N(x, \eta) \setminus \{\Omega^c\}} \left| \hat{f}_I \right| \leq \text{ess sup}_{y \in B_{\eta + \text{diam}(Q)}} |f(y)|, \tag{47}
\]

where we obtain the last inequality by using \( \{45\} \). Note that \( B_{\eta + \text{diam}(Q)}(x) \cap \Omega^c \) has non-zero measure, which implies that

\[
\text{ess sup}_{y \in B_{\eta + \text{diam}(Q)}(x)} |f(y)| \leq \text{osc}(f, B_{\eta + \text{diam}(Q)}(x)). \tag{48}
\]

By combining \( \{47\} \) and \( \{48\} \) we obtain \( \{46\} \) for the case where either \( J \) or \( K \) is equal to \( \Omega^c \). As \( J \) and \( K \) were arbitrary elements of \( N(x, \eta) \) this implies that

\[
\text{osc}(\mathbb{E}_Q f, B_\eta(x)) = \max_{I,J \in N(x, \eta)} \left| \hat{f}_I - \hat{f}_J \right| \leq \text{osc}(f, B_{\eta + \text{diam}(Q)}(x)).
\]

By integrating with respect to \( x \) over \( \mathbb{R}^d \) we obtain \( \{43\} \). To prove \( \{44\} \), suppose that \( \text{diam}(Q) < \eta_0 \). If \( \eta \in (0, \eta_0 - \text{diam}(Q)] \) then \( \eta + \text{diam}(Q) \in (0, \eta_0] \) and so the definition of \( |\cdot|_\beta \) implies that

\[
\int_{\mathbb{R}^d} \text{osc}(f, B_{\eta + \text{diam}(Q)}(x)) \, dx \leq (\eta + \text{diam}(Q))^\beta |f|_\beta.
\]
Thus
\[ \eta^{-\beta} \int_{\mathbb{R}^d} \text{osc} (E_Q f, B_\eta(x)) \, dx \leq \eta^{-\beta} \int_{\mathbb{R}^d} \text{osc} (f, B_{\eta + \text{diam}(Q)}(x)) \, dx \leq \left(1 + \frac{\text{diam}(Q)}{\eta}\right)^{\beta} |f|_\beta. \]

We will now pursue the ‘small’ \( \eta \) bound.

**Lemma A.5.** Let \( Q \in \mathcal{P}(\kappa) \) and let \( S \subseteq Q' \) satisfy \( |S| > 1 \). If \( I \in S \setminus \{\Omega'\} \) and \( \eta > 0 \) then \( A_{S,\eta} \subseteq B_\eta(\partial I) \).

**Proof.** The claim is trivially true if \( A_{S,\eta} \) is empty, henceforth we assume that it is not. Let \( x \in A_{S,\eta} \). We distinguish between two cases: either \( x \in I \) or \( x \notin I \). Suppose that \( x \in I \). As \( |S| > 1 \) and \( N(x, \eta) = S \) there exists some \( J \in Q' \setminus \{I\} \) such that \( B_\eta(x) \cap J \neq \emptyset \).

Actually, as the closure of the interior of \( J \) is \( J \), we have \( B_\eta(x) \cap \text{int}(J) \neq \emptyset \). In this case let \( y \in B_\eta(x) \cap \text{int}(J) \); as \( J \) and \( I \) are convex elements of a measurable partition we have \( J \cap I \subseteq \partial J \cap \partial I \) and so \( y \notin I \). Alternatively, if \( x \notin I \), then let \( y \in B_\eta(x) \cap I \), which is non-empty by a similar argument. In both cases we have a pair of points in \( A_{S,\eta} \): one in \( I \) and the other not. Recalling that elements of \( Q' \) have non-empty interior and then considering the line segment that joins \( x \) and \( y \), it is straightforward to verify that there exists some \( z \in \partial I \) on this line segment. Clearly \( |x - z| < \eta \) and so \( x \in B_\eta(\partial I) \), which completes the proof.

**Lemma A.6.** Let \( Q \in \mathcal{P}(\kappa) \). If \( \eta > 0 \) and \( S \subseteq Q' \) is such that \( |S| > 1 \) and \( m(A_{S,\eta}) > 0 \), then for each \( f \in \mathbb{V}_\beta(\Omega) \) we have

\[ M_S(f) \leq \max_{J,K \in S, J \neq K} \left( \int_J \frac{\text{osc} (f, B_{\eta + \text{diam}(Q)}(x))}{m(J)} \, dx + \int_K \frac{\text{osc} (f, B_{\eta + \text{diam}(Q)}(x))}{m(K)} \, dx \right). \]

**Proof.** Let \( J, K \in S \) be partition elements satisfying

\[ M_S(f) = \left| \hat{f}_J - \hat{f}_K \right|. \]

We may assume that \( J \neq K \), as this case does not contribute to the maximum. Let us first consider the case where \( \Omega' \in \{J, K\} \); without loss of generality let \( K = \Omega' \). For every \( j \in J \) we have \( J \subseteq B_{\eta + \text{diam}(Q)}(j) \). Hence, as \( B_{\eta + \text{diam}(Q)}(j) \cap \Omega' \) has non-empty interior, and therefore non-zero measure, for almost every \( j, j' \in J \) and \( k' \in B_{\eta + \text{diam}(Q)}(j) \cap \Omega' \) we have

\[ |f(j')| = |f(j') - f(k')| \leq \text{osc} (f, B_{\eta + \text{diam}(Q)}(j)). \]

Taking expectations with respect to \( j' \) and \( j \) over \( J \) yields

\[ M_S(f) = \left| \hat{f}_J \right| \leq \int_J \frac{\text{osc} (f, B_{\eta + \text{diam}(Q)}(x))}{m(J)} \, dx, \]

which implies the required conclusion. Alternatively suppose that neither \( J \) nor \( K \) is equal to \( \Omega' \). Fix \( j \in J \) and \( k \in K \). For any \( j' \in J \) we have \( |j - j'| \leq \text{diam}(Q) \) and so \( j' \in B_{\eta + \text{diam}(Q)}(j) \). Similarly, for every \( k' \in K \) we have \( k' \in B_{\eta + \text{diam}(Q)}(k) \). As \( m(A_{S,\eta}) > 0 \), we know that \( A_{S,\eta} \neq \emptyset \). For \( z \in A_{S,\eta} \) the intersection \( B_\eta(z) \cap J \) is non-empty.
and so \( z \in B_{\eta + \text{diam}(Q)}(j) \). Similarly, \( z \in B_{\eta + \text{diam}(Q)}(k) \). Hence, for almost every \( j' \in J \) and \( k' \in K \),
\[
|f(j') - f(k')| \leq |f(j') - f(z)| + |f(k') - f(z)| \\
\leq \text{osc} \left( f, B_{\eta + \text{diam}(Q)}(j) \right) + \text{osc} \left( f, B_{\eta + \text{diam}(Q)}(k) \right).
\]
By taking the expectation with respect to \( j' \) over \( J \) and \( k' \) over \( K \), we find
\[
\left| \hat{f}_J - \hat{f}_K \right| \leq \text{osc} \left( f, B_{\eta + \text{diam}(Q)}(j) \right) + \text{osc} \left( f, B_{\eta + \text{diam}(Q)}(k) \right). \tag{49}
\]
Since (49) holds for every \( j \in J \) and \( k \in K \), we may take expectations again to obtain
\[
\left| \hat{f}_J - \hat{f}_K \right| \leq \int_J \frac{\text{osc} \left( f, B_{\eta + \text{diam}(Q)}(x) \right)}{m(J)} \, dx + \int_K \frac{\text{osc} \left( f, B_{\eta + \text{diam}(Q)}(x) \right)}{m(K)} \, dx.
\]
Taking the maximum over all distinct pairs of \( J, K \in S \) we find
\[
M_S(f) \leq \max_{J,K \in S, J \neq K} \left( \int_J \frac{\text{osc} \left( f, B_{\eta + \text{diam}(Q)}(x) \right)}{m(J)} \, dx + \int_K \frac{\text{osc} \left( f, B_{\eta + \text{diam}(Q)}(x) \right)}{m(K)} \, dx \right).
\]

\( \square \)

Combining the previous two results yields the ‘small’ \( \eta \) bound of Lemma A.7.

**Lemma A.7.** Let \( Q \in \mathcal{P}(\kappa) \). If \( \text{diam}(Q) < \eta_0 \), \( \eta \in (0, \eta_0 - \text{diam}(Q)] \) and \( f \in \mathcal{V}_\beta(\Omega) \) then
\[
\eta^{-\beta} \int_{\mathbb{R}^d} \text{osc} \left( \mathbb{E}_Q f, B_\eta(x) \right) \, dx \leq \left( \max_{I \in Q} \frac{m(B_\eta(\partial I))}{m(I)} \right) \left( 1 + \frac{\text{diam}(Q)}{\eta} \right)^\beta |f|_{\beta}.
\]

**Proof.** By Lemma A.3 we have
\[
\int_{\mathbb{R}^d} \text{osc} \left( \mathbb{E}_Q f, B_\eta(x) \right) \, dx = \sum_{S \subseteq Q'} m(A_{S,\eta}) M_S(f). \tag{50}
\]
Let \( G = \{ S \subseteq Q' : |S| > 1, m(A_{S,\eta}) > 0 \} \). Since \( m(A_{S,\eta}) M_S(f) = 0 \) if \( S \notin G \) we may restrict the sum in (50) to
\[
\int_{\mathbb{R}^d} \text{osc} \left( \mathbb{E}_Q f, B_\eta(x) \right) \, dx = \sum_{S \in G} m(A_{S,\eta}) M_S(f). \tag{51}
\]
Applying Lemma A.6 to each of the terms in (51) yields
\[
\int_{\mathbb{R}^d} \text{osc} \left( \mathbb{E}_Q f, B_\eta(x) \right) \, dx \leq \sum_{S \in G} m(A_{S,\eta}) \max_{J,K \in S, J \neq K} \left( \int_J \frac{\text{osc} \left( f, B_{\eta + \text{diam}(Q)}(x) \right)}{m(J)} \, dx \right.
\]
\[
\left. + \int_K \frac{\text{osc} \left( f, B_{\eta + \text{diam}(Q)}(x) \right)}{m(K)} \, dx \right).
\tag{52}
\]

37
By rearranging the terms in \[52\] to sum over elements of \(Q\) we obtain
\[
\int_{\mathbb{R}^d} \operatorname{osc} ({\mathbb{E}}_Q f, B_\eta(x)) \, dx \leq \sum_{I \in Q} \sum_{S \in G, I \subseteq S} m(A_{S,\eta}) / m(I) \int_I \operatorname{osc} (f, B_{\eta + \text{diam}(Q)}(x)) \, dx \\
\leq \left( \max_{I \in Q} \frac{m(B_\eta(\partial I))}{m(I)} \right) \int_{\mathbb{R}^d} \operatorname{osc} (f, B_{\eta + \text{diam}(Q)}(x)) \, dx,
\]
where we omit the case of \(I = \Omega^c\), as it does not contribute to the sum. Since the sets \(\{A_{S,\eta}\}_{S \subseteq Q}^c\) are disjoint, Lemma A.5 implies that
\[
\sum_{S \in G, I \subseteq S} m(A_{S,\eta}) \leq m(B_\eta(\partial I)).
\]
Thus,
\[
\int_{\mathbb{R}^d} \operatorname{osc} ({\mathbb{E}}_Q f, B_\eta(x)) \, dx \leq \left( \max_{I \in Q} \frac{m(B_\eta(\partial I))}{m(I)} \right) \int_{\mathbb{R}^d} \operatorname{osc} (f, B_{\eta + \text{diam}(Q)}(x)) \, dx.
\]
The required inequality follows by applying the definition of \(|\cdot|_\beta\). \(\square\)

Before proving Proposition A.2 we require a technical lemma for an inequality from convex geometry. For \(U, V \subseteq \mathbb{R}^d\) the Minkowski sum of \(U\) and \(V\) is denoted by \(U + V\) and equal to \(\{u + v : u \in U, v \in V\}\); for basic properties we refer to [17, Section 6.1].

**Lemma A.8.** If \(I\) is a compact convex polytope then for every \(\eta > 0\) we have
\[
m(B_\eta(\partial I) \cap I) \leq m(B_\eta(\partial I) \cap I^c).
\]

**Proof.** Let \(m_{d-1}\) denote \(d - 1\) dimensional Lebesgue measure. By Steiner’s formula [17, Theorem 6.6] there exists a polynomial \(p_I\) with positive coefficients and of degree \(d\) such that \(m(B_\eta(I)) = p_I(\eta)\). The constant coefficient of \(p_I\) is clearly \(m(I)\), while the coefficient of the linear term is \(m_{d-1}(\partial I)\) i.e. the surface area of \(I\). Note that \(m(B_\eta(\partial I) \cap I^c) = p_I(\eta) - m(I)\). We will prove that \(m(B_\eta(\partial I) \cap I) \leq \eta m_{d-1}(\partial I)\). Since \(p_I\) has degree greater than or equal to 2 and positive coefficients, it follows that
\[
m(B_\eta(\partial I) \cap I) \leq \eta m_{d-1}(\partial I) \leq p_I(\eta) - m(I) \leq m(B_\eta(\partial I) \cap I^c),
\]
and would therefore complete the proof.

Let \(\mathcal{F}(I)\) denote the set of set of facets of \(I\). Clearly \(m_{d-1}(\partial I) = \sum_{F \in \mathcal{F}(I)} m_{d-1}(F)\). Let \(y \in B_\eta(\partial I) \cap I\) and denote by \(F\) the (possibly not unique) facet in \(\mathcal{F}(I)\) that minimises the distance from \(y\) to \(\partial I\). Let \(x\) be the point on \(F\) attaining said minimum. If \(x - y\) is not normal to \(F\) then the ball \(B_{|x-y|}(y)\) is not tangent to \(F\) and so there exists \(z \in B_{|x-y|}(y) \cap I^c\). The line segment from \(y\) to \(z\) must intersect \(\partial I\) at some point that is strictly closer to \(y\) than \(x\), which contradicts \(x\) minimising the distance from \(y\) to \(\partial I\).

Hence, \(x - y\) must be normal to \(F\) and so \(y \in F + [0, \eta] \times n_F\), where \(n_F\) is the inward facing unit normal vector to \(F\). This implies that
\[
B_\eta(\partial I) \cap I \subseteq \bigcap_{F \in \mathcal{F}(I)} F + [0, \eta] \times n_F
\]
and so \(m(B_\eta(\partial I) \cap I) \leq \eta \sum_{F \in \mathcal{F}(I)} m_{d-1}(F) = \eta m_{d-1}(\partial I)\) as required. \(\square\)
The proof of Proposition A.2. We begin by bounding
\[ \sup_{|f|_\beta = 1} |E_Q f|_{\beta, \eta_0 - \text{diam}(Q)}. \]

Let \( b : \mathbb{R} \to \mathbb{R} \) be defined by
\[ b(\eta) = \left( 1 + \frac{\text{diam}(Q)}{\eta} \right)^\beta. \]

By taking the minimum of the bounds in Lemmas A.4 and A.7 we have
\[ |E_Q f|_{\beta, \eta_0} - \text{diam}(Q) \leq \sup_{0 < \eta \leq \eta_0} \min \left\{ \max_{I \in Q} \frac{m(B_\eta(\partial I))}{m(I)}, 1 \right\} b(\eta) |f|_\beta. \] (54)

We will now bound \( \max_{I \in Q} \frac{m(B_\eta(\partial I))}{m(I)} \). Lemma A.8 implies that for any \( I \in Q \) we have
\[ \frac{m(B_\eta(\partial I))}{m(I)} \leq 2 \frac{m(I + (\eta/2)B_1(0)) - m(I)}{m(I)}. \] (55)

Noting that \( B_\eta(\partial I) \cap I^c = B_\eta(I) \setminus I \) and \( B_\eta(I) = I + (\eta/2)B_1(0) \), we obtain
\[ \frac{m(B_\eta(\partial I))}{m(I)} \leq 2 \left( 1 + \frac{\eta}{\text{diam}(B_1)} \right)^d - 2 \left( 1 + \frac{\kappa_\eta}{\text{diam}(Q)} \right)^d - 2. \] (57)

By applying (57) to (54), and recalling that \( 1/\text{diam}(B_1) \leq \kappa/\text{diam}(Q) \) as \( Q \in \mathcal{P}(\kappa) \), we find that
\[ \frac{m(B_\eta(\partial I))}{m(I)} \leq 2 \left( 1 + \frac{\eta}{\text{diam}(B_1)} \right)^d - 2 \leq 2 \left( 1 + \frac{\kappa_\eta}{\text{diam}(Q)} \right)^d - 2. \] (58)

By applying (57) to (54) we obtain
\[ |E_Q f|_{\beta, \eta_0 - \text{diam}(Q)} \leq \sup_{0 < \eta \leq \eta_0} \min \left\{ \left( 2 \left( 1 + \frac{\kappa_\eta}{\text{diam}(Q)} \right)^d - 2 \right) b(\eta), b(\eta) \right\} |f|_\beta. \] (58)

It is clear that \( b \) is monotonically decreasing. Note that
\[ 2 \left( 1 + \frac{\kappa_\eta}{\text{diam}(Q)} \right)^d - 1 \right) b(\eta) = 2\eta^{-\beta} \left( \left( 1 + \frac{\kappa_\eta}{\text{diam}(Q)} \right)^d - 1 \right) \eta + \text{diam}(Q))^\beta. \] (59)

The map \( \eta \mapsto (\eta + \text{diam}(Q))^\beta \) is clearly monotonically increasing on \((0, \eta_0]\). As \( d \geq 2 \) and \( \beta \in (0, 1] \), the map \( \eta \mapsto \eta^{-\beta} \left( 1 + \frac{\kappa_\eta}{\text{diam}(Q)} \right)^d - 1 \) is monotonically increasing on
(0, \eta_0] too. Thus the left-hand side of (59) is monotonically increasing. Since both \( b \) and the left-hand side of (59) are continuous on (0, \eta_0], \( b \) is monotonically decreasing and the left-hand side of (59) is monotonically increasing, it follows that if \( \eta' \in (0, \infty) \) solves

\[
2 \left( 1 + \frac{\kappa \eta'}{\text{diam}(Q)} \right)^d - 2 = 1,
\]

then

\[
\sup_{0 < \eta \leq \eta_0} \min \left\{ \left( 2 \left( 1 + \frac{\kappa \eta}{\text{diam}(Q)} \right)^d - 2 \right) b(\eta), b(\eta) \right\} \leq b(\eta').
\]

Solving (60) yields

\[
\frac{\text{diam}(Q)}{\eta'} = \frac{\kappa}{\sqrt{\frac{3}{2} - 1}}.
\]

By substituting this into (58) we obtain the bound

\[
|E_Q f|_{\beta, \eta_0 - \text{diam}(Q)} \leq \left( 1 + \frac{\kappa}{\sqrt{\frac{3}{2} - 1}} \right)^\beta |f|\beta.
\]

Applying Lemma A.1 yields the required bound.

With Proposition A.2 in hand we may now prove Lemmas 5.10 and 5.11.

**Proof of Lemma 5.10**

As \( \lim_{\epsilon \to 0} \text{diam}(Q_\epsilon) = 0 \) there exists \( \epsilon_2 > 0 \) such that for every \( \epsilon \in (0, \epsilon_2] \) we have \( \text{diam}(Q_\epsilon) < \eta_0 \) and

\[
1 + \frac{\text{diam}(Q_\epsilon)}{(\eta_0 - \text{diam}(Q_\epsilon))} < \sqrt{\frac{d}{d-1}}.
\]

By [7] Section 8.5, page 236, this implies

\[
S(1, 1 + \frac{\text{diam}(Q_\epsilon)}{(\eta_0 - \text{diam}(Q_\epsilon))}) = S(\eta_0 - \text{diam}(Q_\epsilon), \eta_0) \leq 2d.
\]

The desired conclusion follows by Proposition A.2.

**Proof of Lemma 5.11**

If \( \text{diam}(Q) < \eta_0 \) then \( |E_Q f|_\beta < \infty \) by Proposition A.2. Alternatively, if \( \text{diam}(Q) \geq \eta_0 \), then repeatedly applying Lemma A.1 yields

\[
|E_Q f|_\beta = \sup_{|f|_\beta \leq 1} |E_Q f|_\beta \leq \sup \{ |E_Q f|_{\beta, 2 \text{diam}(Q)} : |f|_{\beta, 2 \text{diam}(Q)} \leq S(\eta_0, 2 \text{diam}(Q)) \}
\]

\[
\leq S(\eta_0, 2 \text{diam}(Q)) |E_Q f|_{\beta, 2 \text{diam}(Q)},
\]

which is finite by Proposition A.2 applied to the seminorm \( |\cdot|_{\beta, 2 \text{diam}(Q)} \) (i.e. when \( \eta_0 = 2 \text{diam}(Q) \)). In either case we have \( |E_Q f|_\beta < \infty \) and so, as \( |E_Q f|_{L^1} = 1 \), we have \( |E_Q f|_\beta < \infty \) too. As \( Q \) partitions \( \Omega \), for every \( f \in \mathbb{V}_\beta(\Omega) \) the support of \( E_Q f \) is a subset of \( \Omega \). Hence \( E_Q f \in \mathbb{V}_\beta(\Omega) \) for every \( f \in \mathbb{V}_\beta(\Omega) \) and so \( E_Q \in L(\mathbb{V}_\beta(\Omega)) \).
B Code

B.1 Variance

Listing 1: This function centres the observable $obs$ (defined by an anonymous MATLAB function e.g. the code snippet in Section 4.2) and computes the required first and second derivatives at zero to estimate the variance.

```matlab
function [ddLam,v,dv,dlam,ddlam]=variance(P,obs),

%P is a row-stochastic matrix
%obs is a pre-defined anonymous function representing the observable
%lam is the leading eigenvalue
%v is the leading eigenfunction
%dv is dv/dtheta, where theta is the twist parameter
%dlam is dlam/dtheta
%ddlam is d^2lam/dtheta^2
%ddLam is d^2Lam/dtheta^2

%% find v and normalise appropriately
n=size(P,1);
phi=ones(n,1)/n;
[v,~]=eigs(P',1);
v=v/sum(v)*n;

%% centre observable g
x=[1/(2*n):1/(n):1-1/(2*n)]';
g=obs(x);
g=g-g'*v/n; %ensure g has mean zero by subtracting the mean

%% estimate dlam and dv using 1.*v=1 and 1.*dv=0
A=[P' -speye(n) -v; ones(1,n) 0];
b=[-P'*(g.*v); 0];
y=A\b;
dv=y(1:n);
dlam=y(n+1);

%% compute d^2lam/dtheta^2 and d^2Lam/dtheta^2
ddlam=((g.^2)'*v+2*g'*dv)/n;
ddLam=(ddlam-dlam^2);
```

B.2 Rate function
Listing 2: This function centres the observable \( \text{obs} \) (defined by an anonymous MATLAB function e.g. the code snippet in Section 4.3), and performs the required minimisation to evaluate the rate function at points specified in the vector \( s \).

```matlab
function \([r,\text{optz}] = \text{rate\_function}(s,P,\text{obs})\)

%P is a row-stochastic matrix
%obs is a pre-defined anonymous function representing the observable
%s is a vector of arguments of the rate function

% calc acim for centering observable.
[v0,~]=eigs(P',1);
v0=v0/sum(v0);

% set up objects to pass to legendre\_function.m
n=length(P);
[I,J,V]=find(P);
xpts=(I-.5)/n;
xptsorig=1/(2*n):1/n:1-1/(2*n);
gmean=\text{obs}(xptsorig)*v0;

% set up arrays for r and optz and set optimisation options
r=zeros(length(s),1);
optz=r;
options = optimoptions('fminunc','Algorithm','quasi
−newton','SpecifyObjectiveGradient',true,'OptimalityTolerance',1e−6);

% initial seed point for minimisation
z0=0;

% evaluate rate function at points specified in s
for i=1:length(s),
    minfun=@(z)\text{legendre\_function}(z,P,v0,s(i),\text{obs},n,I,J,V,xpts,gmean);
    [\text{optz}(i),r(i),~,~]=fminunc(minfun,z0,options);
    z0=\text{optz}(i); %use previous optimum for next initialisation.
end
r=−r;
```

Listing 3: This function evaluates the “Legendre function” (the function to be minimised) and its derivative. This requires twisting the matrix \( P \) by \( z \) and then computing the leading eigenvalue and eigenvector of the twisted matrix.

```matlab
function \([f,\text{df}] = \text{legendre\_function}(z,P,v0,s,\text{fun},n,I,J,V,xpts,gmean)\)

%evaluate 'legendre' function with fixed parameter \( s \), maximising over \( z \).

% twist \( P \) by \( z \)
gvec = z*(fun(xpts)−gmean);
Vtwist = V.*exp(gvec);
Pt twist = sparse(I,J,Vtwist);

%% Calculate objective f
[v, lam] = powermethod(Ptwist, v0, 1);
f = log(lam)−z*s;

%% Calculate gradient df
if nargout > 1 % gradient required
v = v/sum(v);
[phi, lam] = powermethod(Ptwist', ones(n,1), 1);
phi = phi/(phi'*v);
gvec basic = fun(1/(2*n):1/n:1−1/(2*n))−gmean;
dlam = lam*phi'*(gvec basic'.*v);
df = dlam/lam−s;
end

Listing 4: Estimation of the leading eigenvalue and eigenvector by repeated iteration.

function [v1, lam1] = powermethod(P, v0, lam0),

% P is a row stochastic matrix
% v0 is an initial (guessed) eigenvector
% lam0 is an initial (guessed) eigenvalue
v0 = v0/sum(v0);
v1 = P'*v0;
lam1 = sum(v1);
while abs(lam1−lam0)>1e−15,
lam0 = lam1;
v0 = v1/lam1;
v1 = P'*v0;
lam1 = sum(v1);
end

References

[1] R. Aimino and S. Vaienti, A note on the large deviations for piecewise expanding multidimensional maps, in Nonlinear Dynamics New Directions, Springer, 2015, pp. 1–10.

[2] A. Ayyer, C. Liverani, and M. Stenlund, Quenched CLT for random toral automorphism, Discrete Contin. Dyn. Syst., 24 (2009), pp. 331–348.

[3] W. Bahsoun, Rigorous numerical approximation of escape rates, Nonlinearity, 19 (2006), p. 2529.
[4] 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.

[5] W. Bahsoun and M. Ruziboev, On the stability of statistical properties for Lorenz attractors with a C^{1+\alpha} stable foliation, arXiv preprint arXiv:1701.08601, (2017).

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

[7] K. Böröczky, Finite Packing and Covering, Cambridge tracts in mathematics, Cambridge University Press, 2004.

[8] C. Bose, G. Froyland, C. González-Tokman, and R. Murray, Ulam’s method for Lasota–Yorke maps with holes, SIAM Journal on Applied Dynamical Systems, 13 (2014), pp. 1010–1032.

[9] A. Boyarsky and P. Gora, Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension, Probability and its Applications, Birkhäuser Boston, 1997.

[10] A. Broise, Transformations dilatantes de l’intervalle et théorèmes limites, Astérisque, (1996), pp. 1–109. Études spectrales d’opérateurs de transfert et applications.

[11] M. Dellnitz, G. Froyland, and S. Sertl, On the isolated spectrum of the Perron-Frobenius operator, Nonlinearity, 13 (2000), pp. 1171–1188.

[12] M. Dellnitz and O. Junge, On the approximation of complicated dynamical behavior, SIAM Journal on Numerical Analysis, 36 (1999), pp. 491–515.

[13] J. Ding and A. Zhou, Finite approximations of Frobenius-Perron operators. A solution of Ulam’s conjecture to multi-dimensional transformations, Physica D, 92 (1996), pp. 61–68.

[14] S. Gouëzel, Berry–Esseen theorem and local limit theorem for non-uniformly expanding maps, in Annales de l’Institut Henri Poincare (B) Probability and Statistics, vol. 41, 2005, pp. 997–1024.

[15] ———, Almost sure invariance principle for dynamical systems by spectral methods, The Annals of Probability, 38 (2010), pp. 1639–1671.

[16] ———, Limit theorems in dynamical systems using the spectral method, vol. 89 of Proceedings of Symposia in Pure Mathematics, American Mathematical Society, 2015, pp. 161–193.

[17] P. Gruber, Convex and Discrete Geometry, Grundlehren der mathematischen Wissenschaften, Springer Berlin Heidelberg, 2010.

[18] Y. Guivarc’h and J. Hardy, Théorèmes limites pour une classe de chaînes de
Markov et applications aux difféomorphismes d’Anosov, Annales de l’IHP Probabilité et statistiques, 24 (1988), pp. 73–98.

[19] H. Hennion, *Sur un théorème spectral et son application aux noyaux lipchitziens*, Proceedings of the American Mathematical Society, 118 (1993), pp. 627–634.

[20] H. Hennion and L. Hervé, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, vol. 1766 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2001.

[21] F. Hofbauer and G. Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations*, Mathematische Zeitschrift, 180 (1982), pp. 119–140.

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

[23] P. L. Johnson and C. Meneveau, *Large-deviation joint statistics of the finite-time Lyapunov spectrum in isotropic turbulence*, Physics of Fluids, 27 (2015), p. 085110.

[24] T. Kato, *Perturbation theory for linear operators*, Grundlehren der mathematischen Wissenschaften, Springer Berlin Heidelberg, 1966.

[25] G. Keller, *Stochastic stability in some chaotic dynamical systems*, Monatshefte für Mathematik, 94 (1982), pp. 313–333.

[26] ———, *Generalized bounded variation and applications to piecewise monotonic transformations*, Probability Theory and Related Fields, 69 (1985), pp. 461–478.

[27] G. Keller and C. Liverani, *Stability of the spectrum for transfer operators*, Annali della Scuola Normale Superiore di Pisa-Classe di Scienze, 28 (1999), pp. 141–152.

[28] T.-Y. Li, *Finite approximation for the Frobenius-Perron operator. a solution to Ulam’s conjecture*, Journal of Approximation theory, 17 (1976), pp. 177–186.

[29] C. Liverani, *Decay of correlations for piecewise expanding maps*, Journal of Statistical Physics, 78 (1995), pp. 1111–1129.

[30] I. Melbourne and M. Nicol, *A vector-valued almost sure invariance principle for hyperbolic dynamical systems*, The Annals of Probability, (2009), pp. 478–505.

[31] R. Murray, *Existence, mixing and approximation of invariant densities for expanding maps on \( \mathbb{R}^r \)*, Nonlinear Analysis: Theory, Methods and Applications, 45 (2001), pp. 37–72.

[32] S. V. Nagaev, *Some limit theorems for stationary Markov chains*, Theory of Probability & Its Applications, 2 (1957), pp. 378–406.
[33] L. REY-BELLET AND L.-S. YOUNG, *Large deviations in non-uniformly hyperbolic dynamical systems*, Ergodic Theory and Dynamical Systems, 28 (2008), pp. 587–612.

[34] C. M. ROHWER, F. ANGELETTI, AND H. TOUCHETTE, *Convergence of large-deviation estimators*, Physical Review E, 92 (2015), p. 052104.

[35] J. ROUSSEAU-EGELE, *Un théoreme de la limite locale pour une classe de transformations dilatantes et monotones par morceaux*, The Annals of Probability, 11 (1983), pp. 772–788.

[36] B. SAUSSOL, *Absolutely continuous invariant measures for multidimensional expanding maps*, Israel Journal of Mathematics, 116 (2000), pp. 223–248.

[37] H. H. SCHAEFER, *Banach lattices*, in Banach Lattices and Positive Operators, Springer, 1974, pp. 46–153.

[38] C. T. I. TULCEA AND G. MARINESCU, *Theorie ergodique pour des classes d’operations non complete ment continues*, Annals of Mathematics, 52 (1950), pp. 140–147.

[39] S. M. ULAM, *A collection of mathematical problems*, Interscience Publishers, 1960.

[40] C. WORMELL, *Spectral Galerkin methods for transfer operators in uniformly expanding dynamics*, arXiv preprint arXiv:1705.04431, (2017).