Ergod. Th. & Dynam. Sys. (2019), **39**, 1121–1152
doi:10.1017/etds.2017.67

Robustness of ergodic properties of non-autonomous piecewise expanding maps

MATTEO TANZI†, TIAGO PEREIRA†‡ and SEBASTIAN VAN STRIEN†

† *Department of Mathematics, South Kensington Campus, London SW7 2AZ, UK*
(e-mail: m.tanzi13@imperial.ac.uk, tiago.pereira@imperial.ac.uk, s.van-strien@imperial.ac.uk)
‡ *Institute of Mathematical and Computer Sciences, Universidade de São Paulo, São Carlos 13566-590, São Paulo, Brazil*

(*Received* 25 *May* 2016 *and accepted in revised form* 1 *June* 2017)

Abstract. Recently, there has been an increasing interest in non-autonomous composition of perturbed hyperbolic systems: composing perturbations of a given hyperbolic map F results in statistical behaviour close to that of F. We show this fact in the case of piecewise regular expanding maps. In particular, we impose conditions on perturbations of this class of maps that include situations slightly more general than what has been considered so far, and prove that these are stochastically stable in the usual sense. We then prove that the evolution of a given distribution of mass under composition of time-dependent perturbations (arbitrarily—rather than randomly—chosen at each step) close to a given map F remains close to the invariant mass distribution of F. Moreover, for almost every point, Birkhoff averages along trajectories do not fluctuate wildly. This result complements recent results on memory loss for non-autonomous dynamical systems.

1. *Introduction*

During the last decade there has been an increasing focus on non-autonomous dynamical systems meaning that, rather than iterating a single map $F : M \to M$ on the given phase space M, one looks at the evolution under composition of several different self-maps of M. The main motivation for this point of view is that, in practical applications, the map F_t describing how the state variable evolves from time t to time $t + 1$ should depend on t. In this paper, the only assumption we make on F_t is that at each time it is close to a given map F. So, we will *not* assume that the choice of the maps F_t follows some random distribution nor that the system can be written as a skew product.

Non-autonomous composition of small perturbations of a given dynamical system can lead to the pointwise destruction of typical statistical behaviour. For example, in [**OGY90**] the authors showed that for any point in phase space one can construct a sequence of time-dependent perturbations that makes the point evolve arbitrarily close to a periodic orbit of the unperturbed map. This means that even if the Birkhoff averages along the unperturbed orbit were typical, this is not the case for the non-autonomous dynamics. We would like to answer the question of whether, under the same non-autonomous evolution, the change of statistical behaviour by small perturbations can be observed on a set of positive measure (with respect to the reference measure). We address the problem in the case of multidimensional piecewise maps [**Sau00**].

We first define a collection of perturbations to a given multidimensional piecewise expanding map F, and we prove that the system is stochastically stable for these types of perturbation (meaning that perturbed maps with perturbations of given small magnitude have an invariant density which is close to the invariant density for the unperturbed system). Then we prove that the evolution of sufficiently regular mass distributions under the time-dependent dynamics become, up to a fixed precision, close to the mass distribution which is invariant under F. We then use this result together with a law of large numbers for dependent random variables to prove that, given sufficiently regular observables, for almost every point the accumulation point of the sequence of Birkhoff averages is close to the expectation of the observable with respect to the invariant measure.

Our results can be applied to certain dynamical systems defined on networks whose topology slightly changes over time. In applications, we have in mind that some of the edges of the network are occasionally broken due to mechanical failures. Under certain settings we show that such intermittent mechanical failures do not significantly change the ergodic properties of the system.

The paper is organized as follows. In §2 we state the main results and discuss the existing literature. In §3 we give some applications of the main result, in particular to dynamics on networks with changing topologies. In §4 we give the precise definitions and prove the main theorem. The proof we use relies on spectral stability. In the Appendix, we formulate a related result for $C^{1+\nu}$-expanding maps, presenting a different approach relying on invariant cones. We provide an extensive outline of the ideas of the proofs of both results.

2. *Statements of the results*

We consider the class of maps introduced in [**Sau00**]. The phase space is $\Omega \subset \mathbb{R}^N$, a compact subset of $\mathbb{R}^N$, which is decomposed into a fixed number of domains (allowed to slightly change in the perturbed versions of the maps). The domains can have fractal boundaries, and the restriction to each of them is regular. The precise hypotheses that a map $F : \Omega \to \Omega$ and its perturbations must satisfy are given in properties (ME1)–(ME6) (§4) and (CM1)–(CM2) (§4.2). Under these assumptions, we obtain the following.

THEOREM A. *Let $F_{\hat{\gamma}}$ be a piecewise expanding map on the compact set $\Omega \subset \mathbb{R}^N$ belonging to a collection of maps $\{F_\gamma\}_{\gamma \in \Gamma}$ satisfying (ME1)–(ME6), continuous at $\hat{\gamma} \in \Gamma$ ((CM1)–(CM2)). Then, for each $\varepsilon > 0$, there exists $\delta > 0$ so that:*

(1) *if ν is a Borel probability measure on Γ with $\operatorname{supp} \nu \subset B_\delta(\hat{\gamma})$, then there exists φ_ν, a stationary density for the random dynamical system, obtained by composing*

independently maps of the family according to ν *and* φ_ν *satisfies*

$$\|\varphi_\nu - \varphi_{\hat{\gamma}}\|_1 \leq \varepsilon;$$

in particular, for all $\gamma \in B_\delta(\hat{\gamma})$, F_γ *has an invariant density* φ_γ *and*

$$\|\varphi_\gamma - \varphi_{\hat{\gamma}}\|_1 \leq \varepsilon;$$

(2) *if* $\boldsymbol{\gamma} \in B_\delta(\hat{\gamma})^{\mathbb{N}}$, *then, for every probability measure* $\mu = \varphi m$ *with density* $\varphi \in V_\alpha$ *(*$V_\alpha \subset L^1$ *is defined in (5)), there exists* $\overline{n} := \overline{n}(\varepsilon, \varphi) \in \mathbb{N}$ *such that for every* $n > \overline{n}$ *the Radon–Nikodym derivative* $d(F^n_{\boldsymbol{\gamma}\,*}\mu)/dm$ *has a representative in* V_α, *and*

$$\left\| \frac{d}{dm}(F^n_{\boldsymbol{\gamma}})_*\mu - \varphi_{\hat{\gamma}} \right\|_1 < \varepsilon; \tag{1}$$

(3) *moreover, for any observable* $\psi \in V_\alpha$, *there exists a set* $X_{\boldsymbol{\gamma}}$ *of full measure so that, for every* $x \in X_{\boldsymbol{\gamma}}$,

$$\int \psi \, d\mu_{\hat{\gamma}} - \varepsilon\|\psi\|_1 \leq \liminf_{n\to\infty} \frac{1}{n} S_n(\psi)(x) \leq \limsup_{n\to\infty} \frac{1}{n} S_n(\psi)(x) \leq \int \psi \, d\mu_{\hat{\gamma}} + \varepsilon\|\psi\|_1,$$

where $\mu_{\hat{\gamma}} = \varphi_{\hat{\gamma}} m$ *and* $S_n(\psi)(x) = \psi(x) + \sum_{i=1}^{n-1} \psi \circ F_{\gamma_i} \circ \cdots \circ F_{\gamma_1}(x)$.

Part (1) of the theorem proves stochastic stability of the collection and in particular proves continuous dependence of invariant measures with respect to the size of the perturbation. Part (2) shows that the evolution $(F^n_{\boldsymbol{\gamma}})_*\mu$ of the mass μ with density φ remains close to the invariant measure (indeed, this holds in the L^1 sense for the densities). For many applications one cannot be confident that $\boldsymbol{\gamma}$ is chosen randomly, and therefore the assertion from Part (2) is more useful than having merely stochastic stability. Part (3) shows that one has quasi-Birkhoff behaviour for time averages, meaning that the accumulation points of the time averages remain close to the space average of the observable with respect to the invariant measure of the unperturbed system.

Remark 1. Previous results on stochastic stability [**Cow00**] for piecewise expanding maps require the domains of the partition elements to have piecewise smooth boundaries.

Remark 2. In the Appendix we also formulate a related result in the context of $C^{1+\nu}$-expanding maps, using the contraction properties of the transfer operator on suitable cones. The improved regularity results in a uniform estimate, rather than the L^1 estimate in Part (2) above. We will discuss previous results, give an overview of the literature and also the ideas of the proofs at the end of this section.

Remark 3. Keller [**Kel82b**] obtained a result for the L^1-analogue of inequality (1) for piecewise expanding interval maps.

COROLLARY 1. *Let* $\mu_{\hat{\gamma}} := \varphi_{\hat{\gamma}} m$ *be the invariant measure for* $F_{\hat{\gamma}}$; *then for every neighbourhood* $\mathcal{U}(\mu_{\hat{\gamma}})$ *with respect to the weak topology there exists* $\delta > 0$ *such that for every sequence* $\boldsymbol{\gamma} \in B_\delta(\hat{\gamma})^{\mathbb{N}}$, *and for almost every* $x \in \Omega_{\boldsymbol{\gamma}}$, *there is* $\overline{n}$ *such that*

$$\frac{1}{n}\sum_{i=0}^{n-1} (F^n_{\boldsymbol{\gamma}})_*\delta_x \in \mathcal{U}(\mu_{\hat{\gamma}}) \quad \textit{for all } n > \overline{n}.$$

2.1. *Historical comments.* Certain classes of dynamical systems possess good statistical properties under the effect of small random perturbations. For example, stochastic perturbations (i.e. chosen randomly according to some distribution) of expanding maps on a compact manifold have been thoroughly investigated (e.g. [**Via97**, **AA03**, **AT05**, **BY93**]), as well as those of piecewise expanding maps on the interval (e.g. [**LY73**, **Kel82a**, **Liv95b**]). Some recent studies deal with maps of the interval with neutral fixed points [**SvS13**] and with multidimensional piecewise maps of compact subsets in $\mathbb{R}^N$ [**BG89**, **Cow00**, **Sau00**]. In all these cases one can give a description of the statistical behaviour of the orbits of the randomly perturbed system via a stationary measure which is also close to some absolutely continuous invariant density for the unperturbed system. Key to these results is that perturbations are independent and identically distributed. However, recent developments require the understanding of the asymptotic behaviour of dynamical systems under non-autonomous perturbations [**KR11**]. These perturbations are not independent, and the natural question concerns whether the statistics of the unperturbed and perturbed maps remain close in this more general setting. We provide an affirmative answer for two classes of dynamical systems: piecewise $C^{1+\nu}$ maps of a compact subset of $\mathbb{R}^N$ and (in the Appendix) $C^{1+\nu}$-expanding maps of a compact manifold.

Recent work (among others [**CR07**, **AHNT15**] and [**AR15**]) on non-autonomous composition of dynamical systems (sometimes also referred to as sequential dynamical systems) focused mainly on proving that the system exhibits memory loss, which roughly means that, given a finite precision, the orbits of sufficiently regular densities of states become indistinguishable after a finite number of iterations of the system. More precisely, if the density of the measures μ, ν belong to a suitable cone, then one has *memory loss*:

$$\left\| \frac{d}{dm}(F_\gamma^n)_*\mu - \frac{d}{dm}(F_\gamma^n)_*\nu \right\|_1 \to 0$$

as $n \to \infty$. Memory loss under non-autonomous perturbations holds also for example for contracting systems, where all orbits tend to get indefinitely close, thus losing track of their initial condition. It is easy to give examples where one has memory loss, where the densities $d(F_\gamma^n)_*\mu/dm$ strongly fluctuate. Part (2) of Theorem A shows that the size of the fluctuations, in the above setting, depends only on the size of the perturbation.

A work similar in spirit to ours is [**Kel82b**] where the author derived a result on perturbed operators satisfying the hypotheses of an ergodic theorem by Ionescu-Tulcea and Marinescu [**ITM50**, Theorem 1] and applied it to one-dimensional piecewise expanding maps. The same approach could be used to deal with the multidimensional case imposing conditions on the maps and their perturbations so that they fit into the hypotheses of the theorem. We follow a slightly different argument. Also in [**NSV12**] the authors provided general conditions on transfer operators and on observables that ensure the validity of a central limit theorem for Birkhoff sums. In [**HNTV17**] analogous conditions were provided for the almost sure invariance principle to hold. Other notable works are [**OSY09**, **Ste11**, **SYZ13**] where a coupling technique was used to prove exponential memory loss for the evolution of densities in the smooth expanding case, in the one-dimensional piecewise expanding case, in the two-dimensional Anosov case and in Sinai billiards with slowly moving scatterers. In [**DS16**] the authors looked at the non-autonomous composition of one-dimensional expanding maps which are changing very slowly in time. In this situation, they could describe Birkhoff sums and their fluctuations as diffusion processes in the adiabatic limit of very slow change. It is also worth noticing that in [**BKL02**] the authors defined Banach spaces that allowed them to give a complete picture of the spectral properties of Anosov systems. Combining this result with the result from [**Kel82b**], one can obtain robustness of the evolution of densities. While we were writing this paper we also came across [**GOT13**] where memory loss is discussed in the multidimensional piecewise expanding setting for strongly mixing systems, but using techniques closer to [**Liv95b**].

2.2. *Strategy of the proof.* To prove Theorem A one could proceed in a similar way looking at the application of a non-autonomous sequence of perturbed transfer operators on some invariant cone of functions with finite diameter. This is the approach followed in [**GOT13**] to prove memory loss for the non-autonomous composition. However, to proceed in this way, one needs to restrict to a composition of maps which have a strong mixing property on the whole phase space. This is required because otherwise the supports of the invariant density for the unperturbed map and for an arbitrarily small perturbation might not coincide, making the diameter of any cone of functions containing both densities infinite with respect to the Hilbert metric.

Since we want to treat the more general case we will use another technique that will work for any non-autonomous composition of perturbed versions of a piecewise expanding map with quasi-compact transfer operator having 1 as unique simple eigenvalue. The restriction of the transfer operators to a Banach subspace of L^1 made of quasi-Hölder functions, as defined in [**Sau00**], satisfies a Lasota–Yorke inequality, which implies quasi-compactness and the presence of a spectral gap. Each transfer operator thus induces a splitting on the space V_α that can be written as the direct sum of a one-dimensional eigenspace corresponding to the invariant density, and the subspace of quasi-Hölder functions with zero mean and the successive application of the transfer operator on a probability density makes it converge (with respect to the L^1 norm) exponentially fast to the associated invariant density. One then shows that the invariant density is stochastically stable, implying that sufficiently small perturbations of a given piecewise expanding map have invariant densities close in the L^1 norm. The main idea to prove the result is to follow the evolution of a probability density splitting at each iteration on the eigenspaces corresponding to the invariant densities, and to control the remainders that these projections introduce using consequences of the Lasota–Yorke inequality and spectral properties. We treat all this in §4 where we first introduce a precise definition of the maps we considered, followed by a preliminary section (§4.1) on spectral and perturbation results. We introduce perturbations and perturbative results in §4.2 to conclude with the proof of the main claim in §4.3.

An alternative way to prove Part (2) of Theorem A would have been to show that the setting taken from [**Sau00**] with the perturbation we introduce can be framed in the general theorem of [**Kel82b**]. To do so one should prove that the operators acting on the Banach subspace of L^1, V_α (or some possibly larger space) satisfy all the hypotheses of the

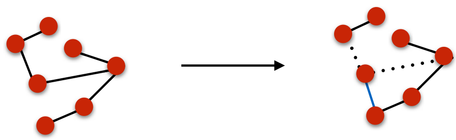

FIGURE 1. Networks with changing topology over time. During the time step depicted above, two edges are deleted (dotted lines) while one is added.

abstract theorem. This has been done by Keller in the one-dimensional case considering the properties of the restriction of transfer operators of one-dimensional piecewise expanding maps to the space of functions with bounded variation.

We deduce information on the decay of correlations of random variables $\psi_i = \psi \circ F_{\gamma_n} \circ \cdots \circ F_{\gamma_1}$ from the spectral properties. We use these properties to apply a strong law of large numbers [**Wal04**] for correlated random variables. This gives an expression for the accumulation points of the Birkhoff averages and enables us to obtain Part (3) of Theorem A.

The Appendix of this paper contains Theorem B, which treats the $C^{1+\nu}$-expanding setting using cones and the Hilbert metric. The outline of the proof is given in the Appendix.

3. *Applications*

Dynamics on non-stationary networks. A typical application of Theorem A we have in mind is that of dynamics on networks of fixed size n in which the topology of the network or the coupling strength is allowed to fluctuate over time. For example, consider the following dynamics:

$$x_i(t+1) = f(x_i(t)) + \alpha \sum_{j=1}^{n} A_{ij}(t) h_{ij}(t, x_j(t), x_i(t)) \quad \text{for } i = 1, \ldots, n, \ t \in \mathbb{N}, \tag{2}$$

where $f: T^n \to T^n$ is an expanding map on the n-dimensional torus, $x_i(t)$ describes the state of the ith node at time t, $\alpha \in \mathbb{R}$ describes the overall coupling strength, $A_{ij}(t) \in \{0, 1\}$ is the adjacency matrix of the network (so it describes whether or not node i is connected to node j) and $h_{ij}: \mathbb{N} \times T^n \times T^n \to T^n$ is a time-dependent coupling map. For the uncoupled case $\alpha = 0$, the system has an invariant measure which is absolutely continuous with respect to the Lebesgue measure. Our results show that, provided $|\alpha|$ is small, a regularly distributed collection of initial points remains almost uniformly distributed as time progresses, even when the topology of the network changes at each time step, as in Figure 1. Independence from time t is often an unreasonable assumption. If a connection between i and j is broken at time t, it most likely will take time to be fixed. For this reason our theorem would apply in this setting, whereas standard results on stochastic stability would not.

Stochastic stability does not imply robustness of measures. In this example we show that a stochastically stable dynamical system, that is, a system which admits a stationary measure under random independent perturbations, can have intricate behaviour under non-

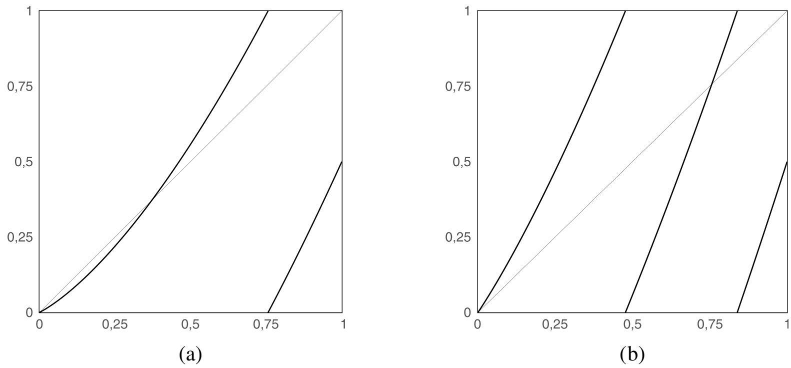

FIGURE 2. Perturbed Pomeau–Manneville maps of the circle (drawn on [0, 1] modulus 1). (a) Graph of $f_{-\varepsilon}$: 0 is an attracting fixed point. (b) Graph of f_ε: the map is uniformly expanding.

autonomous perturbations. For instance, take $\kappa \in (0, 1)$ and perturbations of the Pomeau–Manneville map on the circle

$$f_\gamma(x) = x + x^{1+\kappa} + \gamma x \mod 1$$

or of the Liverani–Saussol–Vaienti circle map

$$f_\gamma(x) = \begin{cases} x(1 + 2^\kappa x^\kappa) + \gamma x & \text{for } x \in [0, 1/2),\\ 2x - 1 + \gamma x & \text{for } x \in [1/2, 1); \end{cases}$$

see Figure 2 for an illustration. For $\gamma = 0$ these maps have an absolutely continuous invariant measure, but, for $\gamma < 0$ close to zero, f_γ has a stable attracting fixed point. In spite of this, Shen and van Strien [**SvS13**] proved that such maps are stochastically stable.

In other words, for almost every sequence $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots)$, chosen so that the γ_i are independent and identically distributed (i.i.d.) uniform random variables in an interval $[-\varepsilon, \varepsilon]$, the pushforward $(f^n_{\boldsymbol{\gamma}})_*(\mu)$ converges to an absolutely continuous measure μ_ε and as $\varepsilon > 0$ tends to 0 the density of this measure converges in the L^1 sense to the density of μ. Here it is crucial that the γ_i are chosen so as to be i.i.d. This is obvious because the pushforward of the Lebesgue measure λ under iterates of the map $f_{-\varepsilon}$ converges to the Dirac measure at the attracting fixed point of $f_{-\varepsilon}$.

One also can construct sequences $\boldsymbol{\gamma}$ so that $(f^n_{\boldsymbol{\gamma}})_*(\lambda)$ accumulates both to singular as well as to absolutely continuous invariant measures. Indeed, given a sequence of integers $0 = k_0 < k_1 < k_2 < \dots$, define $\boldsymbol{\gamma}$ so that, for all integers $i > 0$ and $j \geq 0$,

$$\gamma_i = \begin{cases} \varepsilon \text{ for } k_{2j} & < i \leq k_{2j+1},\\ -\varepsilon \text{ for } k_{2j+1} & < i \leq k_{2j+2}. \end{cases}$$

So, $f^n_{\boldsymbol{\gamma}}$ is a composition of the expanding maps f_ε (the second iterate of this map is expanding) and $f_{-\varepsilon}$, which has an attracting fixed point; see Figure 2. Provided we choose the sequence so that $k_{i+1} - k_i$ is sufficiently large compared to k_i, the sequence of measures $(f^n_{\boldsymbol{\gamma}})_*\lambda$ does not stay close to the absolutely continuous invariant measure of f. Indeed, for a suitable choice of the sequence k_i, for $n = k_{2i} \to \infty$ the measure $(f^n_{\boldsymbol{\gamma}})_*\lambda$ converges to the Dirac measure at the attracting fixed point for f_{γ_1}, while for

$n = k_{2i+1} \to \infty$ the measure $(f_\gamma^n)_*\lambda$ converges to the absolutely continuous invariant measure of f_ε (which is close to the absolutely continuous invariant measure of f when $\varepsilon > 0$ is small).

4. *Proof*

In this section we will consider piecewise expanding maps, and put ourselves in the setting of [**Sau00**].

4.0.1. *Assumptions on the maps.* Suppose that $\Omega \subset \mathbb{R}^N$ is a compact set with $\Omega \subset \text{Clos}(\text{Int}\,\Omega)$, and Γ is a metric space that will serve as the indexing set for the perturbations. Consider a collection of maps $\{F_\gamma\}_{\gamma\in\Gamma}$, $F_\gamma : \Omega \to \Omega$ for which there exist $k \in \mathbb{N}$, γ dependent partitions $\{U_\gamma^{(i)}\}_{1\le i\le k}$ of Ω, neighbourhoods $V^{(i)}$ of $U_\gamma^{(i)}$ for $i = 1, \ldots, k$ and $\gamma \in \Gamma$ and maps $F_\gamma^{(i)} : V^{(i)} \to \Omega$ such that, for every $\gamma \in \Gamma$:

(ME1) $F_\gamma|_{U_\gamma^{(i)}} = F_\gamma^{(i)}|_{U_\gamma^{(i)}}$ and $B_{\varepsilon_0}(F_\gamma(U_\gamma^{(i)})) \subset F_\gamma^{(i)}(V^{(i)})$ for all $i = 1, \ldots, k$;

(ME2) $F_\gamma^{(i)}$ is a $C^{1+\alpha}$ diffeomorphism, meaning that $F_\gamma^{(i)}$ is a C^1 diffeomorphism, and the Jacobian is uniformly Hölder, so for all $\varepsilon < \varepsilon_0$,

$$|\det D_x F_\gamma^{(i)^{-1}} - \det D_y F_\gamma^{(i)^{-1}}| \le c|\det D_z F_\gamma^{(i)^{-1}}|\varepsilon^\alpha, \quad x, y \in B_\varepsilon(z) \cap F_\gamma^{(i)}(V^{(i)});$$

(ME3) $m(\Omega \setminus \bigcup_i U_\gamma^{(i)}) = 0$;

(ME4) the map F_γ is expanding: there exists $s_\gamma \in (0, 1)$ such that for all $i \in \{1, \ldots, k\}$ for all $u, v \in F_\gamma^{(i)}(V^{(i)})$ with $d(u, v) < \varepsilon_0$, $d(F_\gamma^{(i)^{-1}}(u), F_\gamma^{(i)^{-1}}(v)) < s_\gamma d(u, v)$;

(ME5)

$$G^{(\gamma)}(x, \varepsilon) := \sum_i \frac{m(F_\gamma^{(i)^{-1}}(B_\varepsilon(\partial F_\gamma^{(i)} U_\gamma^{(i)})) \cap B_{(1-s_\gamma)\varepsilon})}{m(B_{(1-s_\gamma)\varepsilon}(x))},$$

$$G^{(\gamma)}(\varepsilon) := \sup_x G^{(\gamma)}(\varepsilon, x);$$

then

$$\sup_{\delta\le\varepsilon_0}\left[s_\gamma^\alpha + 2\sup_{\varepsilon\le\delta}\frac{G^{(\gamma)}(\varepsilon)}{\varepsilon^\alpha}\delta^\alpha\right] < \rho < 1,$$

where ρ does not depend on $\gamma \in \Gamma$;

(ME6) there is $\hat\gamma \in \Gamma$ such that the transfer operator of $F_{\hat\gamma}$ has a unique eigenfunction $\varphi_{\hat\gamma}$ in V_α, which is a Banach subspace of L^1 defined in (5) below.

Remark 4. For what concerns hypothesis (ME6), examples in the one-dimensional case of conditions that imply uniqueness of the eigenvalues can be found in several references (for example [**Via97**, **LM85**]). This condition is usually implied by the uniqueness of the absolutely continuous invariant measure plus a mixing condition.

Remark 5. Condition (ME5) requires uniformity of the upper bound on the function describing the complexity of the partition in relation to the expansion of the maps. The condition might seem rather artificial, but in [**GOT13**] it is proven that, considering maps satisfying (ME1)–(ME4), if their partition sets have piecewise C^2 boundaries with uniformly bounded C^2 norm and these hypersurfaces are in one-to-one correspondence and have small Hausdorff distance, then (ME5) is automatically satisfied whenever it is satisfied by one of the maps.

We first report some preliminary results from which we derive spectral properties of the transfer operators of any function F from the collection and its perturbations when they are restricted to the Banach subspace $V_\alpha \subset L^1(\mathbb{R}^N)$ of quasi-Hölder functions (defined in §4.1.3). For a definition of the transfer operator and a discussion of some of its properties, see §A.3. Under assumption (ME6), for each perturbed operator, the space of quasi-Hölder functions splits into the direct sum of invariant subspaces: one corresponds to the invariant direction while the restriction of the transfer operator to the other is a contraction. We conclude by showing how the presence of such a spectral gap implies the result.

4.1. *Preliminaries.* In [**GOT13**] the authors proved memory loss looking at the action of the maps on an invariant cone of functions by generalizing to multidimensional maps a construction introduced in [**Liv95b**] for the one-dimensional case. This procedure is the analogue in the piecewise case of what we present for $C^{1+\nu}$ maps in the Appendix, but it requires that all the maps have some mixing property on the whole phase space, which is always the case in the regular case, while it has to be assumed as an extra hypothesis in the piecewise case. Our argument allows us to drop this hypothesis, although it works only when considering composition of small perturbations of a given map.

We exploit a well-known property exhibited by the transfer operator of some maps: quasi-compactness [**BY93**, **Kel82a**, **Bal00**, **Kel82b**, **Liv95b**, **You99**, **You98**]. For a definition of the transfer operator and a discussion of some of its properties, see §A.3.1 in the Appendix.

4.1.1. *Stochastic stability: stationary case.* Now fix a Borel probability measure ν on the metric space Γ (endowed with the σ-algebra of Borel sets) and consider the asymptotic behaviour of the random trajectories $\{x_i\}_{i\in\mathbb{N}}$ with

$$x_i := F_{\gamma_i} \circ \cdots \circ F_{\gamma_1}(x_0) \quad \text{for all } i \in \mathbb{N},$$

where the $\{\gamma_i\}_{i\in\mathbb{N}}$ are sampled independently from Γ with distribution given by ν. This random dynamical system is equivalent to the skew-product dynamical system

$$\begin{aligned} \mathcal{F} : M \times \Gamma^{\mathbb{N}} &\to M \times \Gamma^{\mathbb{N}} \\ (x, \boldsymbol{\gamma}) &\to (f_{\gamma_1}(x), \sigma(\boldsymbol{\gamma})) \end{aligned}$$

on $(M \times \Gamma^{\mathbb{N}}, m \otimes \nu^{\otimes\mathbb{N}})$, where $\boldsymbol{\gamma} \in \Gamma^{\mathbb{N}}$ and σ is the left-sided shift. Denoting by $\mathcal{L}_\gamma$ the transfer operator associated to F_γ, one can define the *average transfer operator*,

$$\hat{\mathcal{L}}_\nu \varphi := \int_\Gamma \mathcal{L}_\gamma \varphi \, d\nu(\gamma).$$

The density φ_ν satisfying $\mathcal{L}_\nu \varphi_\nu = \varphi_\nu$ is called the stationary density. Whenever $\nu = \delta_\gamma$ for some Γ, $\hat{\mathcal{L}}_\nu = \mathcal{L}_\gamma$ and φ_ν (if it exists) is the invariant density under the map F_γ.

To prove Part (i) of Theorem A, we will use the spectral properties of $\mathcal{L}_{\hat{\gamma}}$ and $\mathcal{L}_\nu$ in a way similar to what has already been shown for one-dimensional piecewise maps in [**Via97**].

4.1.2. *Spectral theorems for the transfer operators.* It is often very useful to restrict the action of the transfer operator to some Banach space contained in L^1. It has been shown in many cases how such a restriction has nice spectral properties that imply, among others, existence of invariant absolutely continuous measures and exponential mixing of correlations between observables. Suppose that $V \subset L^1(M)$ and $(V, \|\cdot\|_V)$ is a Banach space. A theorem by Ionescu-Tulcea and Marinescu gives a criterion to establish the spectral properties of $\mathcal{L}$. We report here this theorem in the case where the ambient space is $L^1(M)$.

THEOREM 1. [**ITM50**] *Let $(V, \|\cdot\|_V)$ be a Banach closed subspace of $L^1(M)$ such that if $\{\varphi_n\}_{n\in\mathbb{N}} \subset V$ and $\|\varphi_n\|_V \le K$ is such that $\varphi_n \to \varphi$ in L^1, then $\varphi \in V$ and $\|\varphi\|_V \le K$. Let $\mathcal{C}(V)$ be the class of linear bounded operators with image in V satisfying the following:*
(1) *there exists H such that $|\mathcal{P}^n|_V \le H$ for all $n \in \mathbb{N}$;*
(2) *there exist $0 < r < 1$ and $R > 0$ such that*

$$\|\mathcal{P}\varphi\|_V \le r\|\varphi\|_V + R\|\varphi\|_1; \tag{3}$$

(3) *$\mathcal{P}(B)$ is compact in L^1 for every bounded B in $(V, \|\cdot\|_V)$.*
Then every $\mathcal{P} \in \mathcal{C}(V)$ has only a finite number of eigenvalues $\{c_1, \ldots, c_p\}$ of modulus 1 with finite-dimensional eigenspaces $\{X_1, \ldots, X_p\}$, and

$$\mathcal{P} = \sum_{i=1}^{p} c_i P_i + P_0,$$

where, if $\{\pi^{(i)}\}_{i=\{1,\ldots,p\}}$, $\pi^{(0)}$ are projections relative to the splitting,

$$V = \bigoplus_{i=1}^{p} X_i \oplus X_0,$$

$P_i := \mathcal{P} \circ \pi^{(i)}$ and $\|P_0^n\|_V = O(q^n)$ with $q \in (0, 1)$.

The theorem can be used to understand the behaviour of the transfer operator for a variety of maps. Most of the requirements are automatically satisfied by the transfer operator, and the only thing that requires an additional proof is inequality (3), often referred as a Lasota–Yorke type of inequality. For such an inequality to hold, the Banach space $(V, \|\cdot\|_V)$ must be chosen carefully.

In the following we need a result that deals with perturbed transfer operators. This is treated in various references and presented in different formulations. Among others we cite [**KL98**, **Kel82b**, **Via97**, **Bal00**]. We report the statement that can be found in [**Via97**] for transfer operators associated to piecewise expanding maps, and that can be generalized without any extra effort to the above setting.

THEOREM 2. [**Via97**] *Suppose that $(V, \|\cdot\|_V)$ is a closed Banach space which is a subspace of $L^1(M)$. Let $C > 0$, $q < 1$, $\lambda < 1$ and $\mathcal{P}_\varepsilon : V \to V$ be a family of linear operators satisfying:*
- *$\int \mathcal{P}_\varepsilon\varphi \, dm = \int \varphi \, dm$ and $\varphi \ge 0$ implies that $\mathcal{P}_\varepsilon\varphi \ge 0$;*
- *$\|\mathcal{P}_\varepsilon^n\varphi\|_V \le C\lambda^n\|\varphi\|_V + C\|\varphi\|_1$*

for every $n \ge 1$, $\varepsilon \ge 0$ and $\varphi \in V$. Suppose that:

- *for $n \geq 1$ there is $\varepsilon(n)$ so that, for all $\varphi \in V$ and all $\varepsilon \in (0, \varepsilon(n))$,*

$$\|\mathcal{P}_0^n \varphi - \mathcal{P}_\varepsilon^n \varphi\|_1 \leq C\lambda^n \|\varphi\|_V; \tag{4}$$

- $\mathrm{spec}(\mathcal{P}_0) = \{1\} \cup \Sigma_0$*, where 1 is a simple eigenvalue and* $\Sigma_0 \subset \{z \in \mathbb{C} : |z| \leq q\}$.

Fix $\widetilde{q} \in (\max\{\sqrt{q}, \sqrt{\lambda}\}, 1)$. Then, for any small enough $\varepsilon > 0$, $\mathrm{spec}(\mathcal{P}_\varepsilon) = \{1\} \cup \Sigma_\varepsilon$*, where 1 is a simple eigenvalue and $\Sigma_\varepsilon \subset \{z \in \mathbb{C} : |z| \leq \widetilde{q}\}$.*

The above theorem states that, under some hypotheses, when dealing with a quasi-compact transfer operator with 1 as unique simple eigenvalue, small perturbations do not jeopardize quasi-compactness.

4.1.3. *Quasi-Hölder spaces V_α.* In this section we report the definition of quasi-Hölder space as presented in [**Sau00**]. This is the Banach space on which we restrict the action of the Perron–Frobenius operator associated to F. Given $\varphi \in L^1(\mathbb{R}^N)$ and S a Borel subset of $\mathbb{R}^N$, define

$$\mathrm{osc}(\varphi, S) := \mathrm{Esup}_S\, \varphi - \mathrm{Einf}_S\, \varphi.$$

For all $\varepsilon > 0$ and $\varphi \in L^1(\mathbb{R}^N)$, the map $x \mapsto \mathrm{osc}(\varphi, B_\varepsilon(x))$ is measurable (in particular it is lower semi-continuous). Given $\alpha \in (0, 1)$, $|f|_\alpha$ is defined (finite or infinite) as

$$|\varphi|_\alpha := \sup_{0<\varepsilon\leq\varepsilon_0} \varepsilon^{-\alpha} \int_{\mathbb{R}^N} \mathrm{osc}(\varphi, B_\varepsilon(x))\, dm(x)$$

and V_α is

$$V_\alpha := \{\varphi \in L^1(\mathbb{R}^N) : |\varphi|_\alpha < \infty\}. \tag{5}$$

The space V_α, endowed with the norm $\|\cdot\|_\alpha := |\cdot|_\alpha + \|\cdot\|_1$, where $\|\cdot\|_1$ is the L^1 norm, is a Banach space.

4.1.4. *Spectral properties of $\mathcal{L}$ on V_α.* The transfer operator $\mathcal{L}$ associated to F satisfying (ME1)–(ME5) fulfils a Lasota–Yorke type of inequality.

PROPOSITION 1. [**Sau00**] *Suppose that F satisfies (ME1)–(ME5). If ε_0 is small enough, there exist $\eta \in (0, 1)$ and $C < 0$ such that, for all $\varphi \in V_\alpha$,*

$$\mathcal{L}\varphi \in V_\alpha \text{ and } |\mathcal{L}\varphi|_\alpha \leq \eta|\varphi|_\alpha + C \int_{\mathbb{R}^N} |\varphi|\, dm.$$

Theorem 1 by Ionescu-Tulcea and Marinescu gives the spectral properties of $\mathcal{L}$. Whenever the transfer operator $\mathcal{L}$ has $\{1\}$ as unique eigenvalue which is also simple, we obtain the following splitting.

PROPOSITION 2. *Let F be a map that satisfies (ME1)–(ME6). Then*

$$\mathrm{spec}(\mathcal{L}) = \{1\} \cup \Sigma_0,$$

where 1 is a simple eigenvalue and Σ_0 is a disc of radius $q < 1$, and

$$V_\alpha = \mathbb{R}\varphi_0 \oplus X_0.$$

4.2. *Perturbations.* In [**Cow00**] the author treats the problem of stochastic stability for the invariant density for multidimensional piecewise expanding maps with piecewise smooth boundaries of the partitions. We address the same problem in the setting presented above, which makes it natural to consider perturbations of the system $F : \Omega \to \Omega$ with different regularity partitions as long as the branches admit an extension to the same neighbourhood.

4.2.1. *Continuity assumptions.* Given any $\hat{\gamma} \in \Gamma$, we say that the collection $\{F_\gamma\}_{\gamma\in\Gamma}$ is continuous at $\hat{\gamma}$ if for every $\varepsilon > 0$ there is a $\delta > 0$ such that the δ-ball $B_\delta(\hat{\gamma})$ has the following properties:

(CM1) for each $\gamma_1, \gamma_2 \in B_\delta(\hat{\gamma})$, the C^1 distance between $F^{(i)}_{\gamma_1}$ and $F^{(i)}_{\gamma_2}$ is at most ε;

(CM2) for every $\gamma \in B_\delta(\hat{\gamma})$ and $1 \le i \le k$, $m(U^{(i)}_\gamma \triangle U^{(i)}) \le \varepsilon$, where $\triangle$ stands for the symmetric difference.

4.2.2. *Notation.* Define the set of multi-indices $\mathcal{I}_n := \{1, \dots, k\}^n$. For $\boldsymbol{i} \in \mathcal{I}_n$ and $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_n) \in \Gamma^n$, call

$$F^n_{\boldsymbol{\gamma}}(x) := F_{\gamma_n} \cdots F_{\gamma_1}(x) \quad \text{and} \quad F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}(x) := F^{(i_n)}_{\gamma_n} \cdots F^{i_1}_{\gamma_1}(x),$$

whenever it is well defined. For $\boldsymbol{\gamma} \in \Gamma^n$, define $\mathcal{L}^n_{\boldsymbol{\gamma}} := \mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_1}$. For all $n \in \mathbb{N}$ and $\boldsymbol{\gamma} \in \Gamma^n$, let us denote by $\{U^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}\}_{\boldsymbol{i}\in\mathcal{I}_n}$ the partition of Ω modulo a negligible subset such that, if $x \in U^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}$, then $x \in U^{i_1}_{\gamma_1}$, and $F^{(i_j)}_{\gamma_j} \cdots F^{(i_1)}_{\gamma_1}(x) \in U^{(i_{j+1})}_{\gamma_{j+1}}$ for $1 \le j < n$. Notice that $F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}$ is well defined on $U^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}$, and its restriction equals $F_{\boldsymbol{\gamma}}$.

We fix $\hat{\gamma} \in \Gamma$. From now on, $F_{\hat{\gamma}}$ represents the 'unperturbed' map, and will be denoted as F, and $\mathcal{L}_{\hat{\gamma}}$ will be denoted as $\mathcal{L}$.

Remark 6. Some of the $U^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}$ might be empty, or of measure zero.

4.2.3. *Perturbation results.* We now prove that if one randomly composes sufficiently small perturbations, then the average transfer operator of the perturbed map satisfies a uniform Lasota–Yorke type of inequality.

LEMMA 1. *Let $\{F_\gamma\}_{\gamma\in\Gamma}$ be perturbations of $F := F_{\hat{\gamma}}$ as in (ME1)–(ME6), continuous at $\hat{\gamma}$ as in (CM1)–(CM2). Then there exist $C > 0$, $\widetilde{\eta} \in (0, 1)$ and $\Gamma' \subset \Gamma$, a neighbourhood of $\hat{\gamma}$, such that, for every probability measure ν with* supp $\nu \subset \Gamma'$ *and all $\varphi \in V_\alpha$,*

$$|\hat{\mathcal{L}}_\nu \varphi|_\alpha \le \widetilde{\eta}|\varphi|_\alpha + C \int |\varphi|\, dm.$$

Proof. The proof of this lemma follows from Proposition 1. Since F_γ satisfies (ME1)–(ME5), Proposition 1 implies that it satisfies a Lasota–Yorke inequality with $\widetilde{\eta}(\gamma) \in (0, 1)$, and $C(\gamma) > 0$. As proved in [**Sau00**], $C(\gamma)$ has a uniform bound for every $\gamma \in \Gamma$ and

$$\eta(\gamma) = (1 + c s^\alpha_\gamma \varepsilon^\alpha_0)\rho.$$

Choosing ε_0 so that $\eta(\hat{\gamma}) < 1$, since $s_\gamma \to s_{\hat{\gamma}}$ for $\gamma \to \hat{\gamma}$ by (CM1), one can pick $\Gamma' \subset \Gamma$, a neighbourhood of $\hat{\gamma}$, such that $\eta(\gamma) < \widetilde{\eta} < 1$ for all $\gamma \in \Gamma'$. This implies that, for all

$\gamma \in \Gamma'$,

$$|\mathcal{L}_\gamma \varphi|_\alpha \le \widetilde{\eta} |\varphi|_\alpha + C \int |\varphi| \, dm.$$

Now

$$|\hat{\mathcal{L}}_\nu \varphi|_\alpha = \left| \int_\Gamma \mathcal{L}_\gamma \varphi \, d\nu(\gamma) \right|_\alpha = \sup_{0<\varepsilon\le\varepsilon_0} \varepsilon^{-\alpha} \int_{\mathbb{R}^N} \operatorname{osc}\left(\int_\Gamma \mathcal{L}_\gamma \varphi \, d\nu(\gamma), \, B_\varepsilon(x) \right) dm(x)$$

and, from the definition of oscillation, since Esup and Einf are respectively a convex and a concave function on the essentially bounded functions,

$$\begin{aligned}
|\hat{\mathcal{L}}_\nu \varphi|_\alpha &\le \sup_{0<\varepsilon\le\varepsilon_0} \varepsilon^{-\alpha} \int_{R^N} \int_\Gamma \operatorname{osc}(\mathcal{L}_\gamma \varphi, \, B_\varepsilon(x)) \, d\nu(\gamma) \, dm(x) \\
&\le \sup_{0<\varepsilon\le\varepsilon_0} \int_\Gamma \varepsilon^{-\alpha} \int_{R^N} \operatorname{osc}(\mathcal{L}_\gamma \varphi, \, B_\varepsilon(x)) \, dm(x) \, d\nu(\gamma) \\
&\le \int_\Gamma |\mathcal{L}_\gamma \varphi|_\alpha \, d\nu(\gamma) \\
&\le \widetilde{\eta} |\varphi|_\alpha + C \int |\varphi| \, dm.
\end{aligned}$$

□

Since $\widetilde{\eta} < 1$, one immediately gets the following lemma.

LEMMA 2.

(1) *For all $n \in \mathbb{N}$ and $\boldsymbol{\gamma} \in (\Gamma')^n$,*

$$|\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_1} \varphi|_\alpha \le \widetilde{\eta}^n |\varphi|_\alpha + \frac{C}{1-\widetilde{\eta}} \int |\varphi| \, dm,$$

which is thus uniformly bounded on n.

(2) *For all $\varepsilon > 0$ and all densities $\varphi \in V_\alpha$, there is a $\widetilde{n}(\varphi, \varepsilon) \in \mathbb{N}$ such that*

$$|\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_1} \varphi|_\alpha \le \frac{C}{1-\widetilde{\eta}} + \varepsilon$$

for all $n > \widetilde{n}(\varphi, \varepsilon)$.

We now prove a perturbation estimate crucial in determining the spectral properties of the perturbed transfer operators using similar estimates to the procedure in [**Via97**].

PROPOSITION 3. *Let $\{F_\gamma\}_{\gamma\in\Gamma}$ be perturbations of $F := F_{\hat{\gamma}}$ as in (ME1)–(ME6), continuous at $\hat{\gamma}$ as in (CM1)–(CM2); then there exist $0 < \widetilde{s} \le 1$ and $\widetilde{C} > 0$ such that for all $n \in \mathbb{N}$ there is $\delta > 0$ satisfying*

$$\|\mathcal{L}^n_{\boldsymbol{\gamma}} \varphi - \mathcal{L}^n \varphi\|_1 \le \widetilde{C} \widetilde{s}^n \|\varphi\|_\alpha \quad \textit{for all } \varphi \in V_\alpha \tag{6}$$

for all $\boldsymbol{\gamma} \in B_\delta(\hat{\gamma})^n$.

Proof.

$$\begin{aligned}
\int |\mathcal{L}^n_{\boldsymbol{\gamma}}\varphi(x) - \mathcal{L}^n\varphi(x)|\, dm(x)
&\le \sum_{\boldsymbol{i}\in\mathcal{I}_n}\Bigg[\int_{F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}(U^{(\boldsymbol{i})}_{\boldsymbol{\gamma}})\cap F^{(\boldsymbol{i})}(U^{(\boldsymbol{i})})} |(\varphi|\det DF^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}|^{-1})\circ(F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}})^{-1}(x)\\
&\quad - (\varphi|\det DF^{(\boldsymbol{i})}|^{-1})\circ {F^{(\boldsymbol{i})}}^{-1}(x)|\, dm(x)\\
&\quad + \int_{F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}(U^{(\boldsymbol{i})}_{\boldsymbol{\gamma}})\setminus F^{(\boldsymbol{i})}(U^{(\boldsymbol{i})})} \left|(\varphi|\det DF^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}|^{-1})\circ(F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}})^{-1}(x)\right| dm(x)\\
&\quad + \int_{F^{(\boldsymbol{i})}(U^{(\boldsymbol{i})})\setminus F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}(U^{(\boldsymbol{i})}_{\boldsymbol{\gamma}})} \left|(\varphi|\det DF^{(\boldsymbol{i})}|^{-1})\circ(F^{(\boldsymbol{i})})^{-1}(x)\right| dm(x)\Bigg]\\
&=: \sum_{\boldsymbol{i}\in\mathcal{I}_n}[(A)_{\boldsymbol{i}} + (B)_{\boldsymbol{i}} + (C)_{\boldsymbol{i}}].
\end{aligned}$$

We first treat $(A)_{\boldsymbol{i}}$ and then $(B)_{\boldsymbol{i}}$ with $(C)_{\boldsymbol{i}}$ for which analogous arguments hold. Notice that

$$\begin{aligned}
&\left|(\varphi\left|\det DF^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}\right|^{-1})\circ(F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}})^{-1} - (\varphi\left|\det DF^{(\boldsymbol{i})}\right|^{-1})\circ {F^{(\boldsymbol{i})}}^{-1}\right|\\
&\le \left|\varphi\circ(F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}})^{-1} - \varphi\circ {F^{(\boldsymbol{i})}}^{-1}\right|\left|\det D{F^{(\boldsymbol{i})}}^{-1}\right|\\
&\quad + \left|\varphi\circ(F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}})^{-1}\right|\left|\det D{F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}}^{-1} - \det D{F^{(\boldsymbol{i})}}^{-1}\right|,
\end{aligned}$$

which upper bounds $(A)_i$ with the sum of two terms. The first one is

$$(1)_i := \int_{F^{(i)}_{\boldsymbol{\gamma}}(U^{(i)}_{\boldsymbol{\gamma}})\cap F^{(i)}(U^{(i)})} |\varphi\circ(F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}})^{-1}(x) - \varphi\circ {F^{(\boldsymbol{i})}}^{-1}(x)||\det D{F^{(\boldsymbol{i})}}^{-1}|\, dm(x)$$

and one can choose a suitable $\delta > 0$ such that, for all $\gamma \in B_\delta(\hat{\gamma})$ and all $i \in \{1, \dots, k\}$,

$$|{F^{(i)}_\gamma}^{-1} - {F^{(i)}}^{-1}| < \xi(\delta)$$

with $\xi(\delta) < \varepsilon_0(1-s)$; so, by induction,

$$\begin{aligned}
|{F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}}^{-1}(x) - {F^{(\boldsymbol{i})}}^{-1}(x)| &\le |{F^{(i_1)}_{\gamma_1}}^{-1}\circ {F^{(i_2,\dots,i_n)}_{(\gamma_2,\dots,\gamma_n)}}^{-1}(x) - {F^{(i_1)}}^{-1}\circ {F^{(i_2,\dots,i_n)}_{(\gamma_2,\dots,\gamma_n)}}^{-1}(x)|\\
&\quad + |{F^{(i_1)}}^{-1}\circ {F^{(i_2,\dots,i_n)}_{(\gamma_2,\dots,\gamma_n)}}^{-1}(x) - {F^{(i_1)}}^{-1}\circ {F^{(i_2,\dots,i_n)}}^{-1}(x)|\\
&\le \xi(\delta) + s|{F^{(i_2,\dots,i_n)}_{(\gamma_2,\dots,\gamma_n)}}^{-1}(x) - {F^{(i_2,\dots,i_n)}}^{-1}(x)|
\end{aligned}$$

for all $x \in F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}(U^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}) \cap F^{(\boldsymbol{i})}(U^{(\boldsymbol{i})})$, yielding

$$|{F^{(\boldsymbol{i})}_{\boldsymbol{\gamma}}}^{-1} - {F^{(\boldsymbol{i})}}^{-1}| \le \xi(\delta)\frac{1}{1-s}.$$

This implies that, for any fixed ε, choosing δ so that $\xi(\delta)/(1-s) < \varepsilon$,

$$\begin{aligned}
(1)_i &\le \int_{F^{(i)}(U^{(i)})} \operatorname{osc}(\varphi, B_\varepsilon({F^{(\boldsymbol{i})}}^{-1}(x)))|\det D{F^{(\boldsymbol{i})}}^{-1}|\, dm(x)\\
&\le \int_{U^{(i)}} \operatorname{osc}(\varphi, B_\varepsilon(y))\, dm(y).
\end{aligned}$$

Taking the sum over $\mathcal{I}_n$,

$$\sum_{\boldsymbol{i}\in\mathcal{I}_n}\int_{U^{(\boldsymbol{i})}} \operatorname{osc}(\varphi,\, B_\varepsilon(y))\, dm(y) \le \int_{\mathbb{R}^N} \operatorname{osc}(\varphi,\, B_\varepsilon(y))\, dm(y) \le \varepsilon^\alpha |\varphi|_\alpha.$$

For the second term,

$$\begin{aligned}(2)_{\boldsymbol{i}} &:= \int_{F_{\boldsymbol{\gamma}}^{(\boldsymbol{i})}(U_{\boldsymbol{\gamma}}^{(\boldsymbol{i})})\cap F^{(\boldsymbol{i})}(U^{(\boldsymbol{i})})} |\varphi\circ (F_{\boldsymbol{\gamma}}^{(\boldsymbol{i})})^{-1}||\det DF_{\boldsymbol{\gamma}}^{(\boldsymbol{i})^{-1}} - \det DF^{(\boldsymbol{i})^{-1}}|\, dm\\ &\le m(\Omega)\xi(\delta)\operatorname{Esup}_\Omega |\varphi|,\end{aligned}$$

where $\xi(\delta)$ is a number that can be made arbitrarily small restricting δ. From compactness of Ω, there exists a $\overline{x}$ such that

$$\begin{aligned}\operatorname{Esup}_\Omega |\varphi| &= \operatorname{Esup}_{B_{\varepsilon_0/2}(\overline{x})} |\varphi|\\ &\le \frac{1}{m(B_{\varepsilon_0/2}(x))}\int_{B_{\varepsilon_0/2}(x)} [|\varphi(y)| + \operatorname{osc}(\varphi,\, B_{\varepsilon_0/2}(y))]\, dm(y).\end{aligned} \tag{7}$$

One obtains

$$(2)_{\boldsymbol{i}} \le C'\xi(\delta)\|\varphi\|_\alpha,$$

from which

$$\sum_{\boldsymbol{i}\in\mathcal{I}_n}(A)_{\boldsymbol{i}} \le \varepsilon^\alpha|\varphi|_\alpha + C'(\#\mathcal{I}_n)\xi(\delta)\|\varphi\|_\alpha.$$

For what concerns $(B)_{\boldsymbol{i}}$,

$$\begin{aligned}(B)_{\boldsymbol{i}} &= \int_{F_{\boldsymbol{\gamma}}^{(\boldsymbol{i})}(U_{\boldsymbol{\gamma}}^{(\boldsymbol{i})})\setminus F^{(\boldsymbol{i})}(U^{(\boldsymbol{i})})} |(\varphi \det DF_{\boldsymbol{\gamma}}^{(\boldsymbol{i})^{-1}})\circ (F_{\boldsymbol{\gamma}}^{(\boldsymbol{i})})^{-1}(x)|\, dm\\ &\le m(F_{\boldsymbol{\gamma}}^{(\boldsymbol{i})}(U_{\boldsymbol{\gamma}}^{(\boldsymbol{i})})\setminus F^{(\boldsymbol{i})}(U^{(\boldsymbol{i})}))\widetilde{s}^n \operatorname{Esup}_\Omega |\varphi|\\ &\le \xi(\delta)C''\widetilde{s}_\gamma^n\|\varphi\|_\alpha,\end{aligned}$$

where we upper bounded $m(F_{\boldsymbol{\gamma}}^{(\boldsymbol{i})}(U_{\boldsymbol{\gamma}}^{(\boldsymbol{i})})\setminus F^{(\boldsymbol{i})}(U^{(\boldsymbol{i})}))$ with $\xi(\delta)$ that thanks to (CM2) can be made arbitrarily small reducing δ. Summing all the contributions,

$$\sum_{\boldsymbol{i}\in\mathcal{I}_n}(B)_{\boldsymbol{i}} \le (\#\mathcal{I}_n)\xi(\delta)C''\widetilde{s}_\gamma^n\|\varphi\|_\alpha.$$

The sum of the $(C)_{\boldsymbol{i}}$ terms can be upper bounded analogously. As already pointed out, for a smaller δ we can make the upper bound arbitrarily small. This allows us, in particular, to obtain an exponential upper bound as (6) with respect to some $\widetilde{s}\in(0,\,1)$. □

We can generalize the above proposition to the case of averaged transfer operators.

PROPOSITION 4. *There exist* $0<\widetilde{s}\le 1$ *and* $\widetilde{C}>0$ *such that for all* $n\in\mathbb{N}$ *there is* $\delta>0$ *satisfying*

$$\|\hat{\mathcal{L}}_\nu^n\varphi - \mathcal{L}^n\varphi\|_1 \le \widetilde{C}\widetilde{s}^n\|\varphi\|_\alpha \quad \textit{for all } \varphi\in V_\alpha$$

for every probability measure ν *with* $\operatorname{supp}\nu\subset B_\delta(\hat{\gamma})$.

Proof.

$$\|\hat{\mathcal{L}}_\nu^n\varphi - \mathcal{L}^n\varphi\|_1 \le \int \left| \int_{\Gamma^n} \mathcal{L}_{\boldsymbol{\gamma}}^n\varphi(x)\, d\nu^{\otimes n}(\boldsymbol{\gamma}) - \mathcal{L}^n\varphi(x)\right| dm(x)$$
$$\le \int\int_{\Gamma^n} |\mathcal{L}_{\boldsymbol{\gamma}}^n\varphi(x) - \mathcal{L}^n\varphi(x)|\, d\nu^{\otimes n}(\boldsymbol{\gamma})\, dm(x)$$
$$\le \int_{\Gamma^n} \|\mathcal{L}_{\boldsymbol{\gamma}}^n\varphi - \mathcal{L}^n\varphi\|_1\, d\nu^{\otimes n}(\boldsymbol{\gamma}).$$

Since ν is supported on $B_\delta(\hat{\gamma})$, almost every sequence $\boldsymbol{\gamma}$ in the above integral will belong to $B_\delta(\hat{\gamma})^n$, allowing a direct application of Proposition 3. □

The above proposition and Proposition 2 give the spectral properties for the perturbed transfer operators.

PROPOSITION 5. *There is a neighbourhood Γ'' of $\hat{\gamma}$, $\Gamma'' \subset \Gamma$, such that*

$$\operatorname{spec} \mathcal{L}_\nu = \{1\} \cup \Sigma_0 \quad V_\alpha = \mathbb{R}\varphi_\nu \oplus X_0$$

for all probability measures ν with supp $\nu \subset \Gamma''$, *with Σ_0 inside a disc of radius $\widetilde{q} \in (0, 1)$, and where φ_ν is the unique stationary density. The projections associated to the splitting are*

$$\pi_1^{(\nu)}\varphi = \left(\int \varphi\, dm\right)\varphi_\nu \quad \pi_0^{(\nu)}\varphi = \varphi - \left(\int \varphi\, dm\right)\varphi_\nu.$$

Remark 7. As already remarked, the stationary measure associated with $\nu = \delta_\gamma$, $\gamma \in \Gamma''$, is the invariant measure for the map F_γ.

This proposition is a direct consequence of Theorem 2. One can easily prove that the invariant densities have uniformly bounded norms.

LEMMA 3. *For all probability measures ν with* supp $\nu \subset \Gamma'$,

$$|\varphi_\nu|_\alpha \le \frac{C}{1-\widetilde{\eta}}.$$

Proof. From Lemma 1,

$$|\varphi_\nu|_\alpha = |\mathcal{L}_\nu\varphi_\nu|_\alpha \le \widetilde{\eta}|\varphi_\nu|_\alpha + C,$$

which implies that

$$|\varphi_\nu|_\alpha \le \frac{C}{1-\widetilde{\eta}}.$$

□

4.3. *Proof of Theorem A.* We first prove Part (1) of Theorem A along the same lines as in [**Via97**], where it is proven for one-dimensional maps.

Proof of Part (1) of Theorem A. By the triangle inequality, for all $n \in \mathbb{N}$ and ν satisfying supp $\nu \subset B_\delta(\hat{\gamma})$,

$$\begin{aligned}\|\varphi_\nu - \varphi_{\hat{\gamma}}\|_1 &\le \|\varphi_\nu - \mathcal{L}_\nu^n\varphi_{\hat{\gamma}}\|_1 + \|\mathcal{L}_\nu^n\varphi_{\hat{\gamma}} - \varphi_{\hat{\gamma}}\|_1\\ &\le \|\mathcal{L}_\nu^n\varphi_\nu - \mathcal{L}_\nu^n\varphi_{\hat{\gamma}}\|_1 + \|\mathcal{L}_\nu^n\varphi_\nu - \mathcal{L}^n\varphi_{\hat{\gamma}}\|_1\\ &\le \widetilde{q}^n\|\varphi_\nu - \varphi_{\hat{\gamma}}\|_\alpha + \widetilde{C}\widetilde{s}^n\|\varphi_{\hat{\gamma}}\|_\alpha, \end{aligned} \tag{8}$$

where in (8) for the first term we used the spectral splitting and the fact that $\varphi_\nu - \varphi_{\hat{\gamma}} \in X_0$, and for the second term we used Proposition 3. By Lemma 3, $\|\varphi_\nu - \varphi_{\hat{\gamma}}\|_\alpha$ is uniformly bounded for $\gamma \in \Gamma'$, and this implies the result choosing n sufficiently large and adjusting $\delta > 0$ accordingly. □

We now prove Part (2) of Theorem A. In the proof, we denote by $\varphi_\gamma \in V_\alpha$ the unique invariant probability density for the map F_γ, where $\gamma \in \Gamma'$.

Proof of Part (2) of Theorem A. We can restate the theorem in terms of the action of the transfer operators on the density of the initial mass distribution. We shall then prove that, for all densities $\varphi \in V_\alpha$ and every $\varepsilon > 0$, there exist $\overline{n} := \overline{n}(\varepsilon, \varphi) \in \mathbb{N}$ and δ independent of φ such that, for all $n \geq \overline{n}$ and for all sequences $\boldsymbol{\gamma} \in B_\delta(\hat{\gamma})^n$,

$$\|\mathcal{L}^n_{\boldsymbol{\gamma}}\varphi - \varphi_{\hat{\gamma}}\|_1 \leq \varepsilon.$$

Let us consider the application of the transfer operators on their arguments split by projections $\pi_1^{(\gamma)}$ and $\pi_0^{(\gamma)}$. For example,

$$\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_1}\varphi = \mathcal{L}_{\gamma_n}\pi_0^{(\gamma_n)}(\mathcal{L}_{\gamma_{n-1}} \cdots \mathcal{L}_{\gamma_1}\varphi) + \mathcal{L}_{\gamma_n}\pi_1^{(\gamma_n)}(\mathcal{L}_{\gamma_{n-1}} \cdots \mathcal{L}_{\gamma_1}\varphi).$$

By induction,

$$\mathcal{L}_{\gamma_n}\mathcal{L}_{\gamma_{n-1}} \cdots \mathcal{L}_{\gamma_1}\varphi = \sum_{(i_1,\ldots,i_n)\in\{0,1\}^n} \mathcal{L}_{\gamma_n}\pi_{i_n}^{(\gamma_n)}\mathcal{L}_{\gamma_{n-1}}\pi_{i_{n-1}}^{(\gamma_{n-1})} \cdots \mathcal{L}_{\gamma_1} \circ \pi_{i_1}^{(\gamma_1)}\varphi.$$

Since $\pi_0^{(\gamma)}$ projects on the X_0 space, which is invariant under $\mathcal{L}_\gamma$, in the above sum only n terms give a non-zero contribution: after projecting on X_0 the action of the operators does not leave this space and if we later project on any of the $\mathbb{R}\varphi_{\gamma'}$ we obtain zero. This implies that only the non-increasing sequences of $\{0, 1\}^n$ may correspond to non-vanishing terms:

$$\mathcal{L}_{\gamma_n}\mathcal{L}_{\gamma_{n-1}} \cdots \mathcal{L}_{\gamma_1}\varphi = \sum_{\substack{(i_1, \ldots, i_n) \in \{0, 1\}^n \\ i_j \geq i_{j+1}}} \mathcal{L}_{\gamma_n}\pi_{i_n}^{(\gamma_n)}\mathcal{L}_{\gamma_{n-1}}\pi_{i_{n-1}}^{(\gamma_{n-1})} \cdots \mathcal{L}_{\gamma_1}\pi_{i_1}^{(\gamma_1)}\varphi.$$

In the above sum:

- the term with $(0, \ldots, 0, 1, \ldots, 1)$ ($i_j = 0$ and $i_{j-1} = 1$) equals
$$\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_j}(\varphi_{\gamma_{j-1}} - \varphi_{\gamma_j});$$
- the term $(0, \ldots, 0)$ equals φ_{γ_n};
- the term $(1, \ldots, 1)$ equals
$$\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_1}(\varphi - \varphi_{\gamma_1}).$$

We can now evaluate the L^1 norm of the following difference:

$$\begin{aligned}
&\|\mathcal{L}_{\gamma_n}\mathcal{L}_{\gamma_{n-1}} \cdots \mathcal{L}_{\gamma_1}\varphi - \varphi_{\hat{\gamma}}\|_1 \\
&\quad\leq \|\varphi_{\gamma_n} - \varphi_{\hat{\gamma}}\|_1 + \|\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_1}(\varphi - \varphi_{\gamma_1})\|_1 + \sum_{j=1}^{n} \|\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_j}(\varphi_{\gamma_{j-1}} - \varphi_{\gamma_j})\|_1 \\
&\quad\leq \|\varphi_{\gamma_n} - \varphi_{\hat{\gamma}}\|_1 + \|\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_1}(\varphi - \varphi_{\gamma_1})\|_1 + \sum_{j=1}^{n} \|\varphi_{\gamma_{j-1}} - \varphi_{\gamma_j}\|_1,
\end{aligned}$$

where in the last inequality we used (P3).

Step 1 For every $\varepsilon > 0$, there exist $\delta' > 0$ and $\overline{n} = \overline{n}(\delta', \varphi)$ such that, for all $k \in \mathbb{N}$ and for all $\boldsymbol{\gamma} \in B_{\delta'}(\hat{\gamma})^{kN}$,

$$\|\mathcal{L}_{\gamma_{k\overline{n}}}\mathcal{L}_{\gamma_{k\overline{n}-1}} \cdots \mathcal{L}_{\gamma_1}\varphi - \varphi_{\hat{\gamma}}\|_1 \leq \varepsilon. \tag{9}$$

We proceed by induction on k. Fix $\varepsilon' < \varepsilon/(2\overline{n} + 2)$ and δ' such that $\|\varphi_\gamma - \varphi_{\gamma'}\|_1 \leq \varepsilon'$ for all $\gamma, \gamma' \in B_{\delta'}(\hat{\gamma})$. In view of Lemma 2, there exist a constant K and $\widetilde{n}(\varphi, \varepsilon')$ such that, for all $n \geq \widetilde{n}$ and all $\gamma_1 \in B_{\delta'}(\hat{\gamma})$,

$$\|\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_1}(\varphi - \varphi_{\gamma_1})\|_\alpha < M.$$

Choose $\overline{n} \geq \widetilde{n}$ such that

$$\widetilde{q}^{\overline{n}-\widetilde{n}} M \leq \varepsilon/2. \tag{10}$$

For $\boldsymbol{\gamma} \in B_{\delta'}(\hat{\gamma})^{\overline{n}}$,

$$\begin{aligned}\|\mathcal{L}_{\gamma_{\overline{n}}}\mathcal{L}_{\gamma_{\overline{n}-1}} \cdots \mathcal{L}_{\gamma_1}(\varphi - \varphi_{\hat{\gamma}})\|_1 &\leq \varepsilon' + \widetilde{q}^{\overline{n}-\widetilde{n}}\|\mathcal{L}_{\gamma_{\widetilde{N}}} \cdots \mathcal{L}_{\gamma_1}(\varphi - \varphi_{\gamma_1})\|_\alpha + \overline{n}\varepsilon' \\ &\leq (\overline{n} + 1)\varepsilon' + \widetilde{q}^{\overline{n}-\widetilde{n}} M \\ &\leq \varepsilon/2 + \varepsilon/2 \leq \varepsilon.\end{aligned}$$

Now call $\varphi' := \mathcal{L}_{\gamma_{(k-1)\overline{n}}} \cdots \mathcal{L}_{\gamma_1}\varphi$. By the same argument,

$$\begin{aligned}\|\mathcal{L}_{\gamma_{k\overline{n}}}\mathcal{L}_{\gamma_{k\overline{n}-1}} \cdots \mathcal{L}_{\gamma_{\overline{n}(k-1)+1}}\varphi' - \varphi_{\hat{\gamma}}\|_1 &\leq (\overline{n} + 1)\varepsilon' + \widetilde{q}^{\overline{n}-\widetilde{n}} M \\ &\leq \varepsilon.\end{aligned}$$

Step 2. Now choose $n > \overline{n}(\delta', \varphi)$. There exist $k, r \in \mathbb{N}$ such that $n = k\overline{n} + r$. Call $\varphi'' = \mathcal{L}_{t_{k\overline{n}}} \cdots \mathcal{L}_{t_1}\varphi$. Using upper bound (9), we have

$$\begin{aligned}\|\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_1}\varphi - \varphi_{\hat{\gamma}}\|_1 &= \|\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_{n-r+1}}\varphi'' - \varphi_{\hat{\gamma}}\|_1 \\ &\leq \|\varphi_{\gamma_n} - \varphi_{\hat{\gamma}}\|_1 + \sum_{j=1}^{r} \|\varphi_{\gamma_{n-j}} - \varphi_{\gamma_{n-j+1}}\|_1 + |\varphi'' - \varphi_{\gamma_{n-r+1}}\|_1 \\ &\leq \varepsilon' + (r + 1)\varepsilon' + \varepsilon \\ &\leq 2\varepsilon.\end{aligned}$$

□

We now consider the proof of Part (3). The main tool we use, along with Theorem A, is a law of large numbers for dependent random variables with sufficiently fast decaying correlations.

THEOREM 3. [**Wal04**] *Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of square integrable random variables such that there exists $r : \mathbb{N}_0 \to \mathbb{R}$ with*

$$|\mathbb{E}[(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j)]| \leq r(|i - j|) \quad i, j \in \mathbb{N}$$

and

$$\sum_{k=1}^{\infty} \frac{r(k)}{k} < +\infty.$$

Then

$$\frac{1}{n}\sum_{k=1}^{n}(X_k - \mathbb{E}[X_k]) \to 0 \quad \textit{almost surely.}$$

Given an observable $\psi \in V_\alpha$, $n \in \mathbb{N}$ and $\boldsymbol{\gamma} \in \Gamma'^{\mathbb{N}}$ (where Γ' is as in Lemma 2), we define

$$\psi_0 := \psi \text{ and } \psi_n := U_{\gamma_1} \cdots U_{\gamma_n} \psi$$

and denote the sum of the observable over the orbit as

$$S_n(\psi)(x) = \sum_{i=0}^{n-1} \psi_i(x).$$

Here $(\psi_n)_{n \in \mathbb{N}_0}$ is a sequence of square integrable random variables on the measure space $(\mathbb{R}^N, m)$ and $\psi_n - \mathbb{E}[\psi_n] = U_{\gamma_1} \cdots U_{\gamma_n}(\widetilde{\psi}_n)$, where

$$\widetilde{\psi}_n := \psi - \int U_{\gamma_1} \circ \cdots \circ U_{\gamma_n} \psi \, dm.$$

We can estimate the covariance.

LEMMA 4. *For any $\psi \in V_\alpha$, there exist a constant $C = C(\psi, \varepsilon_0, N) > 0$ and $q < 1$ depending only on $F_{\hat{\gamma}}$ such that*

$$\mathbb{E}[(\psi_i - \mathbb{E}[\psi_i])(\psi_j - \mathbb{E}[\psi_j])] \leq C q^{j-i}$$

for every $\boldsymbol{\gamma} \in \Gamma'^{\mathbb{N}}$, $i, j \in \mathbb{N}$.

Proof. Without loss of generality, assume that $j > i$ and consider

$$\begin{aligned} R_{ij} &= \mathbb{E}[(\psi_i - \mathbb{E}[\psi_i])(\psi_j - \mathbb{E}[\psi_j])] \\ &= \mathbb{E}[U_{\gamma_1} \cdots U_{\gamma_i} \widetilde{\psi}_i \cdot U_{\gamma_1} \cdots U_{\gamma_j} \widetilde{\psi}_j]; \end{aligned}$$

then, using that properties of the Koopman and transfer operators, we obtain that

$$\begin{aligned} R_{ij} &= \mathbb{E}[\widetilde{\psi}_i \cdot U_{\gamma_{i+1}} \cdots U_{\gamma_j} \widetilde{\psi}_j \cdot \mathcal{L}_{\gamma_i} \cdots \mathcal{L}_{\gamma_1} 1] \\ &= \mathbb{E}[\widetilde{\psi}_j \cdot \mathcal{L}_{\gamma_j} \cdots \mathcal{L}_{\gamma_{i+1}} (\widetilde{\psi}_i \cdot \mathcal{L}_{\gamma_i} \cdots \mathcal{L}_{\gamma_1} 1)]. \end{aligned}$$

Since

$$\int \widetilde{\psi}_i \cdot \mathcal{L}_{\gamma_i} \cdots \mathcal{L}_{\gamma_1} 1 \, dm = \int U_{\gamma_1} \cdots U_{\gamma_i} \bigg(\psi - \int U_{\gamma_1} \cdots U_{\gamma_i} \psi \, dm \bigg) \, dm = 0$$

implies that $\widetilde{\psi}_i \cdot \mathcal{L}_{\gamma_i} \cdots \mathcal{L}_{\gamma_1} 1 \in X_0$, and since $\mathcal{L}_\gamma$ restricted to X_0 are contractions with respect to the $\| \cdot \|_\alpha$ norm, we obtain the bound

$$\|\mathcal{L}_{\gamma_j} \cdots \mathcal{L}_{\gamma_{i+1}} (\widetilde{\psi}_i \cdot \mathcal{L}_{\gamma_i} \cdots \mathcal{L}_{\gamma_1} 1)\|_1 \leq \widetilde{q}^{j-i} \|\widetilde{\psi}_i \cdot \mathcal{L}_{\gamma_i} \cdots \mathcal{L}_{\gamma_1} 1\|_\alpha.$$

By Proposition 3.4 in [**Sau00**], V_α is an algebra and

$$\begin{aligned} \|\widetilde{\psi}_i \cdot \mathcal{L}_{\gamma_i} \cdots \mathcal{L}_{\gamma_1} 1\|_\alpha &\leq C_\# \|\widetilde{\psi}_i\|_\alpha \|\mathcal{L}_{\gamma_i} \cdots \mathcal{L}_{\gamma_1} 1\|_\alpha \qquad (11) \\ &\leq C_\# (|\psi|_\alpha + \|\widetilde{\psi}_i\|_1) + \widetilde{\eta} + \frac{c}{1 - \widetilde{\eta}} \qquad (12) \\ &\leq C_\# (1 + \|\psi\|_\alpha), \end{aligned}$$

where $C_\#$ stands for an uninfluential constant uniform on ψ, i and $\boldsymbol{\gamma}$. Inequality (11) is a consequence of the upper bound in Proposition 3.4 from [**Sau00**]; (12) is derived from the first point in Lemma 2. We conclude noticing that, by the Hölder inequality,

$$\begin{aligned}\mathbb{E}[\widetilde{\psi}_j \cdot \mathcal{L}_{\gamma_j} \cdots \mathcal{L}_{\gamma_{i+1}}(\widetilde{\psi}_i \cdot \mathcal{L}_{\gamma_i} \cdots \mathcal{L}_{\gamma_1} 1)] &\le \|\widetilde{\psi}_j\|_\infty \widetilde{q}^{j-i} C_\#(1 + \|\psi\|_\alpha)\\ &\le \|\psi\|_\infty \widetilde{q}^{j-i} C_\#(1 + \|\psi\|_\alpha).\end{aligned}$$

By Proposition 3.4 in [**Sau00**], any $\psi \in V_\alpha$ is an essentially bounded function with

$$\|\psi\|_\infty \le \frac{\max\{1, \varepsilon^\alpha\}}{\gamma_N \varepsilon_0^N} \|\psi\|_\alpha,$$

where γ_N is the volume of the N-dimensional unit ball. Hence, taking

$$C := C_\# \frac{\max\{1, \varepsilon^\alpha\}}{\gamma_N \varepsilon_0^N} \|\psi\|_\alpha (1 + \|\psi\|_\alpha),$$

we conclude the proof. □

Proof of Part (3) of Theorem A. We know that

$$\begin{aligned}\mathbb{E}[\psi_k] &= \int \psi_k \, dm\\ &= \int U_{\gamma_1} \cdots U_{\gamma_k}(\psi) \, dm\\ &= \int \psi \mathcal{L}_{\gamma_k} \cdots \mathcal{L}_{\gamma_1} 1 \, dm\end{aligned}$$

and, using Theorem A, for every $k > \overline{n}$ and sufficiently small perturbations,

$$\|\mathcal{L}_{\gamma_k} \cdots \mathcal{L}_{\gamma_1} 1 - \varphi_0\|_1 < \varepsilon \quad \Rightarrow \quad \left|\mathbb{E}[\psi_k] - \int \psi \varphi_0 \, dm\right| \le \varepsilon \mathbb{E}[|\psi|]$$

and

$$\begin{aligned}\int \psi \varphi_0 \, dm - \varepsilon \mathbb{E}[|\psi|] &\le \liminf_{n\to\infty} \frac{1}{n} \sum_k \mathbb{E}[\psi_k] \le \limsup_{n\to\infty} \frac{1}{n} \sum_k \mathbb{E}[\psi_k]\\ &\le \int \psi \varphi_0 \, dm + \varepsilon \mathbb{E}[|\psi|].\end{aligned}$$

Thanks to Lemma 4, we can directly apply Theorem 3 to the random variables $X_k = \psi_k$. This implies that, for almost every $x \in M$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \psi_k(x) - \mathbb{E}[\psi_k] = 0, \tag{13}$$

which implies that

$$\begin{aligned}\int \psi \varphi_0 \, dm - \varepsilon \mathbb{E}[|\psi|] &\le \liminf_{n\to\infty} \frac{1}{n} \sum_k \psi_k(x) \le \limsup_{n\to\infty} \frac{1}{n} \sum_k \psi_k(x)\\ &\le \int \psi \varphi_0 \, dm + \varepsilon \mathbb{E}[|\psi|].\end{aligned}$$

□

Proof of Corollary 1. To prove the corollary, we use the Levy–Prokhorov metric for the weak topology. Given two probability measures μ, ν on the Borel σ-algebra $(\Omega, \mathcal{B}(\Omega))$, the Levy–Prokhorov metric is defined as

$$d_{LP}(\mu, \nu) := \inf\{s | \mu(A) \leq \nu(A_s) + s \text{ for all } A \in \mathcal{B}\},$$

where A_s is the set of points at Euclidean distance strictly less than s from A. Since the underlying topological space Ω is separable, it induces the weak topology on the space of Borel probability measures. To prove the corollary it is then enough to show that for fixed $\varepsilon > 0$ there is a $\delta > 0$ such that for every sequence $\boldsymbol{\gamma} \in B_\delta(\hat{\gamma})^{\mathbb{N}}$ and for almost every $x \in \Omega$ there is $\overline{n}$ such that

$$\mu_{n,x} := \frac{1}{n} \sum_{i=0}^{n-1} (F^i_{\boldsymbol{\gamma}})_* \delta_x \in B_\varepsilon(\mu_{\hat{\gamma}}) \quad \text{for all } n > \overline{n},$$

where $B_\varepsilon(\mu_{\hat{\gamma}})$ is the ε-ball around $\mu_{\hat{\gamma}}$ with respect to the metric d_{LP}. We have

$$\mu_{n,x}(A) = \frac{1}{n} \sum_{i=0}^{n-1} \chi_A \circ F^i_{\boldsymbol{\gamma}}(x)$$

and, from equation (13) in the proof of Part (3) of Theorem A with $\psi = \chi_A$,

$$\mu_{n,x}(A) - \frac{1}{n} \sum_i \mathbb{E}[\chi_A \circ F^i_{\boldsymbol{\gamma}}(x)] \to 0$$

for almost every $x \in \Omega$. Using the properties of the transfer operator,

$$\mu_{n,x}(A) - \frac{1}{n} \sum_{i=0}^{n-1} \int_A \mathcal{L}^i_{\boldsymbol{\gamma}}(1)\, dm \to 0 \quad \text{almost everywhere.} \tag{14}$$

L^1 convergence of densities implies weak convergence of the corresponding probability measures; thus,

$$\|\mathcal{L}^i_{\boldsymbol{\gamma}}(1) - \varphi_{\hat{\gamma}}\|_1 \leq \varepsilon' \Rightarrow d_{LP}((F^i_{\boldsymbol{\gamma}})_* m, \mu_{\hat{\gamma}}) \leq \Delta(\varepsilon') \tag{15}$$

with $\Delta(\varepsilon') \to 0$ for $\varepsilon' \to 0$, which implies that

$$\int_A \mathcal{L}^i_{\boldsymbol{\gamma}}(1)\, dm = (F^i_{\boldsymbol{\gamma}})_* m(A) \leq \mu_{\hat{\gamma}}(A_{\Delta(\varepsilon')}) + \Delta(\varepsilon') \quad \text{for all } A \in \mathcal{B}. \tag{16}$$

Now the main problem is that convergence (14) is not uniform in A. To overcome this issue, we use the standard technique to approximate any measurable set A as the union of sets taken from a sufficiently fine (in terms of the Euclidean metric), but finite, collection and use uniformity of (13) for this finite collection to deduce uniformity for any A. Fix $\delta' > 0$. From the topological properties of Ω, one can find a finite collection of balls $\{B_i\}_{i=1}^J$ such that $\operatorname{diam}(B_i) < \delta'$, $m(\bigcup_{i=1}^J B_i) > 1 - \delta'$. The set

$$\mathcal{D} := \left\{ \bigcup_{i \in \mathcal{J}} B_i | \mathcal{J} \subset \{1, \ldots, J\} \right\},$$

given by all possible unions of sets from the collection, is itself finite. This means that we can find a subset $\Omega_{\delta'} \subset \Omega$ of full measure such that for every $x \in \Omega_{\delta'}$ there is $\overline{n} = \overline{n}(x)$ such that

$$\left|\mu_{n,x}(B) - \frac{1}{n}\sum_{i=0}^{n-1}\int_B P^i_{\boldsymbol{\gamma}}(1)\,dm\right| < \delta' \quad \text{for all } B \in \mathcal{D} \text{ and for all } n > \overline{n}'. \tag{17}$$

From Theorem A, we can choose $\delta > 0$ so small that (15) holds for every $i > N_1$ and $\Delta(\varepsilon') < \delta'$. Thus, there is an $N_2 > N_1$ such that

$$\left|\frac{1}{n}\sum_{i=0}^{n-1}\int_B P^i_{\boldsymbol{\gamma}}(1) - \mu_{\hat{\gamma}}(B_{\Delta(\varepsilon')})\right| < 2\Delta(\varepsilon') < 2\delta' \quad \text{for all } n > N_2. \tag{18}$$

Putting (18) and (17) together,

$$\mu_{n,x}(B) \le \mu_{\hat{\gamma}}(B_{\delta'}) + 3\delta'.$$

Now consider $\mathcal{J}_A := \{i | B_i \cap A \neq 0\}$. We have $A \subset (\bigcup_{i\in\mathcal{J}_A} B_i) \cup (\bigcup_{i=1}^{J} B_i)^c$ (trivially) and $(\bigcup_{i\in\mathcal{J}_A}(B_i)_{\delta'}) \subset A_{2\delta'}$ (from the condition on the diameters of the sets of the partition). Thus,

$$\begin{aligned}\mu_{x,n}(A) &\le \mu_{x,n}\left(\bigcup_{i\in\mathcal{J}_A} B_i\right) + \delta' \\ &\le \mu_{\hat{\gamma}}\left(\bigcup_{i\in\mathcal{J}_A}(B_i)_{\delta'}\right) + 4\delta' \\ &\le \mu_{\hat{\gamma}}(A_{2\delta'}) + 4\delta',\end{aligned}$$

which for the right choice of δ' gives the desired result.

Acknowledgements. The authors thank G. Keller, C. Liverani, V. Baladi and D. Turaev for fruitful conversations related to this topic. This work was partially supported by the European Union's Seventh Framework Programme, ERC AdG Grant No. 339523 RGDD (SvS) and FAPESP project 15/08958-4 (TP). We also acknowledge partial support by the EU Marie Curie IRSES Brazilian–European partnership in dynamical systems (FP7-PEOPLE-2012-IRSES 318999 BREUDS).

A. *Appendix. $C^{1+\nu}$-expanding maps*

In this appendix we consider a collection $\{F_\gamma\}_{\gamma\in\Gamma}$ of $C^{1+\nu}$ maps from M into itself (where Γ is some metric space) which are continuous at $\hat{\gamma}$, i.e. so that the map $\gamma \mapsto F_\gamma$ from Γ to $C^{1+\nu}(M, M)$ is *continuous at* $\hat{\gamma}$. We will assume that $F = F_{\hat{\gamma}}$ is expanding and for each sequence $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots) \in (B_\delta(\hat{\gamma}))^{\mathbb{N}}$ we analyze the ergodic properties of compositions of the form

$$F^n_{\boldsymbol{\gamma}} = F_{\gamma_n} \cdots F_{\gamma_1}.$$

We are particularly interested in studying what happens when we let δ tend to zero.

A.1. *Setting and result.* Let M be connected and compact Riemannian manifold, with Riemannian distance d and Riemannian volume m (we normalize so that $m(M) = 1$).

Definition A.1. We say that $F : M \to M$ is a $C^{1+\nu}$*-expanding map* if F is differentiable, $\log |\det D_x F|$ is a locally ν-Hölder function and there exists a constant $\sigma \in (0, 1)$ such that

$$\|D_x F(v)\| \geq \sigma^{-1}\|v\| \quad \text{for all } x \in M \text{ for all } v \in T_x M.$$

It is well known that F has an *absolutely continuous invariant measure* μ_0 with a density φ_0 which is strictly positive and ν-Hölder.

The next theorem shows that for expanding maps all ergodic properties are robust.

THEOREM B. *Let F be a $C^{1+\nu}$-expanding map and $\{F_\gamma\}_{\gamma\in\Gamma}$ a collection of $C^{1+\nu}$ perturbed versions continuous at $\hat{\gamma}$. Then, for every $\varepsilon > 0$, there exists $\delta > 0$ so that:*

(1) **[Via97]** *if $\gamma \in B_\delta(\hat{\gamma})$, then F_γ has an invariant density φ_γ and*

$$\sup_{x\in M} |\varphi_\gamma(x) - \varphi_{\hat{\gamma}}| < \varepsilon;$$

(2) *for every $\boldsymbol{\gamma} \in B_\delta(\hat{\gamma})^{\mathbb{N}}$ and probability measure $\mu = \varphi m$ with a strictly positive ν-Hölder density φ, there exists $\overline{n}$ so that, for $n > \overline{n}$, the Radon–Nikodym derivative $d(F^n_{\boldsymbol{\gamma}})_*\mu/dm$ is strictly positive and ν-Hölder and*

$$\sup_{x\in M} \left| \left(\frac{d}{dm}(F^n_{\boldsymbol{\gamma}})_*\mu \right)(x) - \varphi_{\hat{\gamma}}(x) \right| < \varepsilon; \tag{19}$$

(3) *there exists a set $X_{\boldsymbol{\gamma}} \subset M$ of full measure such that, for each ν-Hölder observable ψ and for each $x \in X_{\boldsymbol{\gamma}}$,*

$$\int \psi\varphi_0 \, dm - \varepsilon \int |\psi| \, dm \leq \liminf_{n\to\infty} \frac{1}{n} S_n(\psi)(x) \leq \limsup_{n\to\infty} \frac{1}{n} S_n(\psi)(x) \leq \int \psi\varphi_0 \, dm + \varepsilon \int |\psi| \, dm,$$

where $S_n(\psi)(x) = \psi(x) + \sum_{i=1}^{n-1} \psi(F_{\gamma_i} \cdots F_{\gamma_1}(x))$.

Remark A.1. Recalling that the unique invariant density for a $C^{1+\nu}$-expanding map on a compact manifold is uniformly bounded and bounded away from zero, the convergence in the sup-norm stated above implies that after a finite number of iterates, any regular density will remain bounded and bounded away from zero under any sufficiently small non-autonomous composition of perturbations of such a map.

Remark A.2. Keller [**Kel82b**] obtained a result for the L^1-analogue of inequality (19) for piecewise expanding interval maps. In our proof of Theorem B we use the well-known cone approach (explained below). In our setting, since we assume that the maps are $C^{1+\nu}$, the cone approach is more direct and gives the stronger uniform (rather than L^1) estimates.

A.2. *Strategy of the proof of Theorem B.* The main ingredient in the proof of Theorem B is the standard cone approach [**Liv95a**, **Liv01**, **Via97**]. For sufficiently small perturbations of a given expanding map, the associated transfer operators leave a cone of strictly positive continuous functions with Hölder logarithm invariant. Endowing the cone with the Hilbert metric, one is able to prove that it has finite diameter and thus the restrictions of the transfer operators to the cone are contractions with respect to this metric and their contracting rates are uniformly bounded away from 1. This immediately implies memory loss for initial distributions of states belonging to the cone. The main claim is implied by a combination of this last result with continuity of the map $t \mapsto \mathcal{L}_t\varphi$ for every fixed function φ inside the cone. We preliminarily define the transfer operator, cones of functions and the Hilbert metric in §A.3 to state results of existence of a.c.i.p. measures in §A.3.3 and stochastic stability in §4.1.1. We conclude with the proof of the result in the final section.

A.3. *Preliminaries.* In this section we review some of the main concepts and techniques to study invariant measure and statistical properties of regular expanding maps that will lead to the proof of Theorem B given in §A.5.

A.3.1. *Transfer operator.* Given a measurable space $(M, \mathcal{B}, m)$ with a non-singular transformation F, one can define two operators on the spaces $L^\infty(M)$ and $L^1(M)$ (the specification of the measure m is omitted whenever there is no risk of confusion), respectively $U : L^\infty(M) \to L^\infty(M)$,

$$U(\varphi) = \varphi \circ F,$$

and its adjoint $\mathcal{L} : L^1(M) \to L^1(M)$, that satisfy for all $\varphi \in L^\infty(M)$ and $\psi \in L^1(M)$,

$$\int_M U\varphi \cdot \psi \, dm = \int_M \varphi \cdot \mathcal{L}\psi \, dm. \tag{20}$$

Here $\mathcal{L}$ is the transfer (or Perron–Frobenius) operator, and has the following properties:

positivity:	$\psi \geq 0 \Rightarrow \mathcal{L}_F\psi \geq 0,$	(P1)
preserves integrals:	$\int \mathcal{L}_F\psi \, dm = \int \psi \, dm,$	(P2)
contraction property:	$\lvert\mathcal{L}_F\psi\rvert_1 \leq \lvert\psi\rvert_1,$	(P3)
composition property:	$\mathcal{L}_{F\circ G} = \mathcal{L}_G \circ \mathcal{L}_F$	(P4)

for any $\psi \in L^1(M)$.

The transfer operator has proven to be an invaluable tool to deduce statistical properties of dynamical systems (such as existence of invariant measures, decay of correlations and central limit theorems for Birkhoff sums [**Liv95a**, **Bal00**, **BG97**, **LM13**, **Via97**]) and their perturbations [**Kel82a**, **Kel82b**, **BY93**, **Via97**]. For example, the probability densities fixed by the transfer operator are the densities of the invariant absolutely continuous probability measures and, likewise, for any $\mu = \varphi m$ a.c.i.p. measure, φ is a fixed point for the transfer operator. The transfer operator also prescribes the evolution of the densities.

For maps satisfying Definition A.1, each point of M has the same finite number of preimages under F (M is compact and connected), and the transfer operator for these maps can be written as

$$\mathcal{L}\varphi(x) = \sum_{i=1}^{k} \varphi(y_i) \cdot |\det D_{y_i} F|^{-1}, \quad F^{-1}(x) = \{y_1, \ldots, y_k\}.$$

A.3.2. *Projective metric.* One of the main tools to investigate properties of the transfer operator is the Hilbert projective metric [**Liv95a**, **Liv01**, **Via97**].

Given a linear vector space E, a cone of E is a subset $C \subset E\backslash\{0\}$ such that if $v \in C$, then $sv \in C$ for all $s > 0$. We consider convex cones, namely those that satisfy

$$s_1 v_1 + s_2 v_2 \in C \quad \text{for all } s_1, s_2 > 0 \text{ and for all } v_1, v_2 \in C.$$

There is a canonical way to define a pseudometric on every cone. Let

$$\alpha(v_1, v_2) := \sup\{s > 0 : \ v_2 - s v_1 \in C\},$$

$$\beta(v_1, v_2) := \inf\{s > 0 : \ s v_1 - v_2 \in C\}$$

and

$$\theta(v_1, v_2) := \log \frac{\beta(v_1, v_2)}{\alpha(v_1, v_2)}$$

with the logarithm extended to a function of $[0, +\infty]$. Here θ has the following properties (see [**Via97**] for a proof):

(i) $\theta(v_1, v_2) = \theta(v_2, v_1)$ for all $v_1, v_2 \in C$;

(ii) $\theta(v_1, v_3) \leq \theta(v_1, v_2) + \theta(v_2, v_3)$ for all $v_1, v_2, v_3 \in C$;

(iii) $\theta(v_1, v_2) = 0$ if and only if there is $s > 0$ such that $v_1 = s v_2$.

Point (iii) implies that θ distinguishes directions only, and for this reason it is called the projective metric.

Remark A.3. Looking closely at α and β, one can notice that it is necessary to evaluate one of the two only. In particular,

$$\alpha(v_1, v_2) = \frac{1}{\beta(v_2, v_1)} \quad \text{for all } v_1, v_2 \in C, \tag{21}$$

which also implies property (i) of θ.

Remark A.4. The projective metric $\theta := \theta_C$ depends on the cone on which it is defined. Vectors belonging to the intersection of two different cones of the same vector space might have different projective distances whether considered as vectors of the first cone or of the second.

Given cones $C_1 \subset C_2$, with projective metrics θ_1 and θ_2, one has

$$\theta_2(v_1, v_2) \leq \theta_1(v_1, v_2) \quad \text{for all } v_1, v_2 \in C_1, \tag{22}$$

which tells us that the projective distance between two vectors decreases after enlarging the cone. We now present two examples that will be useful in what follows.

Example A.1. The cone of strictly positive continuous functions is

$$C_+ := \{\varphi \in C^0(M, \mathbb{R}) \text{ s.t. } \varphi > 0\}.$$

We compute the projective metric:

$$\begin{aligned}\alpha_+(\varphi_1, \varphi_2) &= \sup\{t > 0 : \ \varphi_2(x) - t\varphi_1(x) > 0 \text{ for all } x \in M\} \\ &= \sup\{t > 0 : t < \varphi_2(x)/\varphi_1(x) \text{ for all } x \in M\},\end{aligned}$$

which gives $\alpha_+(\varphi_1, \varphi_2) = \inf_{x\in M} \varphi_2(x)/\varphi_1(x)$. From (21), $\beta_+(\varphi_1, \varphi_2) = \sup_{x\in M} \varphi_2(x)/\varphi_1(x)$ and

$$\theta_+(\varphi_1, \varphi_2) = \log \sup \left\{ \frac{\varphi_2(x)\varphi_1(y)}{\varphi_1(x)\varphi_2(y)} \text{ s.t. } x, y \in M \right\}.$$

Example A.2. For every $a > 0$ and $\nu \in (0, \nu)$, consider the cone

$$C(a, \nu) := \{\varphi \in C_+ \text{ s.t. } d(x_1, x_2) \le \rho_0 \Rightarrow \varphi(x_1) \le \exp(ad(x_1, x_2)^\nu)\varphi(x_2)\}, \tag{23}$$

which is the cone of strictly positive continuous functions with $\log \varphi$ locally $\nu-$Hölder. The computation of the projective metric $\theta_{a,\nu}$ on this cone can be obtained in a similar way as before (see **[Via97]** for details). In particular, we have

$$\alpha(\varphi_1, \varphi_2) = \inf \left\{ \frac{\varphi_2(x)}{\varphi_1(x)}, \frac{\exp(ad(x, y)^\nu)\varphi_2(x) - \varphi_2(y)}{\exp(ad(x, y)^\nu)\varphi_1(x) - \varphi_1(y)} \right. \left. \text{ s.t. } x, y \in M \text{ and } 0 < d(x, y) < \rho_0 \right\}.$$

One usually considers cones instead of the whole linear space because the restriction of linear operators to invariant cones exhibits nice properties. For example, letting E_1, E_2 be two vector spaces, $L : E_1 \to E_2$ a linear map and C_1, C_2 two cones in E_1 and E_2 respectively such that $L(C_1) \subset C_2$, it is easy to verify that

$$\theta_2(L(v_1), L(v_2)) \le \theta_1(v_1, v_2) \quad \text{for all } v_1, v_2 \in C_1.$$

Furthermore, if the image of C_1 has finite diameter, then the restriction of the linear map to the cone is a contraction.

PROPOSITION A.1. **[Via97]** *Let $L : E_1 \to E_2$ be a linear map and $C_1 \subset E_1$ a cone. If*

$$D := \sup\{\theta_2(L(v_1), L(v_2)) \textit{ s.t. } v_1, v_2 \in C_1\}$$

is finite, then

$$\theta_2(L(v_1), L(v_1)) \le q\theta_1(v_1, v_2)$$

with $q = 1 - e^{-D}$.

A.3.3. *Absolutely continuous invariant probability measure.* The above machinery has been used to prove existence of an invariant absolutely continuous probability measure for the class of maps introduced in Definition A.1.

PROPOSITION A.2. [**Via97**] *Let F be a map as in Definition A.1 and let $\mathcal{L}$ be the associated transfer operator. Then, for all sufficiently large $a > 0$, and for all $\lambda \in (\sigma, 1)$,*

$$\mathcal{L}(C(a, \nu)) \subset C(\lambda a, \nu).$$

Moreover, the diameter

$$D_{\lambda a,\nu} := \sup\{\theta_{a,\nu}(\varphi_1, \varphi_2) \text{ s.t. } \varphi_1, \varphi_2 \in C(\lambda a, \nu)\} \tag{24}$$

is finite for every $a > 0$, $\nu > 0$, $0 < \lambda < 1$.

This proposition along with Proposition A.1 imply that the action of $\mathcal{L}$ contracts directions inside the cone $C(a, \nu)$. To get a step closer to obtaining a fixed point, we can restrict to the subset of normalized densities in $C(a, \nu)$,

$$\widetilde{C}(a, \nu) := \left\{\varphi \in C(a, \nu) \text{ s.t. } \int \varphi(x)\, dm(x) = 1\right\},$$

and show that $\mathcal{L}$ is a contraction of this space with respect to the restriction of the projective metric.

PROPOSITION A.3. *The following hold:*
(i) *the restriction of $\theta_{a,\nu}$ to $\widetilde{C}(a, \nu)$ is a metric;*
(ii) $\mathcal{L}(\widetilde{C}(a, \nu)) \subset \widetilde{C}(\lambda a, \nu)$.

Proof. (i) $C(a, \nu)$ has finite diameter and thus $\theta_{a,\nu}$ takes only finite values. $\theta_{a,\nu}(\varphi_1, \varphi_2) = 0$ for $\varphi_1, \varphi_2 \in \widetilde{C}(a, \nu)$ implies that $\varphi_1 = c\varphi_2$, but c must be equal to 1, so $\varphi_1 = \varphi_2$. (ii) is implied by properties (P1) and (P2) of the transfer operator. □

This immediately gives the following corollary.

COROLLARY A.1. *$\mathcal{L}$ is a contraction on the metric space $(\widetilde{C}(a, \nu), \theta_{a,\nu})$.*

Without getting into much detail (which can be found in [**Via97**]), to produce a fixed point it would be sufficient to show that every normalized Cauchy sequence in $C(a, \nu)$ is convergent. The setback is that $(\widetilde{C}(a, \nu), \theta_{a,\nu})$ is not complete. However, as has already been pointed out, $C(a, \nu)$ is a subset of C_+. Normalized Cauchy sequences converge in this cone with respect to θ_+ and the fact that $\theta_+(\varphi_1, \varphi_2) \le \theta_{a,\nu}(\varphi_1, \varphi_2)$ for all $\varphi_1, \varphi_2 \in C(a, \nu)$ (see equation (22)) implies the existence of a fixed point $\varphi_0 \in C_+$ which is an invariant density for the dynamical system $(M, \mathcal{B}, m, F)$. One is also able to prove that φ_0 is in fact a function in $C(\lambda a, \nu)$ and that there exist constants $R > 0$ and $\sigma_1 \in (0, 1)$ such that

$$\sup_{x \in M} |\mathcal{L}^n \varphi(x) - \varphi_0(x)| < R\sigma_1^n \tag{25}$$

for all densities $\varphi \in \widetilde{C}(a, \nu)$. This means that any distribution of mass $\mu = \varphi m$ on M, for $\varphi \in \widetilde{C}(a, \nu)$, will evolve exponentially fast towards the invariant distribution $\mu_0 = \varphi_0 m$ under the iteration of the map.

Remark A.5. Notice that taking a collection of $C^{1+\nu}$ perturbations $\{F_\gamma\}_\Gamma$, if $\delta > 0$ is sufficiently small, then every F_γ for γ in the open ball $B_\delta(\hat{\gamma})$ is a $C^{1+\nu}$-expanding map with uniform lower bound $\overline{\sigma}^{-1}$ on the rate of expansion and on the Hölder constant. This implies that all the associated transfer operators map $C(a, \nu)$ into $C(\overline{\lambda} a, \nu)$ for some sufficiently large a and $\overline{\lambda} \in (0, 1)$. This implies that, for all $\gamma \in B_\delta(\hat{\gamma})$, $\mathcal{L}_\gamma$ has a fixed point $\varphi_\gamma \in C(\overline{\lambda} a, \nu)$ and therefore there is an invariant absolutely continuous probability measure for the perturbed map. The analogue of (25) holds:

$$\sup_{x \in M} |\mathcal{L}_\gamma^n \varphi(x) - \varphi_\gamma(x)| < R\sigma_2^n \tag{26}$$

for some $R > 0$, $\sigma_2 \in (0, 1)$ and for all $\gamma \in B_\delta(\hat{\gamma})$.

A.4. *Stochastic stability.* As for the piecewise case (§4.1.1), also for families of $C^{1+\nu}$ maps one can consider independent compositions of maps sampled according to some measure ν on Γ. In this case, it is known [**Via97**, **BY93**] that if the support of ν is sufficiently close to τ, i.e. $\operatorname{supp} \mu \subset B_\delta(\tau)$ for sufficiently small $\delta > 0$, then the averaged transfer operator has an invariant density φ_μ, and this invariant density converges uniformly to the invariant density φ_τ whenever $\delta \to 0$. φ_μ is called a stationary density and it describes the asymptotic distribution for the random orbits for almost every initial condition (with respect to $\varphi_\mu m$) and for almost every sequence $\{t_i\}_{i \in \mathbb{N}}$ (with respect to $\mu^{\otimes \mathbb{N}}$). The stochastic stability result asserts that stationary densities are uniformly close to the unperturbed invariant density whenever the support of their measure is sufficiently close to τ.

PROPOSITION A.4. [**Via97**] *Given a probability measure μ on T, for every $\varepsilon > 0$ there is $\delta > 0$ such that if* $\operatorname{supp} \mu \subset B_\delta(\tau)$*, then*

$$\sup_{x \in M} |\varphi_\mu(x) - \varphi_\tau(x)| < \varepsilon.$$

Remark A.6. Notice that a particular case of the above setting is when $\mu = \delta_t$ is the singular probability measure concentrated at the point $t \in T$. In this case, $\hat{\mathcal{L}}_\mu$ equals $\mathcal{L}_t$, and the above results imply that, for t sufficiently close to τ, $F^{(t)}$ has an absolutely continuous invariant probability measure with density φ_t, and $\varphi_t \to \varphi_\tau$ uniformly for $t \to \tau$.

The bound (26) implies that the evolution of densities under the iterated action of a perturbed map F_γ converges exponentially fast to the invariant density φ_γ. Proposition A.4 with the above remark imply that for small perturbations, under iteration of some map F_γ, densities in $\widetilde{C}(a, \nu)$ evolve close to $\varphi_{\hat{\gamma}}$.

A.5. *Proof of Theorem B.* Part (1) of Theorem B is given by Proposition A.4. Now we prove Part (2) that, in contrast with the stationary case, tells us what happens to densities when we apply perturbed versions of the map F without requiring any kind of independence of the perturbations.

Proof of Part (2) of Theorem B. Thanks to property (20), we can restate the theorem in terms of the action of the transfer operator on the densities. Thus, we need to prove that for $\{F_\gamma\}_{\gamma \in \Gamma}$ as in the hypotheses, denoting by

$$\mathcal{L}_{\boldsymbol{\gamma}}^n := \mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_1}$$

the transfer operator of the composition $F^n_{\boldsymbol{\gamma}} := F_{\gamma_n} \circ \cdots \circ F_{\gamma_1}$, for every $\varepsilon > 0$ there exist $\delta(\varepsilon) > 0$ and $\overline{n}(\varepsilon) \in \mathbb{N}$ such that for every $\varphi \in \widetilde{C}(a, \nu)$, every $n > \overline{n}$ and $\boldsymbol{\gamma} \in B_\delta(\hat{\gamma})^n$,

$$\mathcal{L}^n_{\boldsymbol{\gamma}}\varphi \in \widetilde{C}(a, \nu) \quad \text{and} \quad \sup_{x \in M} |\mathcal{L}^n_{\boldsymbol{\gamma}}\varphi(x) - \varphi_0(x)| < \varepsilon.$$

$\mathcal{L}^n_{\boldsymbol{\gamma}}\varphi \in \widetilde{C}(a, \nu)$ follows from Remark A.5.

Then we recall that, for every $\varepsilon' > 0$, $\varphi \in C(a, \nu)$ and every $n \in \mathbb{N}$, there exists $\delta(\varepsilon', \varphi, n) > 0$ such that if $\boldsymbol{\gamma} \in B_\delta(\tau)^n$, then

$$\theta_+(\mathcal{L}^n_{\boldsymbol{\gamma}}\varphi, \mathcal{L}^n\varphi) \leq \varepsilon'.$$

This is proven in Proposition 2.14 of [**Via97**] for the case $n = 1$, and can be generalized to any finite $n \in \mathbb{N}$.

If $\gamma \in B_\delta(\hat{\gamma})$ with $\delta > 0$ sufficiently small, for some $a > 0$ and $0 < \nu < 1$, $\mathcal{L}_\gamma$ is a contraction of $(\widetilde{C}(a, \nu), \theta_{a,\nu})$ with contracting constant $q \in (0, 1)$ independent of γ. Fix $\varepsilon' > 0$ and choose $\overline{n} \in \mathbb{N}$ so that $q^{\overline{n}} D_{\lambda a,\nu} < \varepsilon'/2$, with $D_{\lambda a,\nu}$ the diameter of $C(a, \nu)$ (see (24)), and δ less than $\delta(\varepsilon'/2, \varphi_0, \overline{n})$ above, so that $\theta_+(\mathcal{L}^{\overline{n}}_{\boldsymbol{\gamma}}\varphi_0, \varphi_0) < \varepsilon'/2$ for all $\boldsymbol{\gamma} \in B_\delta(\hat{\gamma})^{\overline{n}}$. This implies that, for any $n > \overline{n}$ and any $\boldsymbol{\gamma} \in B_\delta(\hat{\gamma})^{\overline{n}}$,

$$\begin{aligned}
\theta_+(\mathcal{L}^n_{\boldsymbol{\gamma}}\varphi, \varphi_0) &\leq \theta_+(\mathcal{L}^n_{\boldsymbol{\gamma}}\varphi, \mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_{n-\bar{n}}}\varphi_0) + \theta_+(\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_{n-\bar{n}}}\varphi_0, \varphi_0) &(27)\\
&\leq \theta_{a,\nu}(\mathcal{L}^n_{\boldsymbol{\gamma}}\varphi, \mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_{n-\bar{n}}}\varphi_0) + \theta_+(\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_{n-\bar{n}}}\varphi_0, \varphi_0) &(28)\\
&\leq \theta_{a,\nu}(\mathcal{L}_{\gamma_n} \cdots \mathcal{L}_{\gamma_{n-\bar{n}}}\mathcal{L}_{\gamma_{n-\bar{n}-1}} \cdots \mathcal{L}_{\gamma_1}\varphi, \mathcal{L}_{(\gamma_n,\ldots,\gamma_{n-\bar{n}})}\varphi_0) + \varepsilon'/2 &(29)\\
&\leq q^{\overline{n}} D_{\lambda a,\nu} + \varepsilon'/2 &(30)\\
&\leq \varepsilon'.
\end{aligned}$$

While (27) is just the triangle inequality, (28) follows from inequality (22) that upper bounds the metric θ_+ with the metric $\theta_{a,\nu}$ and (30) follows from uniformity of the contraction rate, $q \in (0, 1)$, for $\{\mathcal{L}_\gamma\}_{\gamma \in B_\delta(\hat{\gamma})}$. □

To prove Part (3), we use again the strong law of large numbers in Theorem 3. Given an observable $\psi \in C(a, \nu)$, $n \in \mathbb{N}$ and $\boldsymbol{\gamma} \in \Gamma^{\mathbb{N}}$, we define

$$\psi_0 := \psi, \quad \psi_n := U_{\gamma_1} \cdots U_{\gamma_n}\psi \quad \text{and} \quad S_n(\psi)(x) := \sum_{i=0}^{n-1} \psi_i(x).$$

$(\psi_n)_{n \in \mathbb{N}_0}$ is a sequence of square integrable random variables on the measure space (M, m), and $\psi_n - \mathbb{E}[\psi_n] = U_{\gamma_1} \cdots U\gamma_n(\widetilde{\psi}_n)$, where

$$\begin{aligned}
\widetilde{\psi}_n(x) &:= \psi - \int U_{\gamma_1} \cdots U_{\gamma_n}\psi \, dm\\
&= \psi - \int \psi \mathcal{L}^n_{\boldsymbol{\gamma}}(1) \, dm.
\end{aligned}$$

We estimate the covariance in the following lemma.

LEMMA A.1. *There exist constants $R > 0$ and $r \in (0, 1)$ such that, for every $i \neq j$,*

$$\mathbb{E}[(\psi_i - \mathbb{E}[\psi_i])(\psi_j - \mathbb{E}[\psi_j])] \leq R r^{|j-i+1|}.$$

Proof. Suppose that $j > i$ without loss of generality. Then

$$\begin{aligned}
&\mathbb{E}[(\psi_i - \mathbb{E}[\psi_i])(\psi_j - \mathbb{E}[\psi_j])] \\
&\quad = \mathbb{E}[U_{\gamma_1} \cdots U_{\gamma_i} \widetilde{\psi}_i \cdot U_{\gamma_1} \cdots U_{\gamma_j} \widetilde{\psi}_j] \\
&\quad = \mathbb{E}[\widetilde{\psi}_i \cdot U_{\gamma_{i+1}} \cdots U_{\gamma_j} \widetilde{\psi}_j \cdot \mathcal{L}^i_{\boldsymbol{\gamma}} 1] \\
&\quad = \mathbb{E}[\widetilde{\psi}_j \cdot \mathcal{L}_{\gamma_j} \cdots \mathcal{L}_{\gamma_{i+1}}(\widetilde{\psi}_i \cdot \mathcal{L}^i_{\boldsymbol{\gamma}} 1)] \\
&\quad = \int \widetilde{\psi}_j(x) \cdot \mathcal{L}_{\gamma_j} \cdots \mathcal{L}_{\gamma_{i+1}} \left(\psi \cdot \mathcal{L}^i_{\boldsymbol{\gamma}} 1 - \mathcal{L}^i_{\boldsymbol{\gamma}} 1 \int \psi \mathcal{L}^i_{\boldsymbol{\gamma}} 1 \, dm \right)(x) \, dm(x).
\end{aligned}$$

Now $\varphi_1 := \mathcal{L}^i_{\boldsymbol{\gamma}} 1(x) \cdot \psi(x)$ and $\varphi_2 := (\int \psi \mathcal{L}^i_{\boldsymbol{\gamma}} 1 \, dm) \cdot \mathcal{L}^i_{\boldsymbol{\gamma}} 1(x)$ are positive densities with the same expectation, and both belong to $C(2a, \nu)$. In fact, $\varphi_2 \in C(a, \nu) \subset C(2a, \nu)$, and φ_1 is the product of two densities in $C(a, \nu)$, which is a density of $C(2a, \nu)$. So, letting $r \in (0, 1)$ be the uniform contraction rate of the operators $\{\mathcal{L}_\gamma\}_{\gamma \in \Gamma}$ on $C(2a, \nu)$,

$$\begin{aligned}
\theta_+(\mathcal{L}_{\gamma_j} \cdots \mathcal{L}_{\gamma_{i+1}}(\varphi_1), \mathcal{L}_{\gamma_j} \cdots \mathcal{L}_{\gamma_{i+1}}(\varphi_2)) &\le \theta_{2a,\nu}(\mathcal{L}_{\gamma_j} \cdots \mathcal{L}_{\gamma_{i+1}}(\varphi_1), \mathcal{L}_{\gamma_j} \cdots \mathcal{L}_{\gamma_{i+1}}(\varphi_2)) \\
&\le r^{j-i-1} \theta_{2a,\nu}(\varphi_1, \varphi_2)
\end{aligned}$$

and

$$\begin{aligned}
&\sup_x \left| \mathcal{L}_{\gamma_j} \cdots \mathcal{L}_{\gamma_{i+1}}(\mathcal{L}^i_{\boldsymbol{\gamma}} 1(x) \psi(x)) - \mathcal{L}_{\gamma_j} \cdots \mathcal{L}_{\gamma_{i+1}} \left(\mathcal{L}^i_{\boldsymbol{\gamma}} 1(x) \int \psi \mathcal{L}^i_{\boldsymbol{\gamma}} 1 \, dm \right) \right| \\
&\quad \le R_1[\exp(R_2 r^{j-i+1}) - 1]
\end{aligned}$$

with

$$\begin{aligned}
R_1 &:= \sup\{\varphi(x) | x \in M, \varphi \in C(2a, \nu)\}, \\
R_2 &:= \sup\{\theta_{2a,\nu}(\varphi_1, \varphi_2) | \varphi_1, \varphi_2 \in C(2a, \nu)\}.
\end{aligned}$$

Since $(\int \widetilde{\psi}_i \, dm)$ is uniformly bounded with respect to i, we can upper bound correlations with $R'_1[\exp(R_2 r^{j-i+1}) - 1]$, from which the thesis follows. □

Proof of Part (3) of Theorem B. We know that

$$\begin{aligned}
\mathbb{E}[\psi_k] &= \int \psi_k \, dm \\
&= \int U_{\gamma_1} \cdots U_{\gamma_k}(\psi) \, dm \\
&= \int \psi \mathcal{L}_{\gamma_k} \cdots \mathcal{L}_{\gamma_1} 1 \, dm
\end{aligned}$$

and, using Theorem B, since for every $k > \overline{n}$ and sufficiently small perturbations $\|\mathcal{L}_{\gamma_k} \cdots \mathcal{L}_{\gamma_1} 1 - \varphi_0\|_0 < \varepsilon$,

$$\left| \mathbb{E}[\psi_k] - \int \psi \varphi_0 \, dm \right| \le \varepsilon \mathbb{E}[|\psi|]$$

and

$$\begin{aligned}
\int \psi \varphi_0 \, dm - \varepsilon \mathbb{E}[|\psi|] &\le \liminf_{n \to \infty} \frac{1}{n} \sum_k \mathbb{E}[\psi_k] \\
&\le \limsup_{n \to \infty} \frac{1}{n} \sum_k \mathbb{E}[\psi_k] \le \int \psi \varphi_0 \, dm + \varepsilon \mathbb{E}[|\psi|].
\end{aligned}$$

Thanks to Lemma A.1, we can directly apply Theorem 3 to the random variables $X_k = \psi_k$. This implies that, for almost every $x \in M$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \psi_k(x) - \mathbb{E}[\psi_k] = 0,$$

which implies that

$$\int \psi\varphi_0 \, dm - \varepsilon\mathbb{E}[|\psi|] \le \liminf_{n\to\infty} \frac{1}{n} \sum_k \psi_k(x)$$
$$\le \limsup_{n\to\infty} \frac{1}{n} \sum_k \psi_k(x) \le \int \psi\varphi_0 \, dm + \varepsilon\mathbb{E}[|\psi|]. \qquad \square$$

References

[**AA03**] J. F. Alves and V. Araujo. Random perturbations of nonuniformly expanding maps. *Astérisque* **286** (2003), 25–62.

[**AHNT15**] R. Aimino, H. Hu, M. Nicol, A. Török and S. Vaienti. Polynomial loss of memory for maps of the interval with a neutral fixed point. *Discrete Contin. Dyn. Syst.* **35**(3) (2015), 793–806.

[**AR15**] R. Aimino and J. Rousseau. Concentration inequalities for sequential dynamical systems of the unit interval. *Ergod. Th. & Dynam. Sys.* **36**(8) (2016), 2384–2407.

[**AT05**] V. Araujo and A. Tahzibi. Stochastic stability at the boundary of expanding maps. *Nonlinearity* **18**(3) (2005), 939–958.

[**Bal00**] V. Baladi. *Positive Transfer Operators and Decay of Correlations*. World Scientific, River Edge, NJ, 2000.

[**BG89**] A. Boyarsky and P. Gora. Absolutely continuous invariant measures for piecewise expanding C^2 transformations in $\mathbb{R}^n$. *Israel J. Math.* **67**(3) (1989), 272–286.

[**BG97**] A. Boyarsky and P. Gora. *Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension*. Birkhauser, Boston, MA, 1997.

[**BKL02**] M. Blank, G. Keller and C. Liverani. Ruelle–Perron–Frobenius spectrum for Anosov maps. *Nonlinearity* **15**(6) (2002), 1905–1973.

[**BY93**] V. Baladi and L.-S. Young. On the spectra of randomly perturbed expanding maps. *Comm. Math. Phys.* **156**(2) (1993), 355–385.

[**Cow00**] W. J. Cowieson. Stochastic stability for piecewise expanding maps in R^d. *Nonlinearity* **13** (2000), 1745–1760.

[**CR07**] J.-P. Conze and A. Raugi. Limit theorems for sequential expanding dynamical systems on [0, 1]. *Contemp. Math.* **430** (2007), 89–122.

[**DS16**] N. Dobbs and M. Stenlund. Quasistatic dynamical systems. *Ergod. Th. & Dynam. Sys.* (2016), 1–41 doi:10.1017/etds.2016.9.

[**GOT13**] C. Gupta, W. Ott and A. Török. Memory loss for time-dependent piecewise-expanding systems in higher dimension. *Math. Res. Lett.* **20**(1) (2013), 141–161.

[**HNTV17**] N. Haydn, M. Nicol, A. Török and S. Vaienti. Almost sure invariance principle for sequential and non-stationary dynamical systems. *Trans. Amer. Math. Soc.* **369**(8) (2017), 5293–5316.

[**ITM50**] C. T. Ionescu-Tulcea and G. Marinescu. Théorie ergodique pour des classes d'opérations non complètement continues. *Ann. of Math. (2)* **52**(1) (1950), 140–147.

[**Kel82a**] G. Keller. Stochastic stability in some chaotic dynamical systems. *Monatsh. Math.* **94**(4) (1982), 313–333 in English.

[**Kel82b**] G. Keller. Stochastic stability of some one-dimensional dynamical systems. *Ergodic Theory and Related Topics, Proc. Conf. held in Vitte/Hiddensee, 19–23 October 1981*. Akademie, 1982, pp. 123–127.

[**KL98**] G. Keller and C. Liverani. Stability of the spectrum for transfer operators. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (4)* **28** (1998), 141–152.

[**KR11**] P. E. Kloeden and M. Rasmussen. *Nonautonomous Dynamical Systems*. American Mathematical Society, Providence, RI, 2011, 176.

[**Liv95a**] C. Liverani. Decay of correlations. *Ann. of Math. (2)* **142**(2) (1995), 239–301.

[**Liv95b**] C. Liverani. Decay of correlations for piecewise expanding maps. *J. Stat. Phys.* **78**(3–4) (1995), 1111–1129.

[**Liv01**] C. Liverani. Rigorous numerical investigation of the statistical properties of piecewise expanding maps. A feasibility study. *Nonlinearity* **14**(3) (2001), 463–490.

[**LM13**] S. Luzzatto and I. Melbourne. Statistical properties and decay of correlations for interval maps with critical points and singularities. *Comm. Math. Phys.* **320**(1) (2013), 21–35.

[**LM85**] A. Lasota and M. C. Mackey. *Probabilistic Properties of Deterministic Systems*. Cambridge University Press, New York, 1985.

[**LY73**] A. Lasota and J. A. Yorke. On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.* **186** (1973), 481–488.

[**NSV12**] P. Nándori, D. Szász and T. Varjú. A central limit theorem for time-dependent dynamical systems. *J. Stat. Phys.* **146**(6) (2012), 1213–1220.

[**OGY90**] E. Ott, C. Grebogi and J. A. Yorke. Controlling chaos. *Phys. Rev. Lett.* **64**(11) (1990), 1196–1199.

[**OSY09**] W. Ott, M. Stenlund and L.-S. Young. Memory loss for time-dependent dynamical systems. *Math. Res. Lett.* **16**(3) (2009), 436–475.

[**Sau00**] B. Saussol. Absolutely continuous invariant measures for multidimensional expanding maps. *Israel J. Math.* **116**(1) (2000), 223–248.

[**Ste11**] M. Stenlund. Non-stationary compositions of Anosov diffeomorphisms. *Nonlinearity* **24**(10) (2011), 2991–3018.

[**SvS13**] W. Shen and S. van Strien. On stochastic stability of expanding circle maps with neutral fixed points. *Dyn. Syst.* **28**(3) (2013), 423–452.

[**SYZ13**] M. Stenlund, L.-S. Young and H. Zhang. Dispersing billiards with moving scatterers. *Comm. Math. Phys.* **322**(3) (2013), 909–955.

[**Via97**] M. Viana. *Stochastic Dynamics of Deterministic Systems*. IMPA, Rio de Janeiro, 1997.

[**Wal04**] H. Walk. Strong laws of large numbers by elementary Tauberian arguments. *Monatsh. Math.* **144**(4) (2004), 329–346.

[**You98**] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math. (2)* **147**(3) (1998), 585–650.

[**You99**] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.* **110**(1) (1999), 153–188.