Extensions with shrinking fibers

BENOÎT R. KLOECKNER [ORCID]

LAMA, Univ. Paris-Est Créteil, Univ. Gustave Eiffel, UPEM, CNRS,
F-94010, Créteil, France
(e-mail: benoit.kloeckner@u-pec.fr)

Abstract. We consider dynamical systems $T : X \to X$ that are extensions of a factor $S : Y \to Y$ through a projection $\pi : X \to Y$ with *shrinking* fibers, that is, such that T is uniformly continuous along fibers $\pi^{-1}(y)$ and the diameter of iterate images of fibers $T^n(\pi^{-1}(y))$ uniformly go to zero as $n \to \infty$. We prove that every S-invariant measure $\check{\mu}$ has a unique T-invariant lift μ, and prove that many properties of $\check{\mu}$ lift to μ: ergodicity, weak and strong mixing, decay of correlations and statistical properties (possibly with weakening in the rates). The basic tool is a variation of the Wasserstein distance, obtained by constraining the optimal transportation paradigm to displacements along the fibers. We extend classical arguments to a general setting, enabling us to translate potentials and observables back and forth between X and Y.

Contents

1. *Introduction*

Let $T : X \to X$ be a dynamical system where X is a compact metric space, and assume that T has a topological factor $S : Y \to Y$, that is, there is a continuous onto map $\pi : X \to Y$ such that $\pi T = S\pi$. Each fiber $\pi^{-1}(y) \subset X$ is collapsed under π into a single point y,

CrossMark

and S can thus be thought of as a simplification of T, which may retain certain of its dynamical properties but forget others. When additionally T shrinks the fibers, that is, two points x, x' such that $\pi(x) = \pi(x')$ have orbits that are attracted one to another, one suspects that actually *all* important dynamical features of T survive in S: along the fibers, the dynamic is trivial anyway. It might still happen that S is easier to study than T, in which case one can hope to obtain interesting dynamical properties of T by proving them for S and lifting them back. The present paper aims to develop a systematic machinery to do that in the context of the thermodynamical formalism, that is, the study of equilibrium states (invariant measure optimizing a linear combination of entropy and energy with respect to a potential).

This setting has already been much studied, first in the symbolic case of $X = \{0, 1\}^{\mathbb{Z}}$ (or a subshift), $Y = \{0, 1\}^{\mathbb{N}}$ (or the corresponding one-sided subshift), π the map that forgets negative indexes, T and S the left shifts. Then the strategy outlined above was used for a long time; see, for example, [**Bow08**]. The advantage of one-sided shifts is that an orbit can be looked at backward in time as a non-trivial, contracting Markov chain; one can use this to prove existence, uniqueness and statistical properties of equilibrium states for a wide range of potentials. The same reason makes expanding maps rather easier to study than hyperbolic ones. Recently, several works have used the above approach to study various flavors of hyperbolic dynamics on manifolds or domains of $\mathbb{R}^n$. However, they are often written for specific systems and the technical details are often not obviously generalizable. Moreover, the basic result that an S-invariant measure of S has a *unique* T-invariant lift seems not to be known in general. Our first aim will be to propose a simple and general argument, based on ideas from optimal transportation, to lift invariant measures and show uniqueness. Then we shall use uniqueness and adapt folklore methods to a general framework to lift a rather complete set of properties of invariant measures.

While the dynamical study of uniformly hyperbolic maps is considered reasonably well understood, the study of various kinds of non-uniformly hyperbolic maps has witnessed a great deal of activity in the last two decades; see, for example, [**You98**, **ABV00**, **AMV15**, **ADLP17**], the surveys [**Alv15**, **CP15**] and other references cited below. Even in the uniformly expanding case, new approaches are welcome; see [**CPZ18**]. As is well known, the 'extension' approach can be used to study certain uniformly and non-uniformly hyperbolic maps, when the default of hyperbolicity can be in the contraction or the expansion, or when the potential lacks Hölder regularity (see §1.3 and Remark 1.12).

1.1. *Shrinking fibers: main definition and first main result.*   All measures considered are probability measures, and we denote by $\mathcal{P}(X)$ the set of measures on X and by $\mathcal{P}_T(X)$ the set of T-invariant measures, both endowed with the weak-$*$ topology.

We shall consider the case where the extension T exhibits some contraction along fibers; we introduce a single notion that includes a global property and continuity along fibers (details on moduli of continuity are recalled in §2.1).

Definition 1.1. We say that T is *uniformly continuous* along fibers whenever there is a modulus of continuity $\bar{\omega}_T$ such that, for all x, $x' \in X$ with $\pi(x) = \pi(x')$,

$$d(Tx, Tx') \leq \bar{\omega}_T(d(x, x')).$$

We say that T is an extension of S with *shrinking fibers* (keeping implicit π and S) whenever T is uniformly continuous along fibers and there is a sequence $(a_n)_{n \in \mathbb{N}}$ of positive numbers such that $\lim a_n = 0$ and, for all $x, x' \in X$ with $\pi(x) = \pi(x')$,

$$d(T^n x, T^n x') \leq a_n.$$

If $a_n = C\theta^n$ for some $\theta \in (0, 1)$, (respectively, $a_n = Cn^{-d}$ for some $d > 0$) we may specify that T has *exponentially* (respectively, *polynomially*) shrinking fibers, of *ratio* θ (respectively, *degree d*). We may specify the *shrinking sequence* $(a_n)_n$, for example, by saying that T has $(a_n)_n$-shrinking fibers.

For example, if, for some $\theta \in (0, 1)$ and all $x, x' \in X$ such that $\pi(x) = \pi(x')$, we have $d(Tx, Tx') \leq \theta d(x, x')$, then T has exponentially shrinking fibers; however, the latter property is weaker.

Of course, if T is continuous then it is uniformly continuous along fibers; the first part of the above definition is meant to make it possible to deal quite generally with discontinuous maps S. It will be used mainly in the proof of Theorem 3.1, which is at the core of Theorem A below.

Some of our results actually hold more generally, and to state them in their full scope we introduce the following notion.

Definition 1.2. Given $\check{\mu} \in \mathcal{P}_S(Y)$, we say that T is an extension of S whose *fibers are shrunk on average with respect to $\check{\mu}$* whenever T is uniformly continuous along fibers and there is a sequence $(\bar{a}_n)_{n \in \mathbb{N}}$ of positive numbers such that $\lim_n \bar{a}_n = 0$ and

$$\int \operatorname{diam}(T^n(\pi^{-1}(y))) \, d\check{\mu}(y) \leq \bar{a}_n \quad \text{for all } n \in \mathbb{N}.$$

At certain points we will also need some mild additional regularity for π.

Definition 1.3. We say that the projection π is *non-singular* with respect to positive measures λ_X and λ_Y when $\pi_* \lambda_X$ is equivalent to λ_Y, that is, when, for all Borel sets $B \subset Y$,

$$\lambda_Y(B) > 0 \Leftrightarrow \lambda_X(\pi^{-1}(B)) > 0.$$

Definition 1.4. We say that π *induces a continuous fibration* whenever, for some modulus of continuity $\bar{\omega}_\pi$ and all $y, y' \in Y$, for all $x \in \pi^{-1}(y)$, there exists $x' \in \pi^{-1}(y')$ such that

$$d(x, x') \leq \bar{\omega}_\pi(d(y, y')).$$

Then by the measurable selection theorem, there exists a measurable map $\tau_y^{y'} : \pi^{-1}(y) \to \pi^{-1}(y')$ such that $d(x, \tau_y^{y'}(x)) \leq \bar{\omega}_\pi(d(y, y'))$ for all $x \in \pi^{-1}(y)$.

The modulus of continuity $\bar{\omega}_\pi$ should not be confused with the modulus of continuity of the map π, which whenever needed will be denoted by ω_π.

With these definitions set up, we can state our first results gathered in the following statement.

THEOREM A. *If T is an extension of S with shrinking fibers, then π_* induces a homeomorphism from $\mathcal{P}_T(X)$ to $\mathcal{P}_S(Y)$. In particular, for each S-invariant measure $\check{\mu}$ there is a unique T-invariant measure μ such that $\pi_* \mu = \check{\mu}$.*

Moreover, for all (not necessarily invariant) $\nu \in \mathcal{P}(X)$ such that $\pi_* \nu = \check{\mu}$, we have $T_*^n \nu \to \mu$ in the weak-$*$ topology, and:

(i)   *each of the following adjectives applies to μ (with respect to T) if and only if it applies to $\check{\mu}$ (with respect to S): ergodic, weakly mixing, strongly mixing;*

(ii) *if T is continuous and X, Y have reference measures with respect to which π is non-singular, then each of the following adjectives applies to μ if and only if it applies to $\check{\mu}$: physical, observable;*

(iii) *if S, T are continuous, then μ and $\check{\mu}$ have the same Kolmogorov–Sinai entropy.*

Theorem A is unsurprising, and some parts of it are already known in more or less general settings (e.g. lifting of physicality); however, the uniqueness of the lift was not known in general, and greatly simplifies the proof of further properties. It is, in particular, interesting to compare the existence and uniqueness part of Theorem A to [**APPV09**, §6.1] where Araujo, Pacifico, Pujals and Viana construct a lift of $\check{\mu}$. The first advantage of our result is that we prove uniqueness among all invariant measures, while they get uniqueness only under the property they use in the construction. Second, we need milder assumptions (see Remark 3.2).

Castro and Nascimento have studied in [**CN17**] two kinds of maps, the first one fitting in the theme of the present paper. Namely, they consider the case where S is a non-uniformly expanding map in the family introduced by Castro and Varandas [**CV13**] and T is exponentially contracting along fibers. They focus there on the maximal entropy measure for T, proving it exists, is unique, and enjoys exponential decay of correlations and a central limit theorem for Hölder observables. Leaving aside the statistical properties for now, Theorem A in particular shows that existence and uniqueness of the maximal entropy measure for T does not depend on the specifics of S or on the rate of contraction along fibers (as stated, we actually do not even need T to be a contraction along fibers, only to shrink them globally): item (iii) is a broad generalization of [**CN17**, Theorem A] since, under the sole assumptions that S, T are continuous and that fibers are shrinking, it shows that T has a unique measure of maximal entropy if and only if S does.

1.2. *Further main results: thermodynamical formalism.*   In our subsequent results we shall assume T is Lipschitz, and this hypothesis deserves an explanation. We will often need to work in some functional spaces where observables or potentials are taken, and we made the choice of *generalized Hölder spaces*, that is, spaces of functions with modulus of continuity at most a multiple of some reference, arbitrary modulus. This choice seems a good balance between generality (it includes functions less regular than Hölder, enabling us to consider in particular polynomial rates of shrinking) and clarity (proofs stay pretty simple and the amount of definition needed is significant but not overwhelming). It is often a crucial ingredient that the iterated Koopman operators $f \mapsto f \circ T^k$ are bounded on the chosen functional space, with good control of their norms; requiring T to be Lipschitz is the natural hypothesis to ensure this for generalized Hölder spaces. Were one interested in discontinuous maps (e.g. when S is discontinuous), the principle of proofs could certainly be adapted but one would need (as usual) to work in a suitable functional spaces. Another advantage of our choice is that we can work directly with the Wasserstein distance between measures.

The convergence result ($T^n \nu \to \mu$ whenever $\pi_* \nu = \check{\mu}$) in Theorem A seems new in this generality. It is, however, not as satisfying as those obtained by Galatolo and Lucena in [**GL15**, §5.1] in their particular setting, where instead of $\pi_* \nu = \check{\mu}$ it is only required that $\pi_* \nu$ is absolutely continuous with respect to $\check{\mu}$ (with some regularity assumptions on the density). In this direction, we prove the following variation of [**GL15**, §5.1] (our hypotheses are quite general, but we assume T to be Lipschitz and our convergence is in the Wasserstein metric instead of the particular metric constructed in [**GL15**]).

THEOREM B. *Assume that T is an extension of S with exponentially shrinking fibers; that S admits a conformal measure $\lambda_Y \in \mathcal{P}(Y)$ such that the associated transfer operator has a spectral gap on some Banach space $\mathcal{B}$, whose normalized eigenfunction is denoted by h; that T is Lipschitz; and that π induces a Hölder-continuous fibration. Let μ be the unique T-invariant lift of $h \, d\lambda_Y$. Then, for all $\nu \in \mathcal{P}(X)$ such that $\pi_* \nu$ is absolutely continuous with respect to λ_Y with density in $\mathcal{B}$, the sequence $(T_*^n \nu)_n$ converges to μ exponentially fast in the Wasserstein metric.*

The classical definitions needed are given in §2; in particular, the Wasserstein metric is defined in §2.4 (for now, let us simply say that in our compact setting it metrizes the weak-$*$ topology) and transfer operators are defined in §2.5. A more general (but less precise) result is given in Corollary 3.5.

We now turn to equilibrium states and their statistical properties. It will be convenient to use the following definition (the reader may want to have a look at §2.1 on moduli of continuity and generalized Hölder spaces; in particular, we shall use the very mild modulus of continuity $\omega_{\alpha \log}(r) \simeq (\log(1/r))^{-\alpha}$).

Definition 1.5. Let ω_p, ω_o be moduli of continuity, let $\rho \in (0, \infty]$ and let $\mathcal{T}$ be the name of a limit theorem for discrete-time random processes (e.g. LIL for the law of the iterated logarithm, CLT for the central limit theorem, or ASIP for the almost sure invariance principle; see Definition 2.19). We shall say that *T has unique equilibrium states for potentials in* $\mathrm{Hol}_{\omega_p}(X)$ *of norm less than ρ, with limit theorem $\mathcal{T}$ for observables in* $\mathrm{Hol}_{\omega_o}(X)$ (in short, that T satisfies $\mathrm{UE}(\omega_p[\rho]; \mathcal{T}, \omega_o)$) whenever:

(i) for all potentials $\varphi \in \mathrm{Hol}_{\omega_p}(X)$ such that $\|\varphi\|_{\omega_p} < \rho$, there exists a unique *equilibrium state* μ_φ, that is, a maximizer of the free energy $\mathcal{F}(\mu) = h_{\mathrm{KS}}(T, \mu) + \mu(\varphi)$ over all $\mu \in \mathcal{P}_T(X)$ (where h_{KS} denotes Kolmogorov–Sinai entropy); and

(ii) for all $f \in \mathrm{Hol}_{\omega_o}(X)$, the random process $(f \circ T^k(Z))_{k \in \mathbb{N}}$, where Z is a random variable with law μ_φ, satisfies the limit theorem $\mathcal{T}$ (see Definition 2.19 for further details).

When there is no bound on the norm of potential (i.e. $\rho = \infty$), we may shorten $\mathrm{UE}(\omega_p[\infty]; \mathcal{T}, \omega_o)$ to $\mathrm{UE}(\omega_p; \mathcal{T}, \omega_o)$. When this property is satisfied for all Hölder-continuous potentials or observables, whatever the Hölder exponent, we write ω_* in place of ω_p or ω_o. When we only want to state existence and uniqueness of equilibrium state, we agree to take $\mathcal{T} = \varnothing$ and we can simplify the notation to $\mathrm{UE}(\omega_p; \varnothing)$ as there is no need to specify the observables.

THEOREM C. *Assume that T is L-Lipschitz and that π is β-Hölder and admits a Lipschitz section. Let $\mathcal{T} \in \{\varnothing, \mathrm{LIL}, \mathrm{CLT}, \mathrm{ASIP}\}$.*

(i)   *Assume that the fibers are exponentially shrinking with ratio θ, let $\alpha \in (0, 1]$ and set $\gamma = \alpha\beta/(1 - \log L / \log \theta)$. If S satisfies $\mathrm{UE}(\boldsymbol{\omega}_\gamma[\rho]; \mathcal{T}, \boldsymbol{\omega}_*)$ then T satisfies $\mathrm{UE}(\boldsymbol{\omega}_\alpha[C\rho]; \mathcal{T}, \boldsymbol{\omega}_*)$ for some $C > 0$.*

(ii) *Assume that the fibers are polynomially shrinking with degree $d > 1$, consider $\alpha, \gamma \in (1/d, 1]$ and set $\alpha' = \alpha d - 1$ and $\gamma' = \gamma d - 1$. If S satisfies $\mathrm{UE}(\boldsymbol{\omega}_{\alpha' \log}[\rho]; \mathcal{T}, \boldsymbol{\omega}_{\gamma' \log})$, then T satisfies $\mathrm{UE}(\boldsymbol{\omega}_\alpha[C\rho]; \mathcal{T}, \boldsymbol{\omega}_\gamma)$ for some $C > 0$.*

(iii) *Assume that the fibers are exponentially shrinking, consider $\alpha, \gamma > 1$ and let $\alpha' = (\alpha - 1)/2$ and $\gamma' = (\gamma - 1)/2$. If S satisfies $\mathrm{UE}(\boldsymbol{\omega}_{\alpha' \log}[\rho]; \mathcal{T}, \boldsymbol{\omega}_{\gamma' \log})$ then T satisfies $\mathrm{UE}(\boldsymbol{\omega}_{\alpha \log}[C\rho]; \mathcal{T}, \boldsymbol{\omega}_{\gamma \log})$ for some $C > 0$.*

Many other combinations of moduli of continuity and shrinking speed can be considered; see Theorem 5.5. The main tool is to construct from a potential or an observable on X a suitable potential or observable on Y. For this, we generalize a method that is classical in the symbolic setting: adding a coboundary to make the potential or observable constant along fibers.

Item (i) generalizes [**CN17**, Theorems A and C]: Castro and Nascimento were concerned with the maximal entropy measure, that is, the equilibrium state for the null potential, while item (i) provides in their setting (using the known results for S a Castro–Varandas map [**CV13**]) existence, uniqueness and CLT for Hölder observables for the equilibrium state of any Hölder potential of small enough norm. The generalization is actually far broader, since one can take much more varied base maps S for which equilibrium states have the desired limit theorem (see, for example, [**MT02**, **MN05**, **Gou10**] for the ASIP). Some examples will be provided below.

Interestingly, it appears more efficient to directly lift limit theorems from S to T than to lift decay of correlations and then use them to prove limit theorems for T. Nevertheless, decays of correlations have a long history and are prominent features of invariant measures, and it thus makes sense to lift them as well. In this regard, we obtain the following result.

THEOREM D. *Let T be a Lipschitz extension of S with shrinking fibers, assume that there is a Lipschitz section σ, let μ be T-invariant probability measure and $\check{\mu} := \pi_*\mu$ be the corresponding S-invariant measure.*

(i)   *If, for some $\alpha_0 \in (0, 1]$, the transfer operator $\check{\mathscr{L}}$ of $(S, \check{\mu})$ has a spectral gap in each Hölder space $\mathrm{Hol}_\alpha(Y)$ ($\alpha \in (0, \alpha_0]$) and if fibers are exponentially shrinking, then μ has exponential decay of correlation in each Hölder space.*

(ii) *If, for some $\alpha \in (0, 1]$ and all $n \in \mathbb{N}$, the transfer operator $\check{\mathscr{L}}$ of $(S, \check{\mu})$ satisfies*

$$\mathrm{Hol}_\alpha(\check{\mathscr{L}}^n h) \lesssim \mathrm{Hol}_\alpha(h) \quad and \quad \|\check{\mathscr{L}}^n f\|_\infty \lesssim \frac{\mathrm{Hol}_{\alpha d \log}(f)}{n^p}$$

whenever $h \in \mathrm{Hol}_\alpha(Y)$, $f \in \mathrm{Hol}_{\alpha d \log}(Y)$ with $\check{\mu}(f) = 0$, and if the fibers of π are polynomially shrinking of degree d, then μ has polynomial decay of correlation of degree $\min(\alpha d, p/2)$ for all $\boldsymbol{\omega}_\alpha$-continuous observables.

(iii) *If, for some $\alpha > 0$ and all $n \in \mathbb{N}$, the transfer operator $\check{\mathscr{L}}$ of $(S, \check{\mu})$ satisfies*

$$\mathrm{Hol}_{\alpha \log}(\check{\mathscr{L}}^n h) \lesssim \mathrm{Hol}_{\alpha \log}(h) \quad and \quad \|\check{\mathscr{L}}^n f\|_\infty \lesssim \frac{\mathrm{Hol}_{(\alpha/2) \log}(f)}{n^p}$$

whenever $h \in \mathrm{Hol}_{\alpha \log}(Y)$, $f \in \mathrm{Hol}_{(\alpha/2) \log}(Y)$ with $\breve{\mu}(f) = 0$, and if the fibers of π are exponentially shrinking, then μ has polynomial decay of correlation of degree $\min(\alpha, p/2)$ for all $\boldsymbol{\omega}_{\alpha \log}$-continuous observables.

To prove this result, our main tool is expected: we prove the regularity of the disintegration of μ with respect to π (Theorem 6.3). Such results appeared in the work of Galatolo and Pacifico [**GP10**, Appendix A] (extra difficulty in the proof there seems to be caused by the way disintegration is set up, making it necessary to deal with non-probability measures) and in the recent works of Butterley and Melbourne [**BM17**, Proposition 6] (cf. Theorem 6.3 below) and Araujo, Galatolo and Pacifico [**AGP14**, Theorem A]. Compared to these works, we gain in generality: we can consider very general maps while they tend to restrict to uniformly expanding maps, and we consider an arbitrary S-invariant measure instead of restricting to the absolutely continuous one. Items (ii) and (iii) have no equivalent that I know of in the literature.

1.3. *A few examples.* A commonly studied situation where our framework applies readily is that of skew-products, where $X = Y \times \Phi$ for some compact metric space Φ and

$$T : x = (y, \phi) \mapsto (S(y), R(y, \phi)).$$

The fact that T shrinks fibers then translates into $d(R^n(y, \phi), R^n(y, \phi')) \le a_n$ for all $n \in \mathbb{N}$ and all $\phi, \phi' \in \Phi$ where $a_n \to 0$, $R^1 = R$ and $R^{n+1}(y, \phi) = R(S(y), R^n(y, \phi))$. The projection map is then $\pi : (y, \phi) \to y$ and all needed hypotheses on π in Theorems A–C are easy to check, endowing for example X with the metric $d((y, \phi), (y', \phi')) = \sqrt{d(y, y')^2 + d(\phi, \phi')^2}$. Note that since we will apply our above results, in many cases we will assume T (and thus S and R) to be Lipschitz; and in all cases our 'shrinking fiber' hypothesis implies that $R(y, \phi)$ depends continuously on the variable ϕ when y is fixed.

Remark 1.6. One can easily generalize this setting to fiber bundles: X is then no longer a product, but there exist a compact metric space Φ and a *fibered atlas* $(U_i, h_i)_{i \in I}$, that is, the U_i form an open cover of Y and the h_i are homeomorphisms from $U_i \times \Phi$ to $\pi^{-1}(U_i)$ such that $h_i(y \times \Phi) = \pi^{-1}(y)$. The simplest example of a fiber bundle that is not a product is the Möbius band, together with the usual projection on the circle. In this setting, T is required to send fibers into fibers and is *locally* of the form $(y, \phi) \mapsto (S(y), R(y, \phi))$ (where the charts h_i are used to identify $\pi^{-1}(U_i)$ with a product). Our main results are stated in an even more general framework, and in order to aim for simplicity we shall restrict the examples to product spaces, but fiber bundles seem unjustly under-represented in the dynamical literature.

We shall consider examples spanning all the following weaknesses of the system to be considered: non-uniform expansion in the 'horizontal' direction; slow shrinking in the 'vertical' direction; and low-regularity of potentials (and observables).

Let us recall the classical benchmark for non-uniformly expanding maps, the Pomeau–Manneville family defined on the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ (identified with $[0, 1)$) by

$$S_q : \mathbb{T} \to \mathbb{T}$$

$$y \mapsto \begin{cases} (1 + (2y)^q)y & \text{if } y \in [0, \tfrac{1}{2}], \\ 2y - 1 & \text{if } y \in [\tfrac{1}{2}, 1), \end{cases}$$

where $q \geq 0$ (when $q = 0$ we get the doubling map, which is *uniformly* expanding). Let Φ be a compact metric space endowed with a reference (finite, positive) measure λ_Φ, denote by $\lambda_\mathbb{T}$ the Lebesgue measure on the circle, and endow $X = \mathbb{T} \times \Phi$ with the reference measure $\lambda_X := \lambda_\Phi \times \lambda_\mathbb{T}$.

COROLLARY 1.7. *Assume $T : \mathbb{T} \times \Phi \to \mathbb{T} \times \Phi$ is a continuous skew-product with base map S_q for some $q \in (0, 1)$ and with shrinking fibers.*
(i)   *T admits a unique physical measure μ_T.*
(ii)   *Assume further that $q < \tfrac{1}{2}$, that fibers are exponentially shrinking and that T is Lipschitz. Then μ_T satisfies the ASIP for all Hölder-continuous observables.*
(iii) *If T is L-Lipschitz and fibers are exponentially shrinking with ratio θ and if $q' := q(1 - \log L / \log \theta)$ is less than 1, then for all $\alpha \in (q', 1]$ the map T satisfies $\mathrm{UE}(\boldsymbol{\omega}_\alpha; \mathrm{CLT}, \boldsymbol{\omega}_*)$; that is, each α-Hölder potential φ has a unique equilibrium state μ_φ, for which Hölder-continuous observables satisfy the central limit theorem. Moreover, Hölder observables have exponentially decaying correlations with respect to μ_φ.*

Let us say that a map $S : Y \to Y$ is *uniformly expanding* when it is a self-covering map of degree k, and there is some $\theta \in (0, 1)$ such that, for each $y, y' \in Y$, denoting by $z_1, \ldots, z_k$ and $z_1', \ldots, z_k'$ the inverse images of y and y', there is a permutation σ such that $d(z_i, z_{\sigma(i)}') \leq \theta d(y, y')$ for all i. We assume here for simplicity that Y is a manifold endowed with a volume form yielding a reference measure λ_Y and that $X = Y \times \Phi$ is again endowed with a product measure $\lambda_X = \lambda_Y \times \lambda_\Phi$.

COROLLARY 1.8. *Assume that T is a Lipschitz skew-product with base map S a uniformly expanding map and with polynomially shrinking fibers of degree $d > 2$, and let $\alpha \in (2/d, 1]$. Then, for all α-Hölder potential φ:*
(i)   *T has a unique equilibrium state μ_φ;*
(ii)   *if $\alpha > 5/(2d)$, then T satisfies $\mathrm{UE}(\boldsymbol{\omega}_\alpha; \mathrm{CLT}, \boldsymbol{\omega}_{\alpha-1/d})$: we have the central limit theorem for μ_φ and for all $(\alpha - 1/d)$-Hölder observables;*
(iii) *for all $\gamma \in [\alpha - 1/d, 1]$, all γ-Hölder observables have polynomial decay of correlations of degree $\alpha d / 2 - 1$ with respect to μ_φ,*

Note that the second item is not enough to obtain the third: when α is only slightly above $5/(2d)$, we get decay of correlation of degree only slightly above $1/4$, while degree $1/2$ would be a minimum to obtain the CLT. This is a sign that Theorem D might not be optimal.

Let us now consider low-regularity maps, that is, below the $\mathscr{C}^{1,\alpha}$ regularity.

COROLLARY 1.9. *Assume that N is a manifold admitting a uniformly expanding $\mathscr{C}^{1,\alpha \log}$ map $S : N \to N$ with expanding factor λ, that D is a d-dimensional closed ball with*
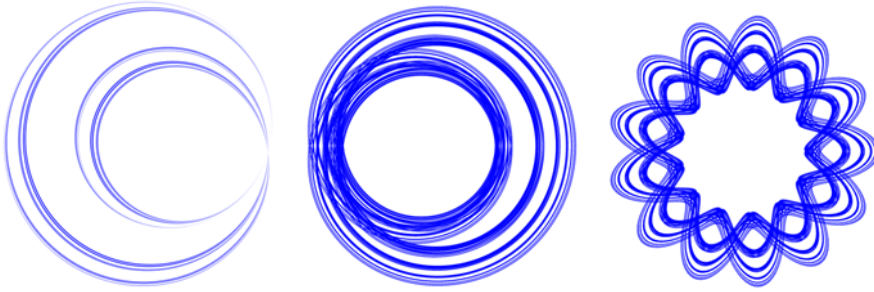
FIGURE 1. Some attractors on the annulus: the images picture $T_i^{n_i}(U)$ with U a neighborhood of the attractor, $n_1 = 4$, $n_2 = 9$ and $n_3 = 8$. The left-hand attractor is homeomorphic to the product of a Cantor set with a circle, pinched at one fiber (to achieve this, the map T_1 pinches two fibers to a common point). The topology of the other two attractors is very intricate, and in particular their fundamental groups seem not to be finitely generated; we expect these attractors to be homotopic to complement of Cantor sets in the plane, which would make them homotopic to one another; they could even be homeomorphic.

some Riemannian metric, and that $T : M = N \times D \to M$ is written as a continuous skew-product $T(y, z) = (S(y), R(y, z))$ with shrinking fibers. Then T admits a unique physical measure μ_T, and the basin of attraction of μ_T has full volume.

If T is Lipschitz and has exponentially shrinking fibers (in particular, when T is uniformly hyperbolic), then μ_T has a polynomial decay of correlations of degree $(\alpha - 1)/2$ for $\mathscr{C}^{(2\alpha-2)\log}$ observables. Moreover, if $\alpha > 3/2$, then μ_T satisfies the central limit theorem for $\mathscr{C}^{(2\alpha-1)\log}$ observables.

The regularities needed on observables are quite weak (they include, in particular, all Hölder observables) but the assumption that T is a skew-product is very strong; it is a whole research project to consider the case where T is a general $\mathscr{C}^{1+\alpha\log}$ uniformly hyperbolic diffeomorphism onto its image, for example, with a 'solenoidal' attractor: one can quotient out by the stable foliation, obtaining a skew-product over a bundle, but only up to a conjugacy as regular as the foliation. Often, this conjugacy is *not* $\mathscr{C}^1$, and the regularity of the foliation's holonomy needs to be finely controlled to overcome this difficulty.

Remark 1.10. The skew-products of the above corollaries need not be diffeomorphisms, and can have intricate attractors: Figure 1 shows some examples with $Y = \mathbb{T}$ and $\Phi = [0, 1]$, and with S an expanding map of the circle.

1.4. *Physical versus SRB.* We close this gallery of examples by stressing the difference between physical measures and Sinai–Ruelle–Bowen (SRB) measures, where we use the distinction advocated by Young [**You02**] (see §2.3 and, in particular, Definition 2.10).

Theorem A combines with a result of Campbell and Quas [**CQ01**] (where they say 'SRB' for 'physical') to yield the following corollary.

COROLLARY 1.11. *There is a $\mathscr{C}^1$ diffeomorphism onto its image $T : U \to U$, where $U \subset \mathbb{R}^3$ is open and bounded, having a compact uniformly hyperbolic attractor $\Lambda = \bigcap_{k \geq 0} T^k(U)$ with unstable dimension one, supporting exactly one physical measure μ_T, but supporting no SRB measure.*

The proof, detailed in §7, can be summed up easily as follows. We construct T as a uniformly hyperbolic skew product (S, R) on $\mathbb{T} \times D^2$; [**CQ01**] shows that taking S a generic $\mathscr{C}^1$ expanding map, it has a unique physical measure $\check{\mu}$, but no absolutely continuous measure. The lift μ of $\check{\mu}$ is a physical measure of T with full basin of attraction, but an SRB measure would project to an absolutely continuous measure of S and thus does not exist. In other words, [**CQ01**] already provides many examples of the above kind, but somewhat degenerate as the stable dimension vanishes; the present work only serves to add some stable dimensions.

Remark 1.12. When $T : M \to M$ is a C^2 diffeomorphism with a hyperbolic attractor $\Lambda = \bigcap_{n_i n \mathbb{N}} T^n(U)$, where U is an open set with $T(\bar{U}) \subset U$, and when in addition there exists a compact $\mathscr{C}^1$ submanifold $Y \in U$ transversal to the stable foliation and intersecting each stable leaf at exactly one point, one can identify Y with the space of leaves and gets a factor $S : Y \to Y$ of T with a nice section $Y \hookrightarrow U$ (e.g. if the unstable dimension is 1 and stable leaves on U are relatively compact, such a Y is easily constructed). One can easily endow Y with a (not necessarily Riemannian) metric that makes S expanding. On the one hand, it follows from [**Klo17**, Theorem 5.8] that S satisfies UE($\boldsymbol{\omega}_*$; CLT, $\boldsymbol{\omega}_*$); on the other hand, the geometric potential $\varphi_T = -\log J^u T$ (where J^u is the determinant of the restriction of DT to the stable distribution) is $\mathscr{C}^1$, in particular Hölder, and one can thus use the present results to construct a Hölder potential $\hat{\varphi}_T$ cohomologous to φ_T and constant on stable leaves. This potential descends on Y into a Hölder potential, for which S has a unique equilibrium state. It follows that T has a unique equilibrium state μ_T for the geometric potential, and that μ_T has the expected statistical properties (CLT, exponential decay of correlations). By the work of Ledrappier [**Led84**], we know that such an equilibrium state is an SRB measure. Moreover, it is a physical measure thanks to the absolute continuity of holonomy of the stable foliation (see, for example, [**AV09**]). Without relying on Markov partitions, we recover in this case the classical result that there is a measure μ_T that is the unique physical measure, the unique SRB measure and the unique equilibrium state of the geometric potential; and that μ_T has good statistical properties.

1.5. *Beyond Lipschitz maps with uniformly shrinking fibers.* The present work can be developed in several directions; for example, one could apply similar ideas for flows. Section 3 only uses an averaged shrinking, and could thus be applied to the examples introduced by Diíaz, Horita, Rios and Sambarino [**DHRS09**] and further studied by Leplaideur, Oliveira and Rios [**LOR11**] and Ramos and Siqueira [**RS17**]. These examples, which are at the frontier between hyperbolic and robustly non-hyperbolic dynamics, are indeed extensions of uniformly expanding maps on a Cantor subset of $\mathbb{R}$, with only some exceptional fibers not being contracted. The ideas of the other sections might be applicable to such examples as well.

As mentioned above, in some interesting cases the map S is not continuous; see, for example, [**Gal18**, **GNP18**]. We expect most of the ideas used to prove Theorems B–D to be adaptable to such a setting, up to devising suitable functional spaces to work on (we do not claim that such an adjustment should always be straightforward); this should make it possible to consider more general invariant measures than the lift of the absolutely

continuous S-invariant measure. In particular, using disintegration with respect to π in its full generality should be useful.

Note also that the ideas presented here can be used without assuming compact fibers, if one has *contraction* properties instead of shrinking (e.g. $d(Tx, Tx') \leq \lambda d(x, x')$ for some $\lambda \in (0, 1)$ and all x, x' in the same fiber, or milder contraction using decay functions as in [**Klo17**]). One would need to assume some moment condition on the lifted measures μ, in order to use a Wasserstein distance (and its modification $W^{\breve{\mu}}$). In the examples I could think of, it would be possible to actually restrict to a stable compact subset of the space though, so that we do not pursue this direction.

1.6. *Organization of the paper.*   In §2 we introduce a number of tools and definitions, many of them classical. Given the variety of properties considered in our main theorem, this section is rather long for a preliminary one. Each subsequent section starts by pointing to the relevant subsections of §2, so that it can be mostly skipped and used as reference.

In §3 we prove existence and uniqueness of the T-invariant lift of an S-invariant measure and study convergences to it under iteration of T (this covers the first part of Theorem A, and Theorem B). Section 4 concludes the proof of Theorem A by considering each preserved property. In this part we consider the more general case of fibers shrunk on average with respect to an S-invariant measure, while all the subsequent sections assume that fibers are all (uniformly) shrinking.

In §5 we consider equilibrium states and establish a correspondence between potentials and observables on X and Y, by adding coboundaries and using the projection. In particular, Theorem C is proved.

Section 6 is devoted to the decay of correlations, and proves Theorem D. To this end, we use another correspondence between observables on X and Y, using disintegration; we prove that disintegration preserves some regularity properties of observables (Theorem 6.3).

Finally, in §7, we explain how to deduce Corollaries 1.7–1.11 from the main theorems and the literature.

2. *Preliminaries*

This section sets up notation and states some results for later use.

Let X be a compact metric space and $T : X \rightarrow X$ be a map (all maps are assumed to be Borel-measurable), admitting a factor $S : Y \rightarrow Y$, that is, Y is a compact metric space, S is a map, and there is a continuous onto map $\pi : X \rightarrow Y$ such that $S\pi = \pi T$ (we denote composition of maps by either juxtaposition or using the usual symbol $\circ$). The sets $\pi^{-1}(y)$ are called *fibers*. We denote by $d(\cdot, \cdot)$ both metrics on X and on Y, the context preventing any ambiguity. Note that to state the more general results, we do not require S, T to be continuous unless specified; in contrast, the continuity of π is crucial in many arguments.

We denote by $\mathscr{B}_X$, $\mathscr{B}_Y$ the Borel σ-algebras of X and Y, with respect to which all measurability conditions are considered unless otherwise specified. Let $\mathcal{P}(X)$, $\mathcal{P}(Y)$ be the sets of probability measures of X, Y and $\mathcal{P}_T(X)$ the set of T-invariant probability measures (similarly, $\mathcal{P}_S(Y)$ is the set of S-invariant probability measures). We denote

by $\int f \, d\mu$, $\int f(x) \, d\mu(x)$ or $\mu(f)$ the integral of an integrable or positive function f with respect to a measure μ.

In order to simplify a few arguments, we always assume (up to changing the metrics by a constant, thus not altering the statements of the theorems in the introduction) that diam X, diam $Y \leq 1$.

Constants denoted by C are positive and can vary from line to line, and we write $a(n) \lesssim b(n)$ to express that, for some $C > 0$ and all $n \in \mathbb{N}$, $a(n) \leq Cb(n)$.

2.1. *Moduli of continuity.* By a *modulus of continuity* we mean a continuous, increasing, concave function $\omega : [0, \infty) \to [0, \infty)$ mapping 0 to 0. We may only define a modulus near 0, and the understanding is that it is extended to the half line; since we shall only be concerned with compact spaces, the specifics of the extension are irrelevant.

A function $f : X \to \mathbb{R}$ is said to have modulus of continuity ω when

$$|f(x) - f(x')| \leq \omega(d(x, x')) \quad \text{for all } x, x' \in X.$$

Every continuous function on a compact metric space is uniformly continuous, hence has a modulus of continuity: concavity of the modulus can be ensured by taking the convex hull of $\{(\eta, \varepsilon) \in [0, \infty)^2 \mid \exists x, x' \in X, d(x, x') \leq \eta, |f(x) - f(x')| \geq \varepsilon\}$.

A function is said to be ω-continuous if there is a constant $C > 0$ such that it has $C\omega$ as a modulus of continuity; the infimum of all such C is denoted by $\mathrm{Hol}_\omega(f)$, and the set of ω-continuous functions $X \to \mathbb{R}$ is a Banach space ('generalized Hölder space') when endowed with the norm

$$\|f\|_\omega := \|f\|_\infty + \mathrm{Hol}_\omega(f)$$

(this claim follows from the corresponding classical claim for the Lipschitz modulus and from the observation that $\omega(d(\cdot, \cdot))$ defines a metric). An observation that will be used without warning is that whenever f has zero average with respect to an arbitrary probability measure, then it takes both non-positive and non-negative values; then $\|f\|_\infty \leq \omega(\mathrm{diam}\, X)\, \mathrm{Hol}_\omega(f)$ and thus $\|f\|_\omega \lesssim \mathrm{Hol}_\omega(f)$.

The most classical moduli of continuity are the Hölder ones, defined for $\alpha \in (0, 1]$ by $\omega_\alpha(r) = r^\alpha$ (so that ω_α-continuous means α-Hölder). We shall have use for a family of more lenient moduli.

Definition 2.1. For each $\alpha \in (0, \infty)$, we denote by $\omega_{\alpha \log}$ the modulus of continuity such that, on $(0, 1]$,

$$\omega_{\alpha \log}(r) = \frac{1}{(\log (r_\alpha/r))^\alpha}$$

where $r_\alpha > 1$ is chosen large enough to ensure monotony and concavity on $(0, 1]$, and $\omega_{\alpha \log}$ is constant for $r \geq 1$.

A $\omega_{\alpha \log}$-continuous function is also said to be α log-Hölder; a function is said to be log-Hölder if it is α log-Hölder for some $\alpha > 0$.

To simplify notation, we write Hol_α instead of $\mathrm{Hol}_{\omega_\alpha}$ and $\mathrm{Hol}_{\alpha \log}$ instead of $\mathrm{Hol}_{\omega_{\alpha \log}}$.

Example 2.2. When $X = \{0, 1\}^{\mathbb{N}}$ with the metric $d(x, x') = 2^{-i(x,x')}$, where $i(x, x')$ is the first index where $x_i \neq x'_i$, a function f is Hölder-continuous when the maximal influence

of the ith component decays exponentially fast, while f is α log-Hölder when the maximal influence of the ith component decays like $i^{-\alpha}$.

Let us show that the modulus $\omega_{\alpha\log}$ being very concave, it is only mildly affected by pre-composition by a high-order iterate of a Lipschitz map.

PROPOSITION 2.3. *For all $\alpha > 0$ and all $L \geq 1$, there exists $D > 0$ such that, for all $n \in \mathbb{N}$ and all $r \in [0, 1]$,*

$$\omega_{\alpha\log}(L^n r) \leq D n^\alpha \omega_{\alpha\log}(r).$$

In particular, if $T : X \to X$ and $\sigma : Y \to X$ are Lipschitz and $f \in \mathrm{Hol}_{\alpha\log}(X)$, then $\mathrm{Hol}_{\alpha\log}(fT^n\sigma) \lesssim n^\alpha \, \mathrm{Hol}_{\alpha\log}(f)$.

Proof. When $L^n r \leq 1$, we have

$$\frac{\omega_{\alpha\log}(L^n r)}{\omega_{\alpha\log}(r)} = \left(1 - \frac{\log L^n}{\log(r_\alpha/r)}\right)^{-\alpha}$$

which is bounded independently of n since $L^n \leq 1/r < r_\alpha/r$; while when $L^n r \geq 1$, we have

$$\frac{\omega_{\alpha\log}(L^n r)}{\omega_{\alpha\log}(r)} = \frac{\omega_{\alpha\log}(1)}{\omega_{\alpha\log}(r)} \lesssim \left(\log \frac{r_\alpha}{r}\right)^\alpha \lesssim (n \log L)^\alpha. \qquad \square$$

2.2. *Sections, disintegration.* The map π can be used to push measures forward: given $\mu \in \mathcal{P}(X)$, $\pi_* \mu : A \mapsto \mu(\pi^{-1}A)$ is a probability measure on Y. Moreover, a T-invariant measure is pushed to an S-invariant measure: for all $f : Y \to \mathbb{R}$,

$$\int f \circ S \, d(\pi_* \mu) = \int f \circ S\pi \, d\mu = \int f \circ \pi T \, d\mu = \int f \circ \pi \, d\mu = \int f \, d(\pi_* \mu).$$

Our first goal, in §3, will be to lift an invariant measure $\check{\mu}$ of S into an invariant measure μ of T, where 'lifting' entails that $\pi_* \mu = \check{\mu}$.

Recall a notion borrowed from the theory of fiber bundles.

Definition 2.4. A *section* of π is a measurable map $\sigma : Y \to X$ such that $\pi\sigma = \mathrm{Id}_Y$.

In other words, $\sigma(y) \in \pi^{-1}(y)$ for all $y \in Y$, that is, σ picks a point in the fiber of its argument. The map $\sigma\pi : X \to X$ then sends each point to the point in its own fiber picked by σ. Note that there is no assumption relating the section with the dynamics. Requiring σ to be measurable is very mild, and in many cases we will require it to be continuous, or even Lipschitz.

PROPOSITION 2.5. (Measurable selection theorem [**KRN65**]) *There exists a section $\sigma :$ $Y \to X$. As a consequence, π_* is onto $\mathcal{P}(Y)$.*

(That π_* is onto follows from the observation that, for all $\nu \in \mathcal{P}(Y)$, $\pi_*(\sigma_* \nu) = \nu$.)

We shall use in a central way the disintegration along a map. We state here the disintegration theorem for π, but it only needs measurability and can be used with other maps, such as S.

PROPOSITION 2.6. *(Disintegration theorem [**Roh52**, **Sim12**]) Let $\mu \in \mathcal{P}(X)$ and $\check{\mu} = \pi_*\mu \in \mathcal{P}(Y)$ be (not necessarily invariant) probability measures. There exist a family $(\xi_y)_{y \in Y}$ of probability measures on X such that $y \mapsto \xi_y$ is Borel-measurable, ξ_y is concentrated on $\pi^{-1}(y)$ for all $y \in Y$, and*

$$\int \xi_y(f) \, d\check{\mu} = \int f \, d\mu \quad \text{for all } f \in L^1(\mu).$$

Moreover, $(\xi_y)_{y \in Y}$ is uniquely defined by these properties up to a $\check{\mu}$-negligible set.

For example, if $\mu = \sigma_*\check{\mu}$ for some section σ, then $\xi_y = \delta_{\sigma(y)}$ for $\check{\mu}$-almost all $y \in Y$.

Given a function $f : X \to \mathbb{R}$ in $L^1(\mu)$, we can define a function in $L^1(\check{\mu})$ by $\xi(f) : y \mapsto \xi_y(f)$ (and then $\check{\mu}(\xi(f)) = \mu(f)$ and $\xi(f)(y)$ only depends on the values of f on $\pi^{-1}(y)$). In §6 we shall study how much regularity $\xi(f)$ retains from the regularity of f; but we have to keep in mind that even when f is continuous, $\xi(f)$ is unambiguously defined only modulo a $\check{\mu}$-negligible set.

Definition 2.7. We say that a measurable function $u : Y \to \mathbb{R}$ *has a continuous version* if there exist a continuous $\bar{u} : Y \to \mathbb{R}$ which is equal to u at $\check{\mu}$-almost every point. If supp $\check{\mu} = Y$, then $\bar{u}$ is unique.

We say that ξ *preserves continuity* if, for all continuous $f : X \to \mathbb{R}$, $\xi(f) : y \mapsto \xi_y(f)$ has a continuous version.

If ω, $\check{\omega}$ are two moduli of continuity, we say that ξ is $(\omega, \check{\omega})$-*bounded* if, for all ω-continuous f, $\xi(f)$ is $\check{\omega}$-continuous and, moreover, the linear map $f \mapsto \xi(f)$ is a continuous operator $\mathrm{Hol}_\omega(X) \to \mathrm{Hol}_{\check{\omega}}(Y)$. If $\check{\omega} = \omega$, then we simply say that ξ is ω-*bounded*.

2.3. *Physicality, observability, SRB.* Assume now that X and Y are equipped with measures λ_X and λ_Y (which *a priori* need not have any particular relation with T, S but will serve as reference measure); for example, X, Y are manifolds equipped with volume forms, or are domains of $\mathbb{R}^m$, $\mathbb{R}^{\check{m}}$ equipped with the Lebesgue measure.

Definition 2.8. The *basin* of a T-invariant measure μ is defined as

$$\mathrm{Ba}(\mu) = \left\{ x \in X \;\middle|\; \frac{1}{n} \sum_{k=0}^{n-1} \delta_{T^k x} \to \mu \right\}$$

(where $\to$ denotes weak-$*$ convergence and δ_x is the Dirac mass at x).

The invariant measure μ is said to be *physical* if its basin has a positive volume:

$$\lambda_X(\mathrm{Ba}(\mu)) > 0.$$

Often, physical measures are said to be those that can be seen in practice, given they drive the behavior of a positive proportion of the points. However, note that in some cases Guihéneuf [**Gui15**] has shown that non-physical measures could be actually observed.

Physical measures do not always exist, and a more general class of measures was proposed in [**CE11**].

Definition 2.9. Given $x \in X$, denote by $\mathrm{p}\omega(x) \subset \mathcal{P}(X)$ the set of cluster points of the sequence $(\sum_{k=0}^{n-1} \delta_{T^k x})_n$. Observe that $\mathrm{p}\omega(x) \subset \mathcal{P}_T(X)$.

Given $\mu \in \mathcal{P}(X)$ and $\varepsilon > 0$, the ε-basin of μ is

$$\mathrm{Ba}_\varepsilon(\mu) = \{x \in X \mid \mathrm{W}(\mathrm{p}\omega(x), \mu) < \varepsilon\}.$$

An invariant measure $\mu \in \mathcal{P}_T(X)$ is said to be *observable* when, for all $\varepsilon > 0$, its ε-basin has positive volume. (Note that while choosing W as metric has an influence on the ε-basins, any metric inducing the weak-$*$ topology yields the same notion of observability.)

It is important to make the distinction between the above notion of *physical* measures and the related but distinct notion of *SRB* measure (beware that some authors use 'SRB' instead of 'physical'). To introduce SRB measure we need to recall some preliminary definitions.

Let U be an open bounded set of a manifold M; a diffeomorphism onto its image $T : U \to U$ has a stable subset $\Lambda := \bigcap_n T^n(U)$, called its *attractor*, on which T induces a homeomorphism. Assuming that $\overline{T(U)} \subset U$ (equivalently, that the closure *in U* of $T(U)$ is compact), the attractor is compact: we have $T^{n+1}(U) = T^n(\overline{T(U)}) \subset T^n(U) \subset \overline{T^n(U)}$, so that $\Lambda = \bigcap_n \overline{T^n(U)}$ is a decreasing intersection of compact sets.

One says that Λ is a *strongly partially hyperbolic* attractor whenever there are continuous sub-bundles E^u, E^s of $T_\Lambda M$ (the restriction of the tangent bundle TM to Λ) of respective dimension d_u, d_s, and there are numbers C and $\lambda_+ > \lambda_- > 0$ such that, for all $x \in \Lambda$ and all $n \in \mathbb{N}$,

$$\|DT_x^n(u)\| \le C\lambda_-^n \|u\| \quad \text{for all } u \in E_x^s \quad \text{and} \quad \|DT_x^n(u)\| \ge C^{-1}\lambda_+^n \|u\| \quad \text{for all } u \in E_x^u.$$

If, moreover, $\lambda_- < 1 < \lambda_+$, one says that Λ is a *uniformly hyperbolic* attractor.

Assuming Λ is a uniformly hyperbolic attractor, we get an invariant 'stable' foliation W^s of U, whose leaf W_x^s through $x \in \Lambda$ is the set of points whose orbit converges to the orbit of x; and an invariant 'unstable' lamination W^u of Λ, whose leaf W_x^u through x is the set of points whose backward orbit converges to the backward orbit of x. The leaves of W^s and W^u are $\mathscr{C}^1$, and they are continuous but not necessarily transversely $\mathscr{C}^1$. Moreover, $T_x W^s = E_x^s$ and $T_x W^u = E_x^u$.

Locally, we can then write the attractor Λ as a product (one factor corresponding to the unstable direction, the other to the stable direction). Given an invariant measure $\mu \in \mathcal{P}_T(U)$ (which must be supported on Λ), we can disintegrate the restriction of μ to a small open set of Λ with respect to the (local) projection π^s in the stable direction, obtaining one the one hand a family of measures $(\mu_L)_L$ supported on each local unstable leaf L, and on the other hand a projected measure $\nu = \pi_*^s \mu$.

Definition 2.10. We say that μ is an *SRB* measure when in this local disintegration, μ_L is absolutely continuous with respect to the Riemannian volume induced on L for ν-almost all L.

2.4. *Wasserstein metric and its vertical version.* We will make use of the Wasserstein metric to metrize the weak-$*$ topology on $\mathcal{P}(X)$. It is defined for $\mu_0, \mu_1 \in \mathcal{P}(X)$ by

$$\mathrm{W}(\mu_0, \mu_1) = \inf_{\gamma \in \Gamma(\mu_0, \mu_1)} \int d(x_0, x_1) \, d\gamma(x_0, x_1) = \sup_{f \text{ 1-Lipschitz}} |\mu_0(f) - \mu_1(f)|$$

where $\Gamma(\mu_0, \mu_1)$ is the set of *couplings* or *transport plans* between μ_0 and μ_1, that is, the set of $\gamma \in \mathcal{P}(X \times X)$ such that $\gamma(A \times X) = \mu_0(A)$ and $\gamma(X \times A) = \mu_1(A)$ for all Borel $A \subset X$. The equality between the two definitions (by transport plans or by duality with Lipschitz functions) is not trivial, and is called Kantorovich duality. The infimum is reached, any transport plan realizing it is said to be optimal, and the set of optimal plans is compact in the weak-$*$ topology (see, for example, [**Vil09**]).

To prove Theorem 3.1 below we introduce a variation of the Wasserstein metric where mass is only allowed to move along fibers. This constraint implies that we need to consider pairs of measure with the same projection. Similar ideas have been developed in [**GP10**, **AGP14**, **GL15**], but in somewhat restricted settings, without taking full advantage of the dual formulations of the Wasserstein metric and of the disintegration theorem.

For each $\check{\mu} \in \mathcal{P}(Y)$, by continuity of π the fiber $\pi_*^{-1}(\check{\mu})$ is a closed subset of $\mathcal{P}(X)$ in the weak-$*$ topology, thus is compact. Set $\Delta_\pi = \{(x_0, x_1) \in X \times X \mid \pi(x_0) = \pi(x_1)\}$. Given any $\mu_0, \mu_1 \in \pi_*^{-1}(\check{\mu})$, we denote by $\Gamma_\pi(\mu_0, \mu_1)$ the set of $\gamma \in \Gamma(\mu_0, \mu_1)$ which are concentrated on Δ_π. We define

$$\mathrm{W}^{\check{\mu}}(\mu_0, \mu_1) = \inf_{\gamma \in \Gamma_\pi(\mu_0, \mu_1)} \int d(x_0, x_1) \, d\gamma(x_0, x_1).$$

We will see in a moment that this is a finite number, but it is already clear that $\mathrm{W} \leq \mathrm{W}^{\check{\mu}}$.

LEMMA 2.11. *As soon as $\pi_*\mu_0 = \pi_*\mu_1$, the set $\Gamma_\pi(\mu_0, \mu_1)$ is non-empty, and as a consequence $\mathrm{W}^{\check{\mu}}(\mu_0, \mu_1) < \infty$. More precisely, if $(\xi_y)_{y \in Y}$ and $(\zeta_y)_{y \in Y}$ are the disintegrations of μ_0 and μ_1 with respect to π, then*

$$\mathrm{W}^{\check{\mu}}(\mu_0, \mu_1) = \int \mathrm{W}(\xi_y, \zeta_y) \, d\check{\mu}(y).$$

Proof. Choose measurably $\eta_y \in \Gamma(\xi_y, \zeta_y)$ for each $y \in Y$ (e.g. $\eta_y = \xi_y \otimes \zeta_y$), and let $\gamma = \int \eta_y \, d\check{\mu} \in \mathcal{P}(X \times X)$, that is, for all Borel $A, B \in X$, $\gamma(A \times B) = \int \eta_y(A \times B) \, d\check{\mu}$. Then $\gamma(A \times X) = \int \xi_y(A) \, d\check{\mu} = \mu_0(A)$ and $\gamma(X \times A) = \int \zeta_y(A) \, d\check{\mu} = \mu_1(A)$ so that $\gamma \in \Gamma(\mu_0, \mu_1)$. Moreover, since η_y projects to two measures supported on $\pi^{-1}(y)$, it is supported on $\pi^{-1}(y) \times \pi^{-1}(y) \subset \Delta_\pi$. It follows that γ is concentrated on Δ_π and $\gamma \in \Gamma_\pi(\mu_0, \mu_1)$. Any $\gamma \in \Gamma_\pi(\mu_0, \mu_1)$ is of the form $\int \eta_y \, d\check{\mu}$, $(\eta_y)_{y \in Y}$ being the disintegration of γ with respect to the map induced by π from Δ_π to Y. Since $\int d(x, x') \, d\gamma(x, x') = \iint d(x, x') \, d\eta_y(x, x') \, d\check{\mu}(y) \geq \int \mathrm{W}(\xi_y, \zeta_y) \, d\check{\mu}(y)$, taking an infimum, we get $\mathrm{W}^{\check{\mu}}(\mu_0, \mu_1) \geq \int \mathrm{W}(\xi_y, \zeta_y) \, d\check{\mu}$.

For each y, the set of optimal transport plans from ξ_y to ζ_y is compact, thus by the measurable selection theorem there is a measurable family $(\eta_y)_{y \in Y}$ such that, for $\check{\mu}$-almost all $y \in Y$, $\int d(x, x') \, d\eta_y(x, x') = \mathrm{W}(\xi_y, \zeta_y)$. It follows that $\mathrm{W}^{\check{\mu}}(\mu_0, \mu_1) \leq \int \mathrm{W}(\xi_y, \zeta_y) \, d\check{\mu}$. $\square$

PROPOSITION 2.12. *For all $\check{\mu} \in \mathcal{P}(Y)$, $\mathrm{W}^{\check{\mu}}$ is a complete metric on the set $\pi_*^{-1}(\check{\mu})$.*

Proof. The expression $\int \mathrm{W}(\xi_y, \zeta_y) \, d\check{\mu}(y)$ is the $L^1(\check{\mu})$ metric for functions from Y to the compact metric space $(\mathcal{P}(X), \mathrm{W})$. By Lemma 2.11, $\mathrm{W}^{\check{\mu}}$ is the restriction of this metric to the closed subset $\pi_*^{-1}(\check{\mu})$. The claim thus follows from the Riesz–Fischer theorem for functions with values in a complete metric space. $\square$

2.5. *Transfer operators, spectral gap and correlations.* Transfer operators are multifaceted objects that are both tools and objects of study. We will use a definition that needs to introduce a generalization of the notion of invariant measure, and then we shall describe other equivalent definitions.

2.5.1. *Quick introduction to transfer operators.* We consider S acting on Y since it is the level at which transfer operators will be most relevant.

Definition 2.13. A measure $\lambda_Y \in \mathcal{P}(Y)$ is said to be a *conformal* measure of S when $S_* \lambda_Y$ is absolutely continuous with respect to λ_Y.

Given a conformal measure λ_Y of S, one defines the *transfer operator*

$$\check{\mathscr{L}} = \check{\mathscr{L}}_{\lambda_Y} : L^1(\lambda_Y) \to L^1(\lambda_Y)$$

of S with respect to λ_Y by $S_*(f \, d\lambda_Y) = \check{\mathscr{L}}(f) \, d\lambda_Y$.

For example, the Lebesgue measure on the circle is a conformal measure for all local $\mathscr{C}^1$ diffeomorphisms (but not for a map that is constant on an interval). The transfer operator simply translates the action of S_* on the set of absolutely continuous measures (with respect to λ_Y) to the space of densities. In particular, finding an absolutely continuous invariant measure is equivalent to finding a non-negative, non-zero eigenfunction of $\check{\mathscr{L}}$ (the eigenvalue is then necessarily 1, since $\int \check{\mathscr{L}}(f) \, d\lambda_Y = \int f \, d\lambda_Y$).

Another classical way to say the same thing is to define $\check{\mathscr{L}}$ as the dual operator of the Koopman operator $f \mapsto f \circ S : L^\infty(\lambda_Y) \to L^\infty(\lambda_Y)$, that is, to characterize it by the property

$$\int f \cdot \check{\mathscr{L}}(g) \, d\lambda_Y = \int f \circ S \cdot g \, d\lambda_Y \quad \text{for all } f \in L^\infty(\lambda_Y), \text{ for all } g \in L^1(\lambda_Y). \quad (1)$$

Invariant measures $\check{\mu}$ are characterized by the property that their transfer operators have the property $\check{\mathscr{L}}_{\check{\mu}} \mathbf{1} = \mathbf{1}$ where $\mathbf{1}$ is the constant function with value one.

2.5.2. *Decay of correlations and spectral gap.* Transfer operators are precious tools to study the decay of correlations.

Definition 2.14. Given $\mu \in \mathcal{P}_T(X)$ and functions $u, v : X \to \mathbb{R}$ (called observables), *correlations* are defined (whenever it makes sense) as

$$C_\mu^n(u, v) = \left| \int u \circ T^n \cdot v \, d\mu - \int u \, d\mu \int v \, d\mu \right|$$

(and of course $C_\mu^n(f, g)$ with $f, g : Y \to \mathbb{R}$ implicitly involves the map S).

We say that $\mu \in \mathcal{P}_T(X)$ *has decay of correlations* $(b_n)_{n \in \mathbb{N}}$ *for ω-continuous observables* whenever, for all $f \in \mathrm{Hol}_\omega(X)$, all $g \in L^1(\mu)$ and all $n \in \mathbb{N}$,

$$C_\mu^n(f, g) \le b_n \, \mathrm{Hol}_\omega(f) \|g\|_{L^1(\mu)}.$$

The link between the transfer operator and decay of correlation is quite direct: assuming $g \in L^1(\check{\mu})$, $f \in L^\infty(\check{\mu})$ and adding a constant to f to ensure $\check{\mu}(f) = 0$, we obtain

$$C_{\check{\mu}}^n(g, f) = \left| \int g \circ S^n \cdot f \, d\check{\mu} \right| = \left| \int g \cdot \check{\mathscr{L}}_{\check{\mu}}^n(f) \, d\check{\mu} \right| \le \|g\|_{L^1(\check{\mu})} \|\check{\mathscr{L}}_{\check{\mu}}^n(f)\|_{L^\infty(\check{\mu})}$$

(other pairs of functional spaces can be considered, such as L^2 and L^2, or inverting the roles of L^1 and L^∞). One thus only has to prove decay of $\check{\mathscr{L}}_{\check{\mu}}^n$ for zero-average observables in some functional space to obtain a corresponding decay of correlations. A particularly nice case, both to find an absolutely continuous invariant probability (Acip) and to prove exponential decay of correlations for it, is when the transfer operator has a spectral gap.

Definition 2.15. Given a Banach space $\mathcal{B}$ of functions $Y \to \mathbb{R}$ whose norm $\|\cdot\|$ is not less than $\|\cdot\|_\infty$ (one could generalize to $\|\cdot\|_{L^1(\check{\mu})}$), one says that $\check{\mathscr{L}}$ has a *spectral gap* on $\mathcal{B}$ whenever:

- $\check{\mathscr{L}}$ preserves $\mathcal{B}$ and acts on it as a bounded operator;
- there is a positive function $h \in \mathcal{B}$ such that $\check{\mathscr{L}}h = h$, which without loss of generality can be assumed to satisfy $\lambda_Y(h) = 1$;
- there are numbers $C \geq 1$, $\delta \in (0, 1)$ such that, for all $f \in \mathcal{B}$ with $\lambda_Y(f) = 0$, $\|\check{\mathscr{L}}^n f\| \leq C(1 - \delta)^n \|f\|$.

Then $hd\lambda_Y$ is an S-invariant probability measure absolutely continuous with respect to λ_Y, satisfying exponential decay of correlations for observables in $\mathcal{B}$ (note that $\check{\mathscr{L}}$ and $\check{\mathscr{L}}_{hd\lambda_Y}$ are conjugated one to another).

We shall need some transfer operators to preserve some functional spaces in the following sense.

Definition 2.16. An operator $\mathscr{P} : L^1(\check{\mu}) \to L^1(\check{\mu})$ is said to *preserve* $\mathrm{Hol}_\omega(Y)$ if $\mathscr{P}(\mathrm{Hol}_\omega(Y)) \subset \mathrm{Hol}_\omega(Y)$ and, moreover, if it induces a bounded operator on $\mathrm{Hol}_\omega(Y)$. The operator $\mathscr{P}$ is said to be *iteratively bounded* with respect to ω if it preserves $\mathrm{Hol}_\omega(Y)$ and, moreover, if there exists $D \geq 0$ such that, for all $n \in \mathbb{N}$ and all $u \in \mathrm{Hol}_\omega(Y)$, $\mathrm{Hol}_\omega(\mathscr{P}^n u) \leq D \, \mathrm{Hol}_\omega(u)$.

For example, if $\mathscr{P}$ has a spectral gap on $\mathrm{Hol}_\omega(Y)$, then it is iteratively bounded with respect to ω (but the latter assumption is much milder than having a spectral gap).

2.5.3. *Disintegrations and transfer operators.* To close this subsection, we shall consider a slightly different point of view on transfer operators, which seems novel in this generality (although it is folklore in the case of Lebesgue measure) and will enable us to relate the transfer operators of T and S. We restrict to the case of a T-invariant measure μ and its S-invariant projection $\check{\mu} = \pi_*\mu$. The transfer operators of (T, μ) and $(S, \check{\mu})$ are denoted by $\mathscr{L}$ and $\check{\mathscr{L}}$.

We denote by $(\xi_y)_{y \in Y}$ the disintegration of μ with respect to the map π, which we recall is characterized by two properties: $\mu = \int \xi_y \, d\check{\mu}(y)$ and $\mathrm{supp}\, \xi_y \subset \pi^{-1}(y)$ for all $y \in Y$. The disintegration theorem can also be applied to T and μ, and yields an essentially unique measurable family of probability measures $(\eta_x)_{x \in X}$ characterized by $\mu = \int \eta_x \, d\mu(x)$ and $\eta_x(T^{-1}(x)) = 1$ for all $x \in X$ (here $T^{-1}(x)$ need not be closed, and while η_x is *concentrated* on $T^{-1}(x)$ its *support* could be larger).

To understand the meaning of this disintegration, one can say that each measure η_x collects the 'derivatives' of T with respect to the measure μ at the points of $T^{-1}(x)$. The clearest situation is when T is at most countable-to-one, in which case η_x is atomic and

the masses of the atoms can be taken as definition of this derivative. If T is one-to-one, then of course $\eta_x = \delta_{T^{-1}x}$.

We can use the disintegration to express the transfer operator.

PROPOSITION 2.17. *We have $\mathscr{L}g(x) = \eta_x(g)$ for all $g \in L^1(\mu)$ and μ-almost all $x \in X$.*

Proof. The proposed formula defines a bounded operator $\tilde{\mathscr{L}}(g)(x) = \eta_x(g)$ of $L^1(\mu)$ into itself, and to prove $\tilde{\mathscr{L}} = \mathscr{L}$ it suffices to check the defining property (1). Let $f \in L^\infty(\mu)$ and $g \in L^1(\mu)$; using the fact that $x = Tx'$ for η_x-almost all x', we get

$$\int f \cdot \tilde{\mathscr{L}}g \, d\mu = \int f(x) \int g(x') \, d\eta_x(x') \, d\mu(x)$$
$$= \iint f(Tx')g(x') \, d\eta_x(x') \, d\mu(x) = \int f \circ T \cdot g \, d\mu. \qquad \square$$

Very often, one works the other way around: the family $(\eta_x)_{x \in X}$ is given and used to define a transfer operator, which is in turn used to construct an invariant measure μ with the prescribed derivatives. Using the disintegration theorem makes transparent the fact that one can go both ways round in a consistent fashion.

Note that, using either definition of transfer operator, we easily get the classical property $\mathscr{L}(f \circ T \cdot g) = f\mathscr{L}(g)$ (where $g \in L^1(\mu)$, $f \in L^\infty(\mu)$). We have $T_*\eta_x = \delta_x$ since η_x is a probability measure supported on $T^{-1}(x)$.

The same applies to S, and we denote by $(\check{\eta}_y)_{y \in Y}$ the disintegration of $\check{\mu}$ with respect to S and by $\check{\mathscr{L}}$ its transfer operator. The same relations as above hold, in particular $\check{\mathscr{L}}u(y) = \check{\eta}_y(u)$. Now, our goal is to relate the transfer operators (or equivalently, the disintegrations) of T and S.

LEMMA 2.18. *For $\check{\mu}$-almost all $y \in Y$, all $u \in L^1(\check{\mu})$ and all $n \in \mathbb{N}$, we have*

$$\pi_*(\textstyle\int \eta_x \, d\xi_y(x)) = \check{\eta}_y, \quad \check{\mathscr{L}}^n u(y) = \xi_y(\mathscr{L}^n(u \circ \pi)).$$

Proof. To prove the first claim, it suffices to check the two defining properties of $(\check{\eta}_y)_{y \in Y}$. First, the measure $\int \eta_x \, d\xi_y(x)$ is concentrated on $T^{-1}(\pi^{-1}(y)) = \{x \in X \mid \pi T(x) = y\} = \pi^{-1}(S^{-1}(y))$ so that its push-forward by π is concentrated on $S^{-1}(y)$. Second, for all $u \in L^1(\check{\mu})$, we have

$$\int \pi_*(\textstyle\int \eta_x \, d\xi_y(x))(u) \, d\check{\mu}(y) = \iint \eta_x(u \circ \pi) \, d\xi_y(x) \, d\check{\mu}(y)$$
$$= \int \eta_x(u \circ \pi) \, d\mu(x) = \int u \circ \pi \, d\mu = \int u \, d\check{\mu}.$$

We prove the second claim using the duality definition of $\check{\mathscr{L}}$:

$$\int \xi_y(\mathscr{L}^n(u\pi)) \cdot v(y) \, d\check{\mu}(y) = \int \xi_y(\mathscr{L}^n(u\pi) \cdot v\pi) \, d\check{\mu}(y)$$
$$= \int \mathscr{L}^n(u\pi) \cdot v\pi \, d\mu = \int u\pi \cdot vS^n\pi \, d\mu$$
$$= \int u \cdot vS^n \, d\check{\mu} = \int \check{\mathscr{L}}^n(u) \cdot v \, d\check{\mu}. \qquad \square$$

2.6. *Statistical properties.*   Let us define precisely the three statistical properties we shall focus on (as will be clear from the proofs, we could consider any statistical theorem insensitive to adding a bounded error term to $\sum_{k=1}^{n} f \circ T^k$).

Definition 2.19. Let $\mathcal{T} \in \{\text{LIL}, \text{CLT}, \text{ASIP}\}$; we shall say that an invariant measure $\mu \in \mathcal{P}_T(X)$ *satisfies $\mathcal{T}$ for all ω-continuous observables* if, for each $f \in \text{Hol}_\omega(X)$, there is $\sigma_f \geq 0$ (meant as a standard deviation, not to be confused with a section) such that, whenever $\sigma_f > 0$, the following statements hold.

When $\mathcal{T} = \text{LIL}$: for μ-almost every $x \in X$,

$$\limsup_{n \to \infty} \frac{\sum_{k=1}^{n} f \circ T^k(x) - n\mu(f)}{\sqrt{2n \log \log n}} = \sigma_f.$$

When $\mathcal{T} = \text{CLT}$: denoting by G_{m,σ^2} the cumulative distribution function of the normal law of mean m and variance σ^2, for all $r \in \mathbb{R}$,

$$\mu\left\{ x \in X : \frac{1}{\sqrt{n}} \sum_{k=1}^{n} f \circ T^k(x) \leq r \right\} \to G_{\mu(f),\sigma_f^2}(r).$$

When $\mathcal{T} = \text{ASIP}$: for some $\lambda \in (0, \frac{1}{2}]$, there exist a probabilistic space Ω and two real-valued processes defined on Ω:
- $(A_k)_{k \in \mathbb{N}}$ with the same law as $(f \circ T^k(Z))_{k \in \mathbb{N}}$, where Z is a random variable with law μ,
- $(B_k)_{k \in \mathbb{N}}$, a sequence of independent Gaussian random variables of mean $\mu(f)$ and variance σ_f^2,

such that almost surely $|\sum_{k=1}^{n} A_k - \sum_{k=1}^{n} B_k| = o(n^\lambda)$.

   We will usually keep the data σ_f, λ implicit but they are part of the statistical theorem, and when we state that a $\text{UE}(\omega_p'[\rho]; \mathcal{T}, \omega_o')$ property for S implies a $\text{UE}(\omega_p[C\rho]; \mathcal{T}, \omega_o)$ property for T, we always implicitly mean that the equilibrium state μ_φ of a potential φ and $\check{\mu} := \pi_*(\mu_\varphi)$ (which will be an equilibrium state for a potential $\check{\varphi}$) satisfy $\mathcal{T}$ *with the same parameters* under a correspondence $f \mapsto \check{f}$ (made explicit in §5), that is, $\sigma_f = \sigma_{\check{f}}$ in the respective statements for μ and $\check{\mu}$ (and, in the case of the ASIP, additionally λ is the same in both statements).

3. *Lifting invariant measures*
This section uses the material of §§2.1, 2.4, and 2.5 additionally for the proof of Theorem B.

3.1. *Existence and uniqueness.*   It is proved in [**APPV09**] that under a shrinking hypothesis each $\check{\mu} \in \mathcal{P}_S(Y)$ has a lift $\mu \in \mathcal{P}_T(X) \cap (\pi_*)^{-1}(\check{\mu})$. Uniqueness seems to be known only under some ergodicity hypotheses (see [**BM17**, Remark 2(b)]). Our first result gives uniqueness in general and a quantified convergence, and generalizes to the case of fibers shrunk *on average* (Definition 1.2).

THEOREM 3.1. (Lifting theorem) *Let $\check{\mu}$ be an S-invariant probability measure, and assume that the fibers of the extension T are shrunk on average with respect to $\check{\mu}$,*

with shrinking sequence $(\bar{a}_n)_n$. *Then there is a unique* $\mu \in \mathcal{P}_T(X)$ *such that* $\pi_*\mu = \check{\mu}$. *Moreover, for all* $\nu \in \mathcal{P}(X)$ *(not necessarily invariant) such that* $\pi_*\nu = \check{\mu}$ *and all* $n \in \mathbb{N}$, *we have* $W(T_*^n\nu, \mu) \le \bar{a}_n$; *in particular,* $T_*^n\nu \to \mu$ *in the weak-$*$ topology.*

If T has $(a_n)_n$-*shrinking fibers then the above statement holds for all* $\check{\mu} \in \mathcal{P}_S(Y)$ *with* $\bar{a}_n = a_n$.

Proof. For all $\mu_0, \mu_1 \in \pi_*^{-1}(\check{\mu})$, and all $\gamma \in \Gamma_\pi(\mu_0, \mu_1)$, denoting by (T^n, T^n) the map from $X \times X$ to itself sending (x_0, x_1) to $(T^n x_0, T^n x_1)$, we have

$$(T^n, T^n)_*\gamma \in \Gamma_\pi(T_*^n\mu_0, T_*^n\mu_1).$$

Since γ is supported on Δ_π, for γ-almost all (x_0, x_1), we have $\pi(x_0) = \pi(x_1)$; using this, that the first marginal of γ is μ_0 and that $\pi_*\mu_0 = \check{\mu}$, we have

$$\int d(x_0, x_1)\, d((T^n, T^n)_*\gamma)(x_0, x_1) = \int d(T^n x_0, T^n x_1)\, d\gamma(x_0, x_1)$$

$$\le \int \mathrm{diam}(T^n(\pi^{-1}(\pi(x_0))))\, d\gamma(x_0, x_1)$$

$$\le \int \mathrm{diam}(T^n(\pi^{-1}(y)))\, d\check{\mu}(y),$$

$$W^{\check{\mu}}(T_*^n\mu_0, T_*^n\mu_1) \le \bar{a}_n. \tag{2}$$

Applying this to any $\nu \in \pi_*^{-1}(\check{\mu})$ and to $T_*^{m-n}\nu$, we get $W(T_*^n\nu, T_*^m\nu) \le \bar{a}_n$ for all $n < m \in \mathbb{N}$, that is, $(T_*^n\nu)_n$ is a Cauchy sequence with respect to $W^{\check{\mu}}$. By Proposition 2.12, it has a limit $\mu \in \pi_*^{-1}(\check{\mu})$ in the metric $W^{\check{\mu}}$, which is also a weak-$*$ limit since $W \le W^{\check{\mu}}$.

Since T is uniformly continuous along fibers, we have, for any $\gamma \in \Gamma_\pi(T_*^n\nu, \mu)$,

$$\int d(x, x')\, d((T, T)_*\gamma)(x, x') \le \int \bar{\omega}_T(d(x, x'))\, d\gamma(x, x')$$

$$\le \bar{\omega}_T\left(\int d(x, x')\, d\gamma(x, x')\right).$$

Taking an infimum, we get $W(T_*^{n+1}\nu, T_*\mu) \le \bar{\omega}_T(W^{\check{\mu}}(T_*^n\nu, \mu))$; the left-hand side converges to $W(\mu, T_*\mu)$ while the right-hand side goes to 0, so that μ is T-invariant.

Reapplying (2) to ν and μ, we get the desired convergence in the Wasserstein metric. □

Remark 3.2. The existence part in [**APPV09**, Corollary 6.2] might at first seem more general in the case of shrinking fibers, as no continuity of the map T (denoted there by F) is explicitly assumed while we assume uniform continuity *along the fibers* in Definition 1.1. However, full continuity is implicitly used in the proof of [**APPV09**, Corollary 6.2]: to obtain that μ_F is invariant, [**APPV09**, Lemma 6.1] is applied to the observable $\psi \circ F$, implicitly assuming it to be continuous.

A related issue found elsewhere in the literature is to construct $\mu \in (\pi_*)^{-1}(\check{\mu})$ as the limit of push-forward measures $T_*^n\nu$, and deducing invariance of μ by writing

$$T_*\mu = T_*\left(\lim_{n\to\infty} T_*^n\nu\right) = \lim_{n\to\infty} T_*^{n+1}\nu = \mu.$$

This is perfectly fine when T is continuous, but without this assumption the second equality may fail.

COROLLARY 3.3. *If T is an extension of S with shrinking fibers, then the map π_* : $\mathcal{P}(X) \to \mathcal{P}(Y)$ induces a homeomorphism from $\mathcal{P}_T(X)$ to $\mathcal{P}_S(Y)$.*

Proof. By Theorem 3.1, π_* induces a bijection $\mathcal{P}_T(X) \to \mathcal{P}_S(Y)$. Since π_* is continuous and $\mathcal{P}_T(X)$ is compact, this induced map is a homeomorphism.   □

We shall denote by $\pi^* : \mathcal{P}_S(Y) \to \mathcal{P}_T(X)$ the inverse map of this homeomorphism (this notation has a differential-geometric flavor: an index star denotes push-forward, while an exponent star denotes pull-back).

3.2. *Stable leaves of invariant measures.*   Imagine that one wishes to draw a random point $x \in X$ whose law is close to μ. Using Theorem 3.1, one could draw a random point $y \in Y$ with law $\check{\mu}$, choose in any way (random or deterministic) an inverse image $x_0 \in \pi^{-1}(y)$, and take $x = T^n(x_0)$ for some large n. However, one may not be able to draw y with precisely the law $\check{\mu}$. One would hopefully be able to draw y with a law very close to $\check{\mu}$, and still get that the law of x is close to μ.

In other words, one asks for conditions on a probability measure $\nu \in \mathcal{P}_T(X)$ ensuring that $T_*^n \nu$ converges to a given invariant measure μ. This idea also connects with the construction of SRB measures by iteratively pushing forward the Lebesgue measure. Define the *stable leaf* of an invariant measure μ by

$$\mathrm{Sl}(\mu, T_*) = \{\nu \in \mathcal{P}(X) \mid T_*^n \nu \to \mu\}$$

where the convergence is in the weak-$*$ topology. Since we are concerned here with the relations between T and S, the question is to relate $\mathrm{Sl}(\mu, T_*)$ to $\mathrm{Sl}(\check{\mu}, S_*)$.

LEMMA 3.4. *Let $\check{\mu} \in \mathcal{P}_S(Y)$ and assume that:*
- *the fibers are $(\bar{a}_n)_n$-shrunk on average with respect to $\check{\mu}$;*
- *T is continuous, and let ω_{T^k} be a modulus of continuity of T^k, for each $k \in \mathbb{N}$;*
- *π induces a continuous fibration with modulus $\bar{\omega}_\pi$.*

Then, for all k, $\ell \in \mathbb{N}$ and all $\nu \in \mathcal{P}(X)$, we have

$$\mathrm{W}(T_*^{k+\ell}\nu, \mu) \le \omega_{T^k} \circ \bar{\omega}_\pi (\mathrm{W}(S_*^\ell \pi_* \nu, \check{\mu})) + \bar{a}_k.$$

Proof. We first prove that, given any $\mu_0 \in \mathcal{P}(X)$ and any $\check{\mu}_1 \in \mathcal{P}(Y)$, there exist $\mu_1 \in \pi_*^{-1}(\check{\mu}_1)$ such that

$$\mathrm{W}(\mu_0, \mu_1) \le \bar{\omega}_\pi(\mathrm{W}(\pi_* \mu_0, \check{\mu}_1)).$$

Let $\check{\gamma}$ be an optimal transport plan from $\pi_* \mu_0$ to $\check{\mu}_1$ and $(\xi_y)_{y \in Y}$ be the disintegration of μ_0 with respect to π. Recall that $\tau_y^{y'}$ is a measurable map from $\pi^{-1}(y)$ to $\pi^{-1}(y')$ such that $d(x, \tau_y^{y'}(x)) \le \bar{\omega}_\pi(d(y, y'))$. We define a measure on $X \times X$ by $\gamma = \int (\mathrm{Id}, \tau_y^{y'})_* \xi_y \, d\check{\gamma}(y, y')$, that is, for $f : X \times X \to \mathbb{R}$,

$$\gamma(f) = \iint f(x, \tau_y^{y'}(x)) \, d\xi_y(x) \, d\check{\gamma}(y, y').$$

The first marginal of γ is μ_0, since when f only depends on its first argument,

$$\gamma(f) = \iint f(x) \, d\xi_y(x) \, d\check{\gamma}(y, y') = \int f(x) \, d\xi_y(x) \, d\pi_* \mu_0(y) = \mu_0(f).$$

Let μ_1 be the second marginal of γ; then $\pi_*\mu_1 = \check{\mu}_1$ since when $f(x, x') = g(\pi(x'))$,

$$\gamma(f) = \iint g(y')\, d\xi_y(x)\, d\check{\gamma}(y, y') = \int g(y')\, d\check{\gamma}(y, y') = \check{\mu}_1(g).$$

We get

$$W(\mu_0, \mu_1) \le \int d(x, x')\, d\gamma(x, x') \le \int d(x, \tau_y^{y'}(x))\, d\xi_y(x)\, d\check{\gamma}(y, y')$$

$$\le \int \bar{\omega}_\pi(d(y, y'))\, d\check{\gamma}(y, y') \le \bar{\omega}_\pi\left(\int d(y, y')\, d\check{\gamma}(y, y')\right)$$

$$= \bar{\omega}_\pi(W(\pi_*\mu_0, \check{\mu}_1)).$$

We now apply this with $\mu_0 = T_*^\ell v$ and $\check{\mu}_1 = \check{\mu}$: there exists $\mu_1 \in \pi_*^{-1}(\check{\mu})$ such that $W(T_*^\ell v, \mu_1) \le \bar{\omega}_\pi(W(\pi_* T_*^\ell v, \check{\mu}))$. Since $\pi_* T_*^\ell v = S_*^\ell \pi_* v$ and $T_*^k \mu = \mu$, we get

$$W(T_*^{k+\ell} v, \mu) \le W(T_*^{k+\ell} v, T_*^k \mu_1) + W(T_*^k \mu_1, \mu)$$

$$\le \omega_{T^k}(W(T_*^\ell v, \mu_1)) + \bar{a}_k$$

$$\le \omega_{T^k} \circ \bar{\omega}_\pi(W(S_*^\ell \pi_* v, \check{\mu})) + \bar{a}_k. \qquad \square$$

COROLLARY 3.5. *If T is continuous, π induces a continuous fibration, and fibers are shrunk on average with respect to $\check{\mu}$, then $\mathrm{Sl}(\mu, T_*) = \pi_*^{-1}(\mathrm{Sl}(\check{\mu}, S_*))$.*

Proof. If $v \in \mathrm{Sl}(\mu, T_*)$, then $S_*^n \pi_* v = \pi_* T_*^n v \to \pi_* \mu = \check{\mu}$, so that $v \in \pi_*^{-1}(\mathrm{Sl}(\check{\mu}, S_*))$.

Assume now that $v \in \mathcal{P}(X)$ is such that $S_*^n \pi_* v \to \check{\mu}$ and let $\varepsilon > 0$. Choose k such that $a_k \le \varepsilon/2$; then there exists $\eta > 0$ such that $r \in [0, \eta] \implies \omega_{T^k} \circ \bar{\omega}_\pi(r) \le \varepsilon/2$. Choose ℓ_0 such that, for all $\ell \ge \ell_0$, $W(S_*^\ell \pi_* v, \check{\mu}) \le \eta$ and apply Lemma 3.4: for all $n \ge k + \ell_0$, we have $W(T_*^n v, \mu) \le \varepsilon$. $\qquad \square$

Proof of Theorem B. According to the statement to be proved, we assume that S has a conformal measure λ_Y and that the corresponding transfer operator $\mathscr{L}$ has a spectral gap on some Banach space of functions $(\mathcal{B}, \|\cdot\|)$ (see Definition 2.15), with eigenfunction h (normalized by $\lambda_Y(h) = 1$). Let $v \in \mathcal{P}(X)$ such that $\pi_* v = f\, d\lambda_Y$ with $f \in \mathcal{B}$, and observe that $\lambda_Y(f) = 1$, so that we can write $f = h + \bar{f}$ where $\lambda_Y(\bar{f}) = 0$. We have $S_*^n(f\, d\lambda_y) = (h + \check{\mathscr{L}}^n \bar{f})\, d\lambda_Y$ and

$$\|\check{\mathscr{L}}^n \bar{f}\|_{L^1(\lambda_Y)} \le \|\check{\mathscr{L}}^n \bar{f}\| \le C(1 - \delta)^n$$

for some $\delta \in (0, 1)$. Since diam $Y \le 1$, the Wasserstein metric is not greater than the total variation distance (take a transport plan that leaves the common mass in place and moves the remaining mass arbitrarily), so that

$$W(h\, d\lambda_Y, (h + \check{\mathscr{L}}^n \bar{f})\, d\lambda_Y) \le \|\check{\mathscr{L}}^n(\bar{f})\, d\lambda_Y\|_{TV} = \|\check{\mathscr{L}}^n \bar{f}\|_{L^1(\lambda_Y)} \le C(1 - \delta)^n.$$

By hypothesis there is some $\theta \in (0, 1)$ such that $a_n \le C\theta^n$ for all $n \in \mathbb{N}$. Applying Lemma 3.4 and denoting by L the Lipschitz constant of T and by $\bar{\omega}_\pi(r) =: K r^\alpha$ the modulus of continuity of the fibration induced by π, we get, for all $k, \ell \in \mathbb{N}$,

$$W(T_*^{k+\ell} v, \mu) \le L^k K (W(h\, d\lambda_Y, (h + \check{\mathscr{L}}^\ell \bar{f})\, d\lambda_Y))^\alpha + a_k \le C L^k (1 - \delta)^{\alpha\ell} + C\theta^k.$$

$$\tag{3}$$

Take $\beta \in (0, 1)$ such that

$$\beta < \frac{\alpha \log (1/(1 - \delta))}{\log L + \alpha \log (1/(1 - \delta))}$$

and define two integer sequences such that $k_n = \beta n + O(1)$, $\ell_n = (1 - \beta)n + O(1)$ and $n = k_n + \ell_n$. Applying (3) yields $W(T_*^n \nu, \mu) \le C \eta^n$ for some $\eta \in (0, 1)$. □

4. Preserved properties of lifted invariant measures

Standing Assumption From now on the map T is assumed to be an extension of S with shrinking fibers.

This assumption will remain active until the end of the paper, and we will only restate it when we want to specify the rate of shrinking or for the most important results.

With the uniqueness of the T-invariant lift of each S-invariant measure comes naturally the problem of which special properties of invariant measures are preserved under lifting (we shall later be specifically concerned with statistical properties). Theorem A is the concatenation of Theorem 3.1 with the main results of the present section. We shall use the material of §§2.2, 2.3 and 2.5.

4.1. Ergodicity and mixing.
It is known that ergodicity is preserved by the lift map π^*; see [**APPV09**] and [**BM17**]. We give an alternative proof, taking advantage of uniqueness in Theorem 3.1.

PROPOSITION 4.1. *For all $\mu \in \mathcal{P}_T(X)$, μ is ergodic if and only if $\pi_* \mu$ is ergodic.*

Proof. This follows from π_* being an affine map inducing a homeomorphism $\mathcal{P}_T(X) \to \mathcal{P}_S(Y)$ (Corollary 3.3), since ergodic measures are the extremal points of the convex set of invariant measures.

Assume, indeed, that μ is not ergodic: then we can write $\mu = p\mu_0 + (1 - p)\mu_1$ where $\mu_0 \ne \mu_1 \in \mathcal{P}_T(X)$ and $p \in (0, 1)$. The three measures $\pi_*\mu$, $\pi_*\mu_0$ and $\pi_*\mu_1$ are S-invariant and satisfy $\pi_*\mu = p\pi_*\mu_0 + (1 - p)\pi_*\mu_1$. Moreover, $\pi_*\mu_0 \ne \pi_*\mu_1$ and thus $\pi_*\mu$ is not ergodic. If $\pi_*\mu$ is not ergodic, then similarly a decomposition lifts and μ is not ergodic either. □

PROPOSITION 4.2. *A measure $\mu \in \mathcal{P}_T(X)$ is weakly mixing if and only if $\pi_* \mu$ is.*

Proof. That $\check{\mu} := \pi_*\mu$ is weakly mixing is equivalent to $\check{\mu} \otimes \check{\mu}$ being ergodic for the diagonal action $S \times S$ on $Y \times Y$ (see, for example, [**Wal82**, Theorem 1.24]).

The map $T \times T$ is an extension of $S \times S$ with factor map $\pi \times \pi : X \times X \to Y \times Y$ and fibers $(\pi \times \pi)^{-1}(y_0, y_1) = \pi^{-1}(y_0) \times \pi^{-1}(y_1)$. If we endow products with the ℓ^∞ combined metric, for example, $d((x_0, x_1), (x_0', x_1')) := \max(d(x_0, x_0'), d(x_1, x_1'))$, then the diameter of $(T \times T)^n([x, x'])$ (where we recall that $[\cdot]$ denotes fibers) is the maximum diameter of $T^n([x])$, $T^n([x'])$, so that $T \times T$ is an extension of $S \times S$ with shrinking fibers.

By Proposition 4.1, the ergodicity of $\check{\mu} \otimes \check{\mu}$ is equivalent to the ergodicity of its lift $\mu \otimes \mu$, and thus $\check{\mu}$ is weakly mixing if and only if μ is. □

We now turn to strong mixing. Recall that μ is said to be *strongly mixing* if, for all f, $g \in L^2(\mu)$, $\mathcal{C}_\mu^n(f, g) \to 0$ as $n \to \infty$. We shall relate observables f, $g : X \to \mathbb{R}$ to observables on Y, which amounts to constructing observables that are constant along fibers. As stressed in [**BM17**], a natural solution is to use the average along the disintegration $(\xi_y)_{y \in Y}$ of μ with respect to $\check{\mu}$ (the disintegration theorem is recalled above as Proposition 2.6). Given a Borel function $f : X \to \mathbb{R}$, we define $\xi(f) : Y \to \mathbb{R}$ by $\xi(f)(y) = \xi_y(f)$ and $\tilde{f} = \xi(f) \circ \pi$. In this way, $\tilde{f}$ is an observable on X which is constant on fibers; moreover,

$$\int \tilde{f} \, d\mu = \int \xi(f) \, d\check{\mu} = \int f \, d\mu.$$

By convexity, $\xi(f)^p \leq \xi(f^p)$, so that $f \in L^p(\mu) \implies \xi(f) \in L^p(\check{\mu})$ for all $p \in [1, \infty]$. It is obvious that, for all $u \in L^p(\check{\mu})$ and $v \in L^{p'}(\check{\mu})$ (where $1/p + 1/p' = 1$, possibly $\{p, p'\} = \{1, \infty\}$), we have $\mathcal{C}_\mu^n(u \circ \pi, v \circ \pi) = \mathcal{C}_{\check{\mu}}^n(u, v)$. We will need a slightly stronger observation.

LEMMA 4.3. *If $u \in L^p(\check{\mu})$ and $g \in L^{p'}(\mu)$, then $\mathcal{C}_\mu^n(u \circ \pi, g) = \mathcal{C}_{\check{\mu}}^n(u, \xi(g))$.*

Proof. Since for all $y \in Y$, ξ_y is supported on $\pi^{-1}(y)$, we have

$$\int u \circ \pi \cdot g \, d\mu = \int \left(\int u \circ \pi(x) \cdot g(x) \, d\xi_y(x) \right) d\check{\mu}(y) = \int u \cdot \xi(g) \, d\check{\mu}.$$

Applying this to $u \circ \pi \circ T^n = u \circ S^n \circ \pi$, we get the desired result. $\square$

The next lemma is inspired by [**AGY06**, Lemma 8.2], and will be used immediately to prove that the strong mixing property lifts to extensions with shrinking fibers, and reused later to study rates of decay of correlations.

LEMMA 4.4. *Assume that the fibers are $(a_n)_n$-shrinking, let $\mu \in \mathcal{P}_T(X)$ and $\check{\mu} = \pi_*\mu$ and let f, $g : X \to \mathbb{R}$ be two observables with f continuous of modulus ω and $g \in L^1(\mu)$. For all k, $m \in \mathbb{N}$, we have*

$$\mathcal{C}_\mu^{k+m}(f, g) \leq \mathcal{C}_{\check{\mu}}^m(\xi(f \circ T^k), \xi(g)) + \omega(a_k)\|g\|_{L^1(\mu)}.$$

Proof. Up to adding a constant, we assume $\mu(f) = 0$. For each $y \in Y$,

$$\sup_{\pi^{-1}(y)} f \circ T^k - \inf_{\pi^{-1}(y)} f \circ T^k < \omega(a_k).$$

After integration with respect to ξ_y, we obtain that $(f \circ T^k)^{\tilde{}} := \xi(f \circ T^k) \circ \pi$ and $f \circ T^k$ are $\omega(a_k)$-close in the uniform norm, so that

$$\mathcal{C}_\mu^{k+m}(f, g) = \mathcal{C}_\mu^m(f \circ T^k, g) \leq \mathcal{C}_\mu^m((f \circ T^k)^{\tilde{}}, g) + \omega(a_k)\|g\|_{L^1(\mu)}.$$

Applying Lemma 4.3, we get $\mathcal{C}_\mu^m((f \circ T^k)^{\tilde{}}, g) = \mathcal{C}_{\check{\mu}}^m(\xi(f \circ T^k), \xi(g))$. $\square$

PROPOSITION 4.5. *A measure $\mu \in \mathcal{P}_T(X)$ is strongly mixing if and only if $\pi_*\mu$ is.*

Proof. Assume first that μ is strongly mixing, and let u, $v : Y \to \mathbb{R}$ be observables in $L^2(\check{\mu})$. Since $\mathcal{C}_\mu^n(u \circ \pi, v \circ \pi) = \mathcal{C}_{\check{\mu}}^n(u, v)$ and μ is strongly mixing, this goes to 0 as n goes to ∞. (This is classical and does not use the shrinking property).

Assume now that $\check{\mu}$ is strongly mixing. Given $f, g \in L^2(\mu)$, define $\xi(f), \tilde{f}, \xi(g), \tilde{g}$ as above and recall that $\xi(f), \xi(g) \in L^2(\check{\mu})$. Fix $\varepsilon > 0$ and let h be a continuous approximation of f, with $\|f - h\|_{L^2(\mu)} < \varepsilon$. We have

$$\left| \int (f - h) \circ T^n \cdot g \, d\mu \right| \leq \left(\int (f - h)^2 \circ T^n \, d\mu \right)^{1/2} \left(\int g^2 \, d\mu \right)^{1/2}$$

$$\leq \left(\int (f - h)^2 \, d\mu \right)^{1/2} \left(\int g^2 \, d\mu \right)^{1/2} \leq \varepsilon \|g\|_{L^2(\mu)}$$

so that

$$\mathcal{C}_\mu^n(f, g) \leq \mathcal{C}_\mu^n(h, g) + \varepsilon \|g\|_{L^2(\mu)}. \tag{4}$$

Let ω be a modulus of continuity of h, and let $(a_n)_n$ be a shrinking sequence. There is a k such that $\omega(a_k) < \varepsilon$. By Lemma 4.4 and using $\|\cdot\|_{L^1(\mu)} \leq \|\cdot\|_{L^2(\mu)}$,

$$\mathcal{C}_\mu^{k+m}(h, g) \leq \mathcal{C}_{\check{\mu}}^m(\xi(h \circ T^k), \xi(g)) + \varepsilon \|g\|_{L^2(\mu)}. \tag{5}$$

Combining (4) and (5), we get $\mathcal{C}_\mu^{k+m}(f, g) \leq \mathcal{C}_{\check{\mu}}^m(\xi(h \circ T^k), \xi(g)) + 2\varepsilon \|g\|_{L^2(\mu)}$. Since $\check{\mu}$ is strongly mixing, there is an m_0 such that, for all $n > m_0 + k$, $\mathcal{C}_\mu^n(f, g) \leq (1 + 2\|g\|_{L^2(\mu)})\varepsilon$, and μ is strongly mixing. □

4.2. *Entropy.* Entropy preservation in Theorem A is unsurprising and, thanks to the uniqueness in Theorem 3.1, follows easily from the relative variational principle established by Ledrappier and Walters [**LW77**]: for all $\check{\mu} \in \mathcal{P}_S(Y)$,

$$\sup_{\mu \in \pi_*^{-1}(\check{\mu})} h_{\mathrm{KS}}(T, \mu) = h_{\mathrm{KS}}(S, \check{\mu}) + \int h(T, \pi^{-1}(y)) \, d\check{\mu}(y)$$

where $h_{\mathrm{KS}}(T, \mu)$ is the Kolmogorov–Sinai entropy and $h(T, K)$ is the topological entropy of T on the (not necessarily invariant) compact set $K \subset X$.

PROPOSITION 4.6. *If S, T are continuous, then $h(T, \mu) = h(S, \pi_*\mu)$ for all $\mu \in \mathcal{P}_T(X)$.*

Proof. Let $\mu \in \mathcal{P}_T(X)$ and $\check{\mu} = \pi_*\mu$. By Theorem 3.1, $\pi_*^{-1}(\check{\mu}) = \{\mu\}$, so that the Ledrappier–Walters relative variational principle reads $h_{\mathrm{KS}}(T, \mu) = h_{\mathrm{KS}}(S, \check{\mu}) + \int h(T, \pi^{-1}(y)) \, d\check{\mu}(y)$, and we are left with proving $h(T, \pi^{-1}(y)) \equiv 0$.

Let $y \in Y$ and $\delta > 0$. There is an $n_0 \in \mathbb{N}$ such that, for all $n > n_0$, $a_n < \delta$. Let E_0 be a maximal (n_0, δ)-separated set of $\pi^{-1}(y)$. For all $x, x' \in E_0$ and all m such that $n_0 < m \leq n$, $d(T^m x, T^m x') \leq a_m < \delta$ so that E_0 must be (n, δ)-separated as well. It follows that the cardinal of an (n, δ)-separated set is bounded independently of n, and therefore $h(T, \pi^{-1}(y)) = 0$. □

4.3. *Absolute continuity, physicality, observability.* Assume here that X and Y are equipped with reference measures λ_X and λ_Y. In general, the lift to an extension with shrinking fibers of an Acip is not itself an Acip; the map S could have an Acip while T does not (e.g. take Y to be a point, T contracting). However, the weaker property of physicality is preserved under a mild regularity assumption on π.

PROPOSITION 4.7. *Assume that T is continuous and X, Y equipped with reference measures with respect to which π is non-singular. A measure $\mu \in \mathcal{P}_T(X)$ is physical if and only if $\pi_*\mu$ is.*

Proof. As usual, we set $\check{\mu} = \pi_* \mu$. For all $x \in X$, we have $\pi_*((1/n) \sum_{k=0}^{n-1} \delta_{T^k x}) = (1/n) \sum_{k=0}^{n-1} \delta_{S^k \pi(x)}$. If $(1/n) \sum_{k=0}^{n-1} \delta_{T^k x}$ converges to some $\nu \in \mathcal{P}_T(X)$, then $(1/n) \sum_{k=0}^{n-1} \delta_{S^k \pi(x)}$ converges to $\pi_* \nu$. This proves that $\mathrm{Ba}(\mu) \subset \pi^{-1}(\mathrm{Ba}(\check{\mu}))$. We get equality by compactness: if $\pi(x) \in \mathrm{Ba}(\check{\mu})$, then any cluster point of the sequence $((1/n) \sum_{k=0}^{n-1} \delta_{T^k x})_n$ is mapped by π_* to $\check{\mu}$. Since T is continuous, such cluster points are T-invariant, so that Theorem 3.1 implies that μ is the unique cluster point of the sequence, hence its limit.

Since π is non-singular, $\lambda_Y(\mathrm{Ba}(\check{\mu})) > 0$ if and only if $\lambda_X(\pi^{-1}(\mathrm{Ba}(\check{\mu}))) > 0$, that is, $\check{\mu}$ is physical if and only if μ is physical. □

Since ergodic Acips are particular cases of physical measures, while they do not necessarily lift to Acips, they do lift to physical measures. This implies that many weakly hyperbolic systems have physical measures (see, for example, Corollaries 1.7, 1.9).

It is not much more difficult to lift observability.

PROPOSITION 4.8. *Assume that T is continuous and X, Y are equipped with reference measures with respect to which π is non-singular. A measure $\mu \in \mathcal{P}_T(X)$ is observable if and only if $\pi_* \mu$ is.*

Proof. Given any $\varepsilon > 0$, since π_* is continuous there exists some $\eta > 0$ such that, for all $\mu_0, \mu_1 \in \mathcal{P}(X)$, $\mathrm{W}(\mu_0, \mu_1) < \eta \implies \mathrm{W}(\pi_* \mu_0, \pi_* \mu_1) < \epsilon$. Let $x \in \mathrm{Ba}_\eta(\mu)$: there exist $\mu_0 \in \mathcal{P}_T(X)$ such that $\mathrm{W}(\mu_0, \mu) < \eta$ and an increasing sequence of positive integers $(n_k)_{k \in \mathbb{N}}$ such that $\mu_0 = \lim_k (1/n_k) \sum_{j=0}^{n_k-1} \delta_{T^j(x)}$. Then $(1/n_k) \sum_{j=0}^{n_k-1} \delta_{S^j \pi(x)} \to \pi_* \mu_0 \in \mathcal{P}_S(Y)$ and $\mathrm{W}(\pi_* \mu_0, \check{\mu}) \leq \varepsilon$, so that $\pi(x) \in \mathrm{Ba}_\varepsilon(\check{\mu})$. We have proved $\mathrm{Ba}_\eta(\mu) \subset \pi^{-1}(\mathrm{Ba}_\varepsilon(\check{\mu}))$; if μ is observable, then $\lambda_X(\mathrm{Ba}_\eta(\mu)) > 0$ and, by non-singularity of π, we deduce that $\check{\mu}$ is observable.

Since π^* is continuous, for all $\varepsilon > 0$ there exists an $\eta > 0$ such that, for all $\nu_0, \nu_1 \in \mathcal{P}_S(Y)$, $\mathrm{W}(\nu_0, \nu_1) < \eta \implies \mathrm{W}(\pi^* \nu_0, \pi^* \nu_1) < \epsilon$. Let $x \in X$ such that $\pi(x) \in \mathrm{Ba}_\eta(\check{\mu})$. There exist $\nu_0 \in \mathcal{P}_S(Y)$ such that $\mathrm{W}(\check{\mu}, \nu_0) < \eta$ and an increasing sequence of positive integers $(n_k)_{k \in \mathbb{N}}$ such that $\nu_0 = \lim_k (1/n_k) \sum_{j=0}^{n_k-1} \delta_{S^j \pi(x)} = \lim_k \pi_*((1/n_k) \sum_{j=0}^{n_k-1} \delta_{T^j x})$. It follows that any cluster point of $((1/n_k) \sum_{j=0}^{n_k-1} \delta_{T^j x})_k$ is mapped by π_* to ν_0. Since T is continuous, these cluster points are T-invariant and there is only one of them, $\pi^* \nu_0$. Moreover, since $\mathrm{W}(\nu_0, \check{\mu}) < \eta$, we have $\mathrm{W}(\pi^* \nu_0, \pi^* \check{\mu}) < \varepsilon$, so that $x \in \mathrm{Ba}_\varepsilon(\mu)$. We have thus proved that $\pi^{-1}(\mathrm{Ba}_\eta(\check{\mu})) \subset \mathrm{Ba}_\varepsilon(\mu)$, from which we deduce that if $\check{\mu}$ is observable, then so is μ. □

5. *Equilibrium states and statistical properties*

In this section we consider some classical objects and properties that form the core of the thermodynamical formalism, and lift them from S to T. This will, for example, be used to recover information about certain hyperbolic maps from information about expanding maps. This is an old strategy, notably well developed in symbolic dynamics, that has been extended more generally through Markov partition and coding. More recently, a 'direct lifting' approach has been frequently used, often in quite specific cases. Our goal is to use this approach in the most general way while keeping all proofs simple. We shall use §§2.1, 2.2 (definition of a section) and 2.6.

5.1. *Equilibrium states.*   Let $\varphi : X \to \mathbb{R}$ be a function, here called a *potential*, to be interpreted physically (up to sign) as a density of energy: a T-invariant measure μ is called a 'state', the total energy of the system in state μ being $-\mu(\varphi)$. The 'free energy' is then $\mathscr{F}(\mu) := h_{\mathrm{KS}}(T, \mu) + \mu(\varphi)$, and we seek *equilibrium states*, that is, invariant measures maximizing free energy. The main questions underlying the 'thermodynamical formalism' are existence, uniqueness and statistical properties of equilibrium states.

Here of course we want to relate this to the corresponding situation for S; since $\check{\mu} = \pi_* \mu$ is the state on Y corresponding to μ and $h_{\mathrm{KS}}(S, \check{\mu}) = h_{\mathrm{KS}}(T, \mu)$ (Proposition 4.6), one only needs to consider the energy term. We would thus like to construct a potential $\check{\varphi} : Y \to \mathbb{R}$ related to φ; using the disintegration of μ to construct $\xi(\varphi)$ as in §4.1 is not suitable here since invariant measures are to be considered all at once and compared. Rather, we will add a suitable *coboundary* to φ, as is classically done in the case of shifts; see [**Bow08**].

Coboundaries are defined as the functions of the form $h - h \circ T : X \to \mathbb{R}$. They are important because, for *all* T-invariant measures μ, we have $\mu(h - h \circ T) = \mu(h) - \mu(h \circ T) = 0$: adding a coboundary to a potential does not change its energy with respect to any state. We will construct a potential $\hat{\varphi} = \varphi + h - h \circ T$ that is constant on fibers (then $\hat{\varphi} = \check{\varphi} \circ \pi$ will define $\check{\varphi} : Y \to \mathbb{R}$).

LEMMA 5.1. *Assume that fibers are $(a_n)_n$-shrinking, and that π admits a continuous section $\sigma : Y \to X$. Let $\varphi : X \to \mathbb{R}$ be a ω-continuous potential where $\sum_{n \geq 0} \omega(a_n) < \infty$, and set $h = \sum_{n=0}^{\infty} (\varphi T^n \sigma \pi - \varphi T^n)$. Then $h : X \to \mathbb{R}$ is well defined and $\hat{\varphi} := \varphi + h - h \circ T$ is constant on each fiber.*

If T is continuous, so is h. If T is L-Lipschitz, then, for all $x, x' \in X$ and all $N \in \mathbb{N}$,

$$|h(x) - h(x')| \leq 2 \operatorname{Hol}_\omega(\varphi) \left(\sum_{n=0}^{N} \omega\big(L^n \omega_{\sigma\pi}(d(x, x'))\big) + \sum_{n > N} \omega(a_n) \right)$$

where $\omega_{\sigma\pi}$ is any modulus of continuity of $\sigma\pi$ such that $\omega_{\sigma\pi}(r) \geq r$ for all $r \in [0, \operatorname{diam} X]$.

(The assumption on $\omega_{\sigma\pi}$ can always be obtained up to increase the modulus, and is only meant to simplify the conclusion.)

Proof. Let $H = \operatorname{Hol}_\omega(\varphi)$. For all $x \in X$, $\sigma\pi(x)$ and x lie on the same fiber, so that $d(T^n \sigma\pi(x), T^n(x)) \leq a_n$ and $|\varphi T^n \sigma\pi(x) - \varphi T^n(x)| \leq H\omega(a_n)$. The convergence of $\sum \omega(a_n)$ ensures the uniform convergence of the series defining h which is therefore well defined, and continuous whenever T is.

Next, we have

$$\hat{\varphi} = \varphi + \sum_{n=0}^{\infty} (\varphi T^n \sigma \pi - \varphi T^n) - \sum_{n=0}^{\infty} (\varphi T^n \sigma \pi T + \varphi T^{n+1})$$

$$= \varphi + \varphi \sigma \pi - \varphi + \sum_{n=0}^{\infty} (\varphi T^{n+1} \sigma \pi - \varphi T^{n+1} - \varphi T^n \sigma S \pi + \varphi T^{n+1})$$

$$= \varphi \sigma \pi + \sum_{n=0}^{\infty} (\varphi T^{n+1} \sigma \pi - \varphi T^n \sigma S \pi)$$

which is constant on fibers since π factors on the right. Assume now that T is L-Lipschitz and $\sigma\pi$ has modulus of continuity $\omega_{\sigma\pi}$. Then

$$
\begin{aligned}
|h(x) - h(x')| &\leq \sum_{n=0}^{\infty} |\varphi T^n \sigma\pi(x) - \varphi T^n(x) - \varphi T^n \sigma\pi(x') + \varphi T^n(x')| \\
&\leq \sum_{n=0}^{N} |\varphi T^n \sigma\pi(x) - \varphi T^n \sigma\pi(x')| + \sum_{n=0}^{N} |\varphi T^n(x) - \varphi T^n(x')| \\
&\quad + \sum_{n>N} |\varphi T^n \sigma\pi(x) - \varphi T^n(x)| + \sum_{n>N} |\varphi T^n \sigma\pi(x') - \varphi T^n(x')| \\
&\leq \sum_{n=0}^{N} H\omega(L^n \omega_{\sigma\pi}(d(x, x'))) + \sum_{n=0}^{N} H\omega(L^n d(x, x')) + 2\sum_{n>N} H\omega(a_n).
\end{aligned}
$$

We conclude by using $d(x, x') \leq \omega_{\sigma\pi}(d(x, x'))$. □

THEOREM 5.2. *Assume that T is an extension of S with $(a_n)_n$-shrinking fibers, that π has modulus of continuity $\omega_\pi(r) \geq r$ and admits a Lipschitz section $\sigma : Y \to X$, and that T is L-Lipschitz. Let $\omega, \check{\omega}$ be two moduli of continuity with $\check{\omega} \gtrsim \omega$.*

If, for some constant $D > 0$ and for all $r \in [0, \mathrm{diam}\, X]$, there exists some $N = N(r) \in \mathbb{N}$ such that

$$
\sum_{n>N} \omega(a_n) \leq D\check{\omega}(r) \quad and \quad \sum_{n=0}^{N} \omega(L^n \omega_\pi(r)) \leq D\check{\omega}(r), \tag{6}
$$

then:

(i)   *for all ω-continuous potential $\varphi : X \to \mathbb{R}$ there is an $\check{\omega}$-continuous potential $\check{\varphi} : Y \to \mathbb{R}$ such that φ differs from $\hat{\varphi} = \check{\varphi} \circ \pi$ by a coboundary;*

(ii) *for all $\mu \in \mathcal{P}_T(X)$, writing $\check{\mu} = \pi_* \mu$, we have*

$$
h_{\mathrm{KS}}(T, \mu) + \mu(\varphi) = h_{\mathrm{KS}}(S, \check{\mu}) + \check{\mu}(\check{\varphi})
$$

   (in particular, π^* realizes a bijection between equilibrium states of $\check{\varphi}$ and equilibrium states of φ);

(iii) *we can realize $\varphi \mapsto \check{\varphi}$ as a continuous linear map from $\mathrm{Hol}_\omega(X)$ to $\mathrm{Hol}_{\check{\omega}}(Y)$.*

Proof. Given $\varphi \in \mathrm{Hol}_\omega(X)$, let h be the function defined by Lemma 5.1. The hypotheses are tailored to ensure that h is $\check{\omega}$-continuous (more precisely, $\mathrm{Hol}_{\check{\omega}}(h) \leq 4D\, \mathrm{Hol}_\omega(\varphi)$). It follows that $\mathrm{Hol}_{\check{\omega}}(h \circ T) \leq 4LD\, \mathrm{Hol}_\omega(\varphi)$, and the potential $\hat{\varphi} = \varphi + h - h \circ T$ is $\check{\omega}$-continuous. Since σ is Lipschitz, $\check{\varphi} := \hat{\varphi} \circ \sigma$ is also $\check{\omega}$-continuous. Since $\hat{\varphi}$ is constant on fibers, $\check{\varphi} \circ \pi = \hat{\varphi}$.

The equality of free energies follows from the equality of entropies (Proposition 4.6) and from $\mu(\varphi) = \mu(\hat{\varphi}) = \pi_* \mu(\check{\varphi})$.

The fact that $\varphi \mapsto \check{\varphi}$ is continuous linear follows from the construction. □

From here the game consists of finding the optimal choice of $N(r)$ depending on the available assumptions. We will restrict in the following to the case where π is Hölder-continuous, a common situation in hyperbolic dynamics.

COROLLARY 5.3. *Assume that T is L-Lipschitz for some $L \geq 1$, that σ is Lipschitz and that π is β-Hölder.*

(i)   *If the fibers are exponentially shrinking with ratio $\theta \in (0, 1)$ and $\omega = \omega_\alpha$ is a Hölder modulus of continuity, then the conclusions of Theorem 5.2 hold with $\check{\omega} = \omega_\gamma$ where $\gamma = \alpha\beta/(1 - \log L/\log \theta)$.*

(ii) *If the fibers are polynomially shrinking with degree $d > 0$, and $\omega = \omega_\alpha$ where $\alpha > 1/d$, then the conclusions of Theorem 5.2 hold with $\check{\omega} = \omega_{\alpha' \log}$ where $\alpha' = \alpha d - 1$.*

(iii) *If the fibers are exponentially shrinking and $\omega = \omega_{\alpha \log}$ with $\alpha > 1$, then the conclusions of Theorem 5.2 hold with $\check{\omega} = \omega_{\alpha' \log}$ where $\alpha' = (\alpha - 1)/2$.*

Note that we can always replace $\check{\omega}$ with a larger modulus if needed. In particular, in the case of exponentially shrinking fibers, if each Hölder-continuous potential on Y has a unique equilibrium state for S, then each Hölder-continuous potential on X has a unique equilibrium state for T.

Proof. We apply Theorem 5.2 three times. For (i), take $N(r) = \gamma \log r/(\alpha \log \theta) + O(1)$; then

$$\sum_{n>N} \omega(a_n) \lesssim \theta^{\alpha N} \lesssim r^\gamma \quad \text{and} \quad \sum_{n=0}^{N} \omega(L^n \omega_\pi(r)) \lesssim L^{\alpha N} r^{\alpha\beta} \lesssim r^{\gamma'}$$

with $\gamma' = \gamma(\log L/\log \theta) + \alpha\beta = \gamma$.

For (ii), take $N(r) = \eta \log(r_{\alpha'}/r)/\log L + O(1)$ with any $\eta < \beta$; then

$$\sum_{n>N} \omega(a_n) \lesssim \frac{1}{N^{\alpha d - 1}} \lesssim \frac{1}{(\log(r_{\alpha'}/r))^{\alpha'}}$$

and

$$\sum_{n=0}^{N} \omega(L^n \omega_\pi(r)) \lesssim L^{\alpha N} r^{\alpha\beta} \lesssim r^{\alpha(\beta-\eta)} \ll \omega_{\alpha' \log}(r).$$

For (iii), take $N(r) = (\log(r_{\alpha'}/r))^{1/2} + O(1)$; then, using Proposition 2.3 for the second term,

$$\sum_{n>N} \omega(a_n) \lesssim \frac{1}{N^{\alpha-1}} \lesssim \frac{1}{(\log(r_{\alpha'}/r))^{\alpha'}}$$

and

$$\sum_{n=0}^{N} \omega(L^n \omega_\pi(r)) \lesssim \frac{N^{\alpha+1}}{(\log(r_\alpha/r^\beta))^\alpha} \lesssim \frac{1}{(\log(r_{\alpha'}/r))^{\alpha'}}. \qquad \square$$

5.2. *Statistical properties.*   We would now like to lift statistical properties, assuming them for $\check{\mu} \in \mathcal{P}_S(Y)$ and deducing them for its lift $\mu \in \mathcal{P}_T(X)$. One can in principle lift a decay of correlations (which we will consider next) and then use it to prove statistical properties, but it is in fact simpler to use Theorem 5.2 on observables to lift statistical properties directly.

PROPOSITION 5.4. *Assume π has a section $\sigma : Y \to X$. Consider $\mu \in \mathcal{P}_T(X)$, $\check{\mu} = \pi_* \mu$, $\mathcal{T} \in \{\text{LIL}, \text{CLT}, \text{ASIP}\}$, and let ω, $\check{\omega}$ be two moduli of continuity. If*

(i)   *for each $f \in \mathrm{Hol}_\omega(X)$ there is a continuous $h : X \to \mathbb{R}$ such that $\hat{f} = f + h - h \circ T$ is constant on fibers and $\check{f} = \hat{f} \circ \sigma$ belongs to $\mathrm{Hol}_{\check\omega}(Y)$, and*

(ii) *$\check\mu$ satisfies $\mathcal{T}$ for all $\check\omega$-continuous observables,*

then μ satisfies $\mathcal{T}$ for all ω-continuous observables, with the same parameters as $\check\mu$ (see Definition 2.19).

Proof. This is classical and straightforward. For all $f \in \mathrm{Hol}_\omega(X)$, we have

$$\sum_{k=1}^{n} \hat{f} \circ T^k = \sum_{k=1}^{n} f \circ T^k + h \circ T - h \circ T^{n+1} = \sum_{k=1}^{n} f \circ T^k + O(1)$$

where the $O(1)$ is bounded in the uniform norm, and $\sum_{k=1}^{n} \hat{f} \circ T^k = (\sum_{k=1}^{n} \check{f} \circ S^k) \circ \pi$.

Then, when $(\check\mu, \check{f})$ satisfy the LIL with some variance $\sigma_{\check{f}} > 0$, we have

$$\frac{\sum_{k=1}^{n} f \circ T^k(x) - n\mu(f)}{\sqrt{2n \log \log n}} = \frac{\sum_{k=1}^{n} \check{f} \circ S^k(\pi(x)) - n\check\mu(\check{f})}{\sqrt{2n \log \log n}} + o(1),$$

and the superior limit is $\sigma_{\check{f}}$ for all $x \notin \pi^{-1}(E)$ for some $\check\mu$-negligible set E. Since $\check\mu = \pi_*\mu$, $\pi^{-1}(E)$ is μ-negligible.

In the case of the CLT, for all $\varepsilon > 0$, for all n large enough,

$$\left\| \sum_{k=1}^{n} \hat{f} \circ T^k - \sum_{k=1}^{n} f \circ T^k \right\|_\infty \leq \varepsilon \sqrt{n}$$

and therefore

$$\mu\left\{ x \in X : \frac{1}{\sqrt{n}} \sum_{k=1}^{n} f \circ T^k(x) \leq r \right\} \geq \mu\left\{ x \in X : \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \hat{f} \circ T^k(x) \leq r - \varepsilon \right\}$$

$$= \check\mu\left\{ y \in Y : \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \check{f} \circ S^k(y) \leq r - \varepsilon \right\}$$

$$\to G_{\check\mu(\check{f}), \sigma_{\check{f}}}(r - \varepsilon)$$

so that, by continuity of $G_{\check\mu(\check{f}), \sigma_{\check{f}}} = G_{\mu(f), \sigma_{\check{f}}}$,

$$\liminf_{n \to \infty} \mu\left\{ x \in X : \frac{1}{\sqrt{n}} \sum_{k=1}^{n} f \circ T^k(x) \leq r \right\} \geq G_{\mu(f), \sigma_{\check{f}}}(r).$$

The superior limit is treated in the same way, and we get the CLT for μ, with the same variance.

In the case of the ASIP, we have processes $(\check{A}_k)_{k \in \mathbb{N}}$, whose law is the same as that of $(\check{f} \circ S^k(\check{Z}))_{k \in \mathbb{N}}$ where $\check{Z}$ has law $\check\mu$, and $(B_k)_{k \in \mathbb{N}}$, independent Gaussian of mean $\check\mu(\check{f}) = \mu(f)$ and variance $\sigma_{\check{f}}$, such that $|\sum_{k=1}^{n} \check{A}_k - \sum_{k=1}^{n} B_k| = o(n^\lambda)$ almost surely. We first construct a random variable Z with law μ such that $\check{Z} = \pi(Z)$: up to enriching Ω, we can assume we have a uniform random variable V on $[0, 1]$ independent of all previous random variables, and by measurable selection we have a measurable family of measurable maps $\Xi_y : [0, 1] \to X$ such that $\Xi_y(V)$ has law ξ_y, where $(\xi_y)_{y \in Y}$ is the disintegration of μ

with respect to π. Then $Z = \Xi_{\check{Z}}(V)$ is the desired random variable; now $(\check{A}_k)_{k\in\mathbb{N}}$ has the same law as $(\hat{f} \circ T^k(Z))_{k\in\mathbb{N}}$. Since $f \circ T^k(Z) = \hat{f} \circ T^k(Z) - h(T^k Z) + h(T^{k+1} Z)$, there is a process $(H_k)_k$ with the same law as $(h(T^k Z) - h(T^{k+1} Z))_k$ such that $(A_k)_k = (\check{A}_k - H_k)_k$ has the same law as $(f \circ T^k(Z))_k$; in particular, $\sum_{k=1}^n H_k$ has the same law as $h(TZ) - h(T^{n+1}Z)$ and is bounded almost surely by $2\|h\|_\infty$. Finally, almost surely

$$\left| \sum_{k=1}^n A_k - \sum_{k=1}^n B_k \right| \leq \left| \sum_{k=1}^n \check{A}_k - \sum_{k=1}^n B_k \right| + \left| \sum_{k=1}^n H_k \right| = o(n^\lambda) + O(1) = o(n^\lambda). \qquad \square$$

Theorem C follows directly from Proposition 5.4 and Corollary 5.3. More generally, applying Theorem 5.2, we obtain the following result.

THEOREM 5.5. *Assume that T is an L-Lipschitz extension of S with $(a_n)_n$-shrinking fibers, that the factor map π is β-Hölder-continuous and that there is a Lipschitz section $\sigma : Y \to X$.*

We consider moduli of continuity ω_p, ω_o, $\check{\omega}_p$, $\check{\omega}_o$ where p stands for 'potential' and o for 'observable' and a limit theorem $\mathcal{T} \in \{\text{LIL, CLT, ASIP}\}$.

If S satisfies $\mathrm{UE}(\check{\omega}_p[\rho]; \mathcal{T}, \check{\omega}_o)$ and if, for some constant $D > 0$ and for all $r \in [0, \mathrm{diam}\, X]$, there exists some $N = N(r) \in \mathbb{N}$ such that, for each $i \in \{o, p\}$,

$$\sum_{n>N} \omega_i(a_n) \leq D\check{\omega}_i(r) \quad and \quad \sum_{n=0}^N \omega_i(L^n r^\beta) \leq D\check{\omega}_i(r), \tag{7}$$

then T satisfies $\mathrm{UE}(\omega_p[C\rho]; \mathcal{T}, \omega_o)$ (with the same parameters in $\mathcal{T}$ as for S).

6. *Decay of correlations*

The fact that strong mixing is preserved by the lift map hints at the fact that decay of correlation, which quantifies mixing for regular enough observables, might also lift from $(S, \check{\mu})$ to (T, μ). This section uses §§2.1, 2.2 and 2.5.

Assume we have some decay of correlations for observables of a given regularity for $(S, \check{\mu})$. Given $f, g : X \to \mathbb{R}$, we have $\mathcal{C}_\mu^n(\tilde{f}, \tilde{g}) = \mathcal{C}_{\check{\mu}}^n(\xi(f), \xi(g))$, and Lemma 4.4 relates $\mathcal{C}_\mu^n(f, g)$ to $\mathcal{C}_\mu^n(\tilde{f}, \tilde{g})$ (up to composition with T^n and a shift in n). The crucial missing piece is to understand whether the disintegration $(\xi_y)_{y\in Y}$ preserves regularity. For example, if f is Hölder does it follow that $\xi(f)$ is Hölder?

6.1. *Regularity of the disintegration.* The following result is close to [**BM17**, Proposition 3], but we do not assume a skew-product structure and add continuity to the conclusion.

LEMMA 6.1. *Assume that T is continuous and that there is a continuous section σ. Fix $\mu \in \mathcal{P}_T(X)$, $\check{\mu} = \pi_*\mu$ and let $(\xi_y)_y$ be the disintegration of μ with respect to π and $\check{\mathcal{L}}$ be the transfer operator of $(S, \check{\mu})$.*

Then, for all $y \in Y$ and all continuous $f : X \to \mathbb{R}$, we have $\xi_y(f) = \lim_n \check{\mathcal{L}}^n(f T^n \sigma)(y)$. If, moreover, $\check{\mathcal{L}}$ sends continuous functions to continuous functions, then $(\xi_y)_{y\in Y}$ preserves continuity.

Proof. For all $n, m \in \mathbb{N}$ and all continuous $f : X \to \mathbb{R}$, using $g = \check{\mathscr{L}}^m(g S^m)$, $\|\check{\mathscr{L}}(g)\|_\infty \leq \|g\|_\infty$ and $\pi T^m \sigma(y) = \pi \sigma S^m(y)$, we have

$$\|\check{\mathscr{L}}^{n+m}(f T^{n+m} \sigma) - \check{\mathscr{L}}^n(f T^n \sigma)\|_\infty = \|\check{\mathscr{L}}^{n+m}(f T^n \circ T^m \sigma) - \check{\mathscr{L}}^{n+m}(f T^n \circ \sigma S^m)\|_\infty$$
$$\leq \|f T^n \circ T^m \sigma - f T^n \circ \sigma S^m\|_\infty$$
$$\leq \omega(a_n)$$

where ω is a modulus of continuity of f and $(a_n)_n$ is a shrinking sequence.

It follows that $\zeta_y^n := \check{\mathscr{L}}^n(f T^n \sigma)(y)$ converges as $n \to \infty$, uniformly in $y \in Y$, and that the limit $\zeta_y := \lim_n \zeta_y^n$ defines for each $y \in Y$ a continuous linear form on $\mathscr{C}^0(X)$, that is, a measure on X.

If $\check{\mathscr{L}}$ sends continuous functions to continuous functions, then, for all $n \in \mathbb{N}$, the function $y \mapsto \zeta_y^n(f)$ is continuous, and by uniform convergence so is $y \mapsto \zeta_y(f)$.

All that remains is to check that $(\zeta_y)_{y \in Y}$ coincides with $(\xi_y)_{y \in Y}$ on a set of full $\check{\mu}$ measure. Since $\check{\mathscr{L}}^n(\mathbf{1} T^n \sigma) = \mathbf{1}$ and $\check{\mathscr{L}}^n(f T^n \sigma) \geq 0$ whenever $f \geq 0$, ζ_y is a probability measure for each y. If $f \equiv 0$ on $\pi^{-1}(y)$, then $f T^n \sigma \equiv 0$ on $S^{-n}(y)$ and $\check{\mathscr{L}}^n(f T^n \sigma)(y) = 0$, so that $\zeta_y(f) = 0$; that is, ζ_y is concentrated on $\pi^{-1}(y)$. Finally,

$$\int \zeta_y(f) \, d\check{\mu} = \lim_n \int \check{\mathscr{L}}^n(f T^n \sigma) \, d\check{\mu} = \lim_n \int f T^n \sigma \, d\check{\mu}$$
$$= \int f \, d\left(\lim_n T_*^n(\sigma_* \check{\mu})\right) = \int f \, d\mu$$

and, by uniqueness in the disintegration theorem, $\zeta_y = \xi_y$ for $\check{\mu}$-almost all $y \in Y$. $\square$

We shall now consider functions $f : X \to \mathbb{R}$ with a specified amount of regularity, that is, $f \in \mathrm{Hol}_\omega(X)$ for some modulus ω. We will need a stronger hypothesis on the transfer operator of S.

LEMMA 6.2. *Assume that $\check{\mathscr{L}}$ is iteratively bounded with respect to ω. Then, for all $f \in \mathrm{Hol}_\omega(X)$, there is a version of $\xi(f)$ such that, for all $y, y' \in Y$ and all $k, n \in \mathbb{N}$,*

$$|\xi_y(f T^k) - \xi_{y'}(f T^k)| \leq 2 \mathrm{Hol}_\omega(f) \omega(a_{n+k}) + C \mathrm{Hol}_\omega(f T^{n+k} \sigma) \omega(d(y, y')).$$

Proof. Set, as above, $\zeta_y^n(f) = \check{\mathscr{L}}^n(f T^n \sigma)(y)$ for all $y \in Y$. Then $(\zeta_y^n)_{y \in Y}$ is the disintegration of $T_*^n(\sigma_* \check{\mu})$, while $(\xi_y)_y$ is the disintegration of $\mu = T_*^n \mu$. There exists $\gamma \in \Gamma_\pi(\mu, \sigma_* \check{\mu})$ (actually γ is unique, equal to $(\mathrm{Id}, \sigma\pi)_* \mu$), and $\gamma^n := (T^n, T^n)_* \gamma$ is in $\Gamma_\pi(\mu, T_*^n(\sigma_* \check{\mu}))$. Let $(\eta_y^n)_{y \in Y}$ be the disintegration of γ^n with respect to the map $\Delta_\pi \to Y$ sending (x, x') to $\pi(x) = \pi(x')$. Then, for $\check{\mu}$-almost all y, $\eta_y^n \in \Gamma(\xi_y, \zeta_y^n)$ and for η_y^n-almost all (x, x'), we have $x = T^n(w)$ and $x' = T^n(w')$ for some w, w' with $\pi(w) = \pi(w')$. Now we get

$$|\xi_y(f T^k) - \zeta_y^n(f T^k)| = \left| \int f T^k(x) \, d\eta_y^n(x, x') - \int f T^k(x') \, d\eta_y^n(x, x') \right|$$
$$= \left| \int (f(x) - f(x')) \, d((T^k, T^k)_* \eta_y^n)(x, x') \right|$$
$$\leq \mathrm{Hol}_\omega(f) \int \omega(d(x, x')) \, d((T^k, T^k)_* \eta_y^n)(x, x')$$
$$\leq \mathrm{Hol}_\omega(f) \omega(a_{n+k})$$

since, for $(T^k, T^k)_* \eta_y^n$-almost all (x, x'), $x = T^{k+n}(w)$ and $x' = T^{n+k}(w')$ for some w, w' in the same fiber. We then have

$$|\xi_y(fT^k) - \xi_{y'}(fT^k)| \leq |\xi_y(fT^k) - \zeta_y^n(fT^k)| + |\zeta_y^n(fT^k) - \zeta_{y'}^n(fT^k)|$$
$$+ |\zeta_{y'}^n(fT^k) - \xi_{y'}(fT^k)|$$
$$\leq 2\operatorname{Hol}_\omega(f)\omega(a_{n+k}) + |\check{\mathscr{L}}^n(fT^{n+k}\sigma)(y) - \check{\mathscr{L}}^n(fT^{n+k}\sigma)(y')|$$
$$\leq 2\operatorname{Hol}_\omega(f)\omega(a_{n+k}) + C\operatorname{Hol}_\omega(fT^{n+k}\sigma)\omega(d(y, y')). \qquad \square$$

THEOREM 6.3. (Regularity of disintegrations) *Let T be a L-Lipschitz extension of S with $(a_n)_n$-shrinking fibers, and assume that there is a Lipschitz section σ. Let $\alpha \in (0, 1]$ and let $\mu \in \mathcal{P}_T(X)$, $\check{\mu} = \pi_* \mu$.*

(i)   *If the transfer operator $\check{\mathscr{L}}$ of $(S, \check{\mu})$ is iteratively bounded with respect to ω_α and if fibers are exponentially shrinking with ratio $\theta \in (0, 1)$, then the disintegration ξ of μ with respect to π is $(\omega_\alpha, \omega_\beta)$-bounded for $\beta = \alpha/(1 - \log L/\log\theta)$.*

(ii)   *If the transfer operator $\check{\mathscr{L}}$ of $(S, \check{\mu})$ is iteratively bounded with respect to ω_α and if $(a_n)_n$ is polynomial of degree d, then $(\xi_y)_{y \in Y}$ is $(\omega_\alpha, \omega_{\alpha d \log})$-bounded and, moreover, the maps*

$$\mathcal{D}_k : \operatorname{Hol}_\alpha(X) \to \operatorname{Hol}_{\alpha d \log}(Y)$$
$$f \mapsto \xi(fT^k)$$

have operator norm bounded above by $Ck^{\alpha d}$.

(iii)   *If the transfer operator $\check{\mathscr{L}}$ of $(S, \check{\mu})$ is iteratively bounded with respect to $\omega_{\alpha \log}$ and if $(a_n)_n$ is exponential, then $(\xi_y)_{y \in Y}$ is $(\omega_{\alpha \log}, \omega_{(\alpha/2) \log})$-bounded.*

Note that the norm bound in (ii) is not a specific feature but will be needed later, as the trivial bound of $CL^{\alpha k}$ is much too weak in this case.

Note that Butterley and Melbourne [**BM17**, Proposition 6] obtain the better exponent $\beta = \alpha$ in item (i), but only in a restricted setting.

Proof. We apply Lemma 6.2 three times.

For item (i) we take $\omega = \omega_\alpha$, $k = 0$ and get, for all $f \in \operatorname{Hol}_\alpha(X)$, y, $y' \in Y$ and k, $n \in \mathbb{N}$,

$$|\xi_y(f) - \xi_{y'}(f)| \lesssim \operatorname{Hol}_\alpha(f)(\theta^{\alpha n} + L^{\alpha n}d(y, y')^\alpha).$$

Taking $n = \beta \log d(y, y')/(\alpha \log\theta)$ yields $\theta^{\alpha n} \simeq L^{\alpha n}d(y, y')^\alpha \simeq d(y, y')^\beta$.

For item (ii) we take $\omega = \omega_\alpha$ and consider arbitrary k to get the norm estimate. Given $y \neq y' \in Y$, we would like to choose $m = n + k = -c \log d(y, y') + O(1)$ for some small constant $c > 0$ to be specified later on. This is possible whenever $k \leq -c \log d(y, y')$; in this case we get $L^{\alpha m} \simeq d(y, y')^{-c\alpha \log L}$ and thus

$$|\xi_y(fT^k) - \xi_{y'}(fT^k)| \lesssim \operatorname{Hol}_\alpha(f)(\omega_{\alpha d \log}(d(y, y')) + d(y, y')^{\alpha(1 - c \log L)}).$$

Choosing $c < 1/\log L$ ensures the last term above is (much) less than $\omega_{\alpha d \log}(d(y, y'))$. We are left with the case $k > -c \log d(y, y')$, but then $1 \leq k/(c \log(1/d(y, y')))$ and

$$|\xi_y(f) - \xi_{y'}(f)| \leq \sup f - \inf f \lesssim \operatorname{Hol}_\alpha(f) \lesssim \operatorname{Hol}_\alpha(f)k^{\alpha d}\omega_{\alpha d \log}(d(y, y')).$$

For item (iii) we take $\omega = \boldsymbol{\omega}_{\alpha\log}$, $k = 0$ and get $\omega(a_n) \simeq n^{-\alpha}$. Given $y \neq y' \in Y$, we choose $n \simeq \boldsymbol{\omega}_{(1/2)\log}(d(y, y')) \simeq (-\log d(y, y'))^{1/2}$ so that by Proposition 2.3,

$$\mathrm{Hol}_{\alpha\log}(f T^n \sigma) \lesssim (-\log d(y, y'))^{\alpha/2} \, \mathrm{Hol}_{\alpha\log}(f);$$

then Lemma 6.2 yields $|\xi_y(f) - \xi_{y'}(f)| \lesssim \mathrm{Hol}_{\alpha\log}(f) \, \boldsymbol{\omega}_{(\alpha/2)\log}(d(y, y'))$ as desired. $\qquad\square$

6.2. *Lifting decay of correlations.*   Combining Lemma 4.4 with Theorem 6.3, we can finally lift decay of correlations.

Proof of Theorem D.   For item (i) we start from $f \in \mathrm{Hol}_\alpha(X)$ of zero μ-average and $g \in L^1(\mu)$; then $f \in \mathrm{Hol}_{\alpha'}(X)$ for all $\alpha' \leq \alpha$, with $\mathrm{Hol}_{\alpha'}(f) \leq \mathrm{Hol}_\alpha(f)$. Up to choosing a smaller α, the hypothesis on the transfer operator enables us to assume that $\check{\mathscr{L}}$ has a spectral gap on $\mathrm{Hol}_\alpha(Y)$, and is thus $\boldsymbol{\omega}_\alpha$-iteratively bounded. From Theorem 6.3, we get that $\xi(f T^k)$ is in $\mathrm{Hol}_\beta(Y)$ for some $\beta \in (0, \alpha)$, with

$$\mathrm{Hol}_\beta(\xi(f T^k)) \lesssim \mathrm{Hol}_\alpha(f T^k) \lesssim L^{\alpha k} \, \mathrm{Hol}_\alpha(f)$$

where L is a Lipschitz constant for T. Lemma 4.4 then yields for all $k, m \in \mathbb{N}$:

$$\mathcal{C}_\mu^{k+m}(f, g) \lesssim (1 - \delta)^m L^{\alpha k} \, \mathrm{Hol}_\alpha(f) \|\xi(g)\|_{L^1(\check{\mu})} + \theta^{\alpha k} \, \mathrm{Hol}_\alpha(f) \|g\|_{L^1(\mu)}$$

where δ is the spectral gap of $\check{\mathscr{L}}$ and θ is the ratio of shrinking (recall $\|\xi(g)\|_{L^1(\check{\mu})} = \|g\|_{L^1(\mu)}$). Taking sequences $k_n = tn + O(1)$ and $m_n = (1 - t)n + O(1)$ summing to n with $t \in (0, 1)$ small enough provides exponential decay of $\mathcal{C}_\mu^n(f, g)$.

For (ii) we start again from $f \in \mathrm{Hol}_\alpha(X)$ of zero μ-average and $g \in L^1(\mu)$; by hypothesis $\check{\mathscr{L}}$ is iteratively $\boldsymbol{\omega}_\alpha$-bounded and $\check{\mu}$ has polynomial decay of correlations of degree p for αd log-Hölder observables. Theorem 6.3 ensures that $\xi(f T^k)$ is in $\mathrm{Hol}_{\alpha d \log}(Y)$ with norm at most $C k^{\alpha d}$. Then Lemma 4.4 yields, for all $k, m \in \mathbb{N}$,

$$\mathcal{C}_\mu^{k+m}(f, g) \lesssim \mathrm{Hol}_\alpha(f) \|g\|_{L^1(\mu)} \left(\frac{k^{\alpha d}}{m^p} + \frac{1}{k^{\alpha d}} \right).$$

Given n, to optimize over pairs (k, m) such that $k + m = n$, one is led to make both terms of the same order of magnitude, that is, to take $k \simeq (n - k)^{p/2\alpha d}$. If $p > 2\alpha d$, then $m \ll k$ and thus $k \simeq n$, and we get a polynomial decay of degree αd. If $p < 2\alpha d$, then $k \ll m$ and thus $m \simeq n$, $k \simeq m^{p/2\alpha d}$ and we get a polynomial decay of degree $p/2$. If $p = 2\alpha d$, we take $k \simeq m$ both of the same order as n and we get a polynomial decay of correlations of degree $\alpha d = p/2$.

For item (iii) we start from $f \in \mathrm{Hol}_{\alpha\log}(X)$ of zero μ-average and $g \in L^1(\mu)$; by hypothesis $\check{\mathscr{L}}$ is iteratively $\boldsymbol{\omega}_{\alpha\log}$-bounded and $\check{\mu}$ has polynomial decay of correlations of degree p in $\mathrm{Hol}_{(\alpha/2)\log}(Y)$. Theorem 6.3 ensures that $\xi(f T^k)$ is in $\mathrm{Hol}_{\alpha/2}(Y)$ with norm at most $C \, \mathrm{Hol}_{\alpha\log}(f T^k) \lesssim k^\alpha \, \mathrm{Hol}_{\alpha\log}(f)$ (Proposition 2.3). Then Lemma 4.4 yields, for all $k, m \in \mathbb{N}$,

$$\mathcal{C}_\mu^{k+m}(f, g) \lesssim \mathrm{Hol}_{\alpha\log}(f) \|g\|_{L^1(\mu)} \left(\frac{k^\alpha}{m^p} + \frac{1}{k^\alpha} \right)$$

and, as above, we get $\mathcal{C}_\mu^n(f, g) \lesssim \mathrm{Hol}_{\alpha\log}(f) \|g\|_{L^1(\mu)} / n^{\min(\alpha, p/2)}$. $\qquad\square$

7. *Proofs of corollaries given in the introduction*

Proof of Corollary 1.7. To prove the first part of item (i), observe that the unique absolutely continuous invariant measure $\check{\mu}_S$ of S_q is ergodic and has full support [**Tha80**], and it must thus be the unique physical measure; then Theorem A implies that the unique lift μ_T of $\check{\mu}_S$ is the unique physical measure of T. To obtain the ASIP, consider an α-Hölder observable f. By Corollary 5.3, for some $\gamma > 0$, there exists a γ-Hölder observable $\check{f} : \mathbb{T} \to \mathbb{R}$ such that $\hat{f} := \check{f} \circ \pi$ differs from f by a coboundary. Then as in Proposition 5.4, the ASIP for (T, μ_T, f) follows from the ASIP for $(S_q, \check{\mu}, \check{f})$, which has been proved by Melbourne and Nicol [**MN05**] (the return time $\mathcal{R}$ of S_q to $[\frac{1}{2}, 1]$ is n on an interval of size $\simeq 1/n^{1+1/q}$, so that when $q < \frac{1}{2}$ we have $\mathcal{R}$ in $L^{2+\delta}$ for some $\delta > 0$). Note that ASIPs with better rates have been obtained recently by Cuny, Dedecker, Korepanov and Merlevède for intermittent maps [**CDKM18**]; Theorem 5.5 enables us to lift these rates.

For item (iii), [**Klo17**, Theorem A] states that S_q satisfies UE($\boldsymbol{\omega}_\gamma$, CLT, $\boldsymbol{\omega}_*$) for all $\gamma > q$ and that the transfer operators of γ-Hölder potentials have a spectral gap on Hölder spaces of small enough exponent (use [**TK05**] for the CLT; see the comment below Corollary F in [**Klo17**]). Theorem C, item (i) (where $\beta = 1$: π is Lipschitz) then ensures that T satisfies UE($\boldsymbol{\omega}_\alpha$, CLT, $\boldsymbol{\omega}_*$) for all $\alpha > q'$. The decay of correlation follows from Theorem D.   □

Proof of Corollary 1.8. Setting $\alpha' = \alpha d - 1 > 1$, [**Klo17**, Theorem E] shows that:

(i)   the transfer operator associated to a α' log-Hölder potential $\check{\varphi}$, defined by

$$\check{\mathscr{L}}_{\check{\varphi}} f(z) = \sum_{z \in S^{-1}(y)} e^{\varphi(y)} f(y),$$

acts on $(\alpha' - 1)$ log-Hölder observables; it can be normalized, that is, up to adding to $\check{\varphi}$ a constant and a coboundary, one can assume $\check{\mathscr{L}}_{\check{\varphi}} \mathbf{1} = \mathbf{1}$; and once normalized there is a unique S-invariant probability measure $\check{\mu}_{\check{\varphi}}$ that is also fixed by the dual operator $\check{\mathscr{L}}_{\check{\varphi}}^*$;

(ii) the transfer operator decays polynomially with degree $\alpha' - 1$ in the uniform norm for all $u \in \mathrm{Hol}_{(\alpha'-1)\log}(Y)$ such that $\check{\mu}_{\check{\varphi}}(u) = 0$, that is, $\|\check{\mathscr{L}}_{\check{\varphi}}^n u\|_\infty \lesssim \mathrm{Hol}_{(\alpha'-1)\log}(u)/n^{\alpha'-1}$;

(iii) when $\alpha' > 3/2$, using [**TK05**] as above, $\check{\mu}_{\check{\varphi}}$ satisfies the CLT for all $(\alpha' - 1)$ log-Hölder observables.

While it is not stated in [**Klo17**], $\check{\mu}_{\check{\varphi}}$ is the unique equilibrium state for $\check{\varphi}$ (see Ledrappier [**Led74**] and Walters [**Wal75**]; the statements there are for one-sided subshifts of finite type, but the assumption really used is the existence of a one-sided generator, which holds here), so that S satisfies UE($\boldsymbol{\omega}_{\alpha'\log}, \varnothing$), and when $\alpha' > 3/2$ it also satisfies UE($\boldsymbol{\omega}_{\alpha'\log}$, CLT, $\boldsymbol{\omega}_{(\alpha'-1)\log}$). Theorem C enables us to deduce for T both UE($\boldsymbol{\omega}_\alpha; \varnothing$) when $\alpha > 2/d$, and UE($\boldsymbol{\omega}_\alpha$; CLT, $\boldsymbol{\omega}_{\alpha-1/d}$) when $\alpha > 5/(2d)$ (take $\gamma = \alpha - 1/d$, so that $\gamma' = \alpha' - 1$).

Since the transfer operator of a uniformly expanding map with respect to the equilibrium state of a Hölder potential is well known to have a spectral gap (and thus is iteratively $\boldsymbol{\omega}_\gamma$-bounded), Theorem D, item (ii) applies (with the exponent $\gamma = \alpha - 1/d$ instead

of α, and $p = \alpha d - 2$), implying a polynomial rate of decay of correlations of degree $\alpha d / 2 - 1$. □

Proof of Corollary 1.9. By [**Klo17**] (see also [**FJ01b**, **FJ01a**]), S has a unique absolutely continuous measure $\check{\mu}$, which has polynomial decay of correlations of degree $(\alpha - 1)$ for all $(\alpha - 1)$ log-Hölder observables (in particular, it is ergodic). Let μ_T be the unique T-invariant lift of $\check{\mu}$ provided by Theorem A: then μ_T is physical, and since $\check{\mu}$ is also the unique physical measure of S, T admits no other physical measure. Better still, its basin of attraction is the inverse image by π of the basin of attraction of $\check{\mu}$ (Corollary 3.5), thus by Fubini's theorem it has full volume.

Moreover, in [**Klo17**] it is shown that the transfer operator of S for the geometric potential $\varphi_S(y) = -\log \det(DS_y)$ (or any other α log-Hölder potential) has polynomial decay of degree $\gamma = \alpha - 1$ in the uniform norm for all γ-log Hölder observables (implying the central limit theorem as soon as $\alpha > 3/2$). If T has exponentially shrinking fibers, then we can apply the last item of Theorem D to $(2\alpha - 2)$ log-Hölder observables with $p = \gamma = \alpha - 1$ to obtain the desired decay of correlation, and the last item of Theorem C to get the central limit theorem for $(2\alpha - 1)$ log-Hölder observables. □

Proof of Corollary 1.11. We construct T as a Smale DE ('derived from expanding') example [**Sma67**]. As their name indicates, DE examples start from an expanding map of a manifold; we will take $S : \mathbb{T} \to \mathbb{T}$ to be a uniformly expanding circle map of class $\mathscr{C}^1$ (since we start from a one-dimensional base map, this kind of example can be called a 'solenoidal' example: the attractor will topologically be a solenoid). For some $\lambda > 1$, we have $S'(y) \geq \lambda$ for all $y \in \mathbb{T}$; we then take a skew-product

$$T : \mathbb{T} \times D^2 \to \mathbb{T} \times D^2$$
$$(y, z) \mapsto (S(y), R(y, z))$$

where D^2 is the open unit disk of $\mathbb{R}^2$, and R is smooth and chosen so that:
- T is a diffeomorphism onto its image (i.e. for all $y \in \mathbb{T}$, $R(y, \cdot) : D^2 \to D^2$ is a diffeomorphism onto its image and whenever $y, y' \in \mathbb{T}$ have the same image under S, $R(y \cdot)$ and $R(y', \cdot)$ are disjoint);
- we assume that whenever $S(y) = S(y')$, the images of $R(y, \cdot)$ and $R(y', \cdot)$ have disjoint closures (in particular, the closure in $\mathbb{T} \times D^2$ of the image $\mathrm{Im}(T)$ is compact);
- $\|D_z R(y, z)\| \leq \lambda^{-1}$ (in particular, T is an extension of S with exponentially shrinking fibers).

We identify $\mathbb{T} \times D^2$ with an open subset U of $\mathbb{R}^3$ (e.g. a solid torus of revolution, with the angle of cylindrical coordinates corresponding to the y variable). By assumption, $\Lambda = \bigcap_n T^n(U)$ is a compact attractor, and one can easily check that the restriction of the projection π to Λ is still onto $\mathbb{T}$; we denote this restriction by the same letter π.

We first check that Λ is uniformly hyperbolic (this argument is classical and can be skipped by the experienced reader). The stable bundle is trivially constructed over the whole of U as $E_x^s = \{0\} \times T_z D^2$ (where $x = (y, z)$), and the main point is to find a transversal bundle E^u that is T-invariant. We consider the space of all continuous one-dimensional sub-bundles $E \subset T_\Lambda U$ transversal to E^s; such a bundle is parametrized by

a field $(L^E_x)_{x=(y,z)\in\Lambda}$ of linear maps $T_y\mathbb{T} \to T_z D^2$, simply setting $E_x = \{(u, L_x(u)) \in T_x U : u \in T_y\mathbb{T}\}$ (i.e. $L_x(u)$ is the unique $v \in T_z D^2$ such that $u + v \in E_x$), and we obtain a complete metric by using the operator norm: $d(E, F) = \max_x \|L^E_x - L^F_x\|$. Now the facts that S is at least λ-expanding and that $\|D_z R\| \leq \lambda^{-1}$ ensure that T acts on this space of bundles as a contraction: writing $x' = (y', z') = T^{-1}(x)$, the definition $(T_*E)_x = DT_{x'}(E_{x'})$ translates as $(T_*L)_x(u) = D_z R_{x'} \circ L_{x'}(DS^{-1}_{x'}(u))$, so that $d(T_*E, T_*F) \leq \lambda^{-2} d(E, F)$. There is thus a unique T-invariant continuous sub-bundle transverse to E^s. Up to changing the Riemannian metric, we can make it coincide on E^u with the pull-back of the metric of $\mathbb{T}$; then DT is at least λ-expanding along E^u in this metric, so that Λ is uniformly hyperbolic.

Now the usual theory ensures we have an unstable lamination W^u of Λ (the stable foliation is trivial, its connected components of leaves being the vertical slices $\{y\} \times D^2$), and the definition of an SRB makes sense. We shall use the following lemma.

LEMMA 7.1. *If μ is an SRB measure of $T = (S, R)$ a uniformly hyperbolic skew product, then the projection $\check{\mu}$ of μ to the first factor is absolutely continuous.*

Proof. Consider a small open set $V \subset \mathbb{T}$, and partition $\pi^{-1}(V) \subset \Lambda$ into small enough subsets $V_1, \ldots, V_k$ such that each V_k is given a product structure by W^s, W^u. For each $i \in \{1, \ldots, k\}$, let $\mu^i = \mu_{|V_i}$ and write its disintegration with respect to the projection on the stable direction as $\mu^i = \int \mu^i_L \, d\nu^i(L)$. If μ is SRB, there are positive integrable functions f^i_L such that $d\mu^i_L = f^i_L \, d\mathrm{Vol}$ where $d\mathrm{Vol}$ is the volume (i.e. Lebesgue measure) on $\mathbb{T}$. Then, for any continuous $g : \mathbb{T} \to \mathbb{R}$,

$$\check{\mu}_{|V}(g) = \sum_{i=1}^k \iint g(y) f^i_L(y) \, d\mathrm{Vol}(y) \, d\nu^i(L) = \int g(y) \left(\sum_{i=1}^k \int f^i_L(y) \, d\nu^i(L) \right) d\mathrm{Vol}(y)$$

is absolutely continuous.   □

Note that we did not use invariance of μ and that the converse of this lemma is not obvious: there are (not necessarily invariant) measures that project to the Lebesgue measure without having absolutely continuous disintegrations.

The work of Campbell and Quas [**CQ01**] shows that, taking S generic, we can assume it has a unique physical measure $\check{\mu}$, with full basin of attraction, but singular with respect to $d\mathrm{Vol}$ (and thus S has no Acip). Then its lift μ is a T-invariant measure that is physical, with full basin of attraction. Moreover, T has no SRB measure, since it would have an absolutely continuous projection.   □

REFERENCES

[**ABV00**]   J. F. Alves, C. Bonatti and M. Viana. SRB measures for partially hyperbolic systems whose central direction is mostly expanding. *Invent. Math.* **140**(2) (2000), 351–398.

[**ADLP17**]   J. F. Alves, C. L. Dias, S. Luzzatto and V. Pinheiro. SRB measures for partially hyperbolic systems whose central direction is weakly expanding. *J. Eur. Math. Soc. (JEMS)* **19**(10) (2017), 2911–2946.

[AGP14]   V. Araújo, S. Galatolo and M. J. Pacifico. Decay of correlations for maps with uniformly contracting fibers and logarithm law for singular hyperbolic attractors. *Math. Z.* **276**(3–4) (2014), 1001–1048.

[AGY06]   A. Avila, S. Gouëzel and J.-C. Yoccoz. Exponential mixing for the Teichmüller flow. *Publ. Math. Inst. Hautes Études Sci.* **104** (2006), 143–211.

[Alv15]   J. F. Alves. SRB measures for partially hyperbolic attractors, 2015, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.709.4072&rep=rep1&type=pdf.

[AMV15]   V. Araújo, I. Melbourne and P. Varandas. Rapid mixing for the Lorenz attractor and statistical limit laws for their time-1 maps. *Comm. Math. Phys.* **340**(3) (2015), 901–938.

[APPV09]   V. Araujo, M. J. Pacifico, E. R. Pujals and M. Viana. Singular-hyperbolic attractors are chaotic. *Trans. Amer. Math. Soc.* **361**(5) (2009), 2431–2485.

[AV09]   F. Abdenur and M. Viana. Flavors of partial hyperbolicity, 2009, http://w3.impa.br/~viana/out/flavors.pdf.

[BM17]   O. Butterley and I. Melbourne. Disintegration of invariant measures for hyperbolic skew products. *Israel J. Math.* **219**(1) (2017), 171–188.

[Bow08]   R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms (Lecture Notes in Mathematics, 470)*, revised edn. Springer, Berlin, 2008. With a preface by David Ruelle, edited by Jean-René Chazottes.

[CDKM18]   C. Cuny, J. Dedecker, A. Korepanov and F. Merlevède. Rates in almost sure invariance principle for slowly mixing dynamical systems. *Ergod. Th. & Dynam. Sys.* doi:10.1017/etds.2019.2. Published online 26 February 2019.

[CE11]   E. Catsigeras and H. Enrich. SRB-like measures for C^0 dynamics. *Bull. Pol. Acad. Sci. Math.* **59**(2) (2011), 151–164.

[CN17]   A. Castro and T. Nascimento. Statistical properties of the maximal entropy measure for partially hyperbolic attractors. *Ergod. Th. & Dynam. Sys.* **37**(4) (2017), 1060–1101.

[CP15]   S. Crovisier and R. Potrie. Introduction to partially hyperbolic dynamics. *Lecture Notes for School on Dynamical Systems (ICTP, Trieste)*, 3, 2015.

[CPZ18]   V. Climenhaga, Y. Pesin and A. Zelerowicz. Equilibrium states in dynamical systems via geometric measure theory. *Bull. Amer. Math. Soc.* **56** (2019), 569–610.

[CQ01]   J. T. Campbell and A. N. Quas. A generic C^1 expanding map has a singular S-R-B measure. *Comm. Math. Phys.* **221**(2) (2001), 335–349.

[CV13]   A. Castro and P. Varandas. Equilibrium states for non-uniformly expanding maps: decay of correlations and strong stability. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **30**(2) (2013), 225–249.

[DHRS09]   L. J. Díaz, V. Horita, I. Rios and M. Sambarino. Destroying horseshoes via heterodimensional cycles: generating bifurcations inside homoclinic classes. *Ergod. Th. & Dynam. Sys.* **29**(2) (2009), 433–474.

[FJ01a]   Fan A. and Y. Jiang. On Ruelle–Perron–Frobenius operators. I. Ruelle theorem. *Comm. Math. Phys.* **223**(1) (2001), 125–141.

[FJ01b]   Fan A. and Y. Jiang. On Ruelle–Perron–Frobenius operators. II. Convergence speeds. *Comm. Math. Phys.* **223**(1) (2001), 143–159.

[Gal18]   S. Galatolo. Quantitative statistical stability and speed of convergence to equilibrium for partially hyperbolic skew products. *J. Éc. Polytech. Math.* **5** (2018), 377–405.

[GL15]   S. Galatolo and R. Lucena. Spectral gap and quantitative statistical stability for systems with contracting fibers and Lorenz like maps. *Discrete Contin. Dyn. Syst.* **40**(3) (2020), 1309–1360.

[GNP18]   S. Galatolo, I. Nisoli and M. J. Pacifico. Decay of correlations, quantitative recurrence and logarithm law for contracting Lorenz attractors. *J. Stat. Phys.* **170**(5) (2018), 862–882.

[Gou10]   S. Gouëzel. Almost sure invariance principle for dynamical systems by spectral methods. *Ann. Probab.* **38**(4) (2010), 1639–1671.

[GP10]   S. Galatolo and M. J. Pacifico. Lorenz-like flows: exponential decay of correlations for the Poincaré map, logarithm law, quantitative recurrence. *Ergod. Th. & Dynam. Sys.* **30**(6) (2010), 1703–1737.

[Gui15]   P.-A. Guihéneuf. Dynamical properties of spatial discretizations of a generic homeomorphism. *Ergod. Th. & Dynam. Sys.* **35**(5) (2015), 1474–1523.

[Klo17]   B. Kloeckner. An optimal transportation approach to the decay of correlations for non-uniformly expanding maps. *Ergod. Th. & Dynam. Sys.* **40**(3) (2020), 714–750.

[KRN65]   K. Kuratowski and C. Ryll-Nardzewski. A general theorem on selectors. *Bull. Acad. Pol. Sci. Sér. Sci. Math. Astronom. Phys.* **13** (1965), 397–403.

[Led74]   F. Ledrappier. Principe variationnel et systèmes dynamiques symboliques. *Z. Wahrsch. Verw. Gebiete* **30**(3) (1974), 185–202.

[**Led84**] F. Ledrappier. Propriétés ergodiques des mesures de Sinai. *Publ. Math. Inst. Hautes Études Sci.*
         59(1) (1984), 163–188.

[**LOR11**] R. Leplaideur, K. Oliveira and I. Rios. Equilibrium states for partially hyperbolic horseshoes.
         Ergod. Th. & Dynam. Sys. **31**(1) (2011), 179–195.

[**LW77**] F. Ledrappier and P. Walters. A relativised variational principle for continuous transformations.
         J. Lond. Math. Soc. (2) **16**(3) (1977), 568–576.

[**MN05**] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems.
         Comm. Math. Phys. **260**(1) (2005), 131–146.

[**MT02**] I. Melbourne and A. Török. Central limit theorems and invariance principles for time-one maps of
         hyperbolic flows. *Comm. Math. Phys.* **229**(1) (2002), 57–71.

[**Roh52**] V. A. Rohlin. On the fundamental ideas of measure theory. *Amer. Math. Soc. Trans.* **1952**(71)
         (1952), 55.

[**RS17**] V. Ramos and J. Siqueira. On equilibrium states for partially hyperbolic horseshoes: uniqueness
         and statistical properties. *Bull. Braz. Math. Soc. (N.S.)* **48**(3) (2017), 347–375.

[**Sim12**] D. Simmons. Conditional measures and conditional expectation; Rohlin's disintegration theorem.
         Discrete Contin. Dyn. Syst. **32**(7) (2012), 2565–2582.

[**Sma67**] S. Smale. Differentiable dynamical systems. *Bull. Amer. Math. Soc.* **73** (1967), 747–817.

[**Tha80**] M. Thaler. Estimates of the invariant densities of endomorphisms with indifferent fixed points.
         Israel J. Math. **37**(4) (1980), 303–314.

[**TK05**] M. Tyran-Kamińska. An invariance principle for maps with polynomial decay of correlations.
         Comm. Math. Phys. **260**(1) (2005), 1–15.

[**Vil09**] C. Villani. *Optimal Transport (Grundlehren der Mathematischen Wissenschaften [Fundamental
         Principles of Mathematical Sciences], 338).* Springer, Berlin, 2009. .

[**Wal75**] P. Walters. Ruelle's operator theorem and *g*-measures. *Trans. Amer. Math. Soc.* **214** (1975),
         375–387.

[**Wal82**] P. Walters. *An Introduction to Ergodic Theory (Graduate Texts in Mathematics, 79).* Springer, New
         York, 1982.

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

[**You02**] L.-S. Young. What are SRB measures, and which dynamical systems have them? *J. Stat. Phys.*
         108(5–6) (2002), 733–754.