FROM RATES OF MIXING TO RECURRENCE TIMES VIA LARGE DEVIATIONS

JOSÉ F. ALVES, JORGE M. FREITAS, STEFANO LUZZATTO, AND SANDRO VAIENTI

Abstract. A classic approach in dynamical systems is to use particular geometric structures to deduce statistical properties, for example the existence of invariant measures with stochastic-like behaviour such as large deviations or decay of correlations. Such geometric structures are generally highly non-trivial and thus a natural question is the extent to which this approach can be applied. In this paper we show that in many cases stochastic-like behaviour itself implies that the system has certain non-trivial geometric properties, which are therefore necessary as well as sufficient conditions for the occurrence of the statistical properties under consideration. As a by product of our techniques we also obtain some new results on large deviations for certain classes of systems which include Viana maps and multidimensional piecewise expanding maps.

1. Introduction and statement of results

Let $f\colon M \to M$ be a piecewise C^{1+} endomorphism defined on a Riemannian manifold M, and let d denote the distance in M and m a normalized volume form on the Borel sets of M that we call *Lebesgue measure.* Here C^{1+} denotes the class of continuously differentiable maps with Hölder continuous derivative and the precise conditions on the "piecewise" will be stated below. A basic problem is the study of the statistical properties of the map f, starting from questions about the existence of an ergodic invariant measure μ which is absolutely continuous with respect to Lebesgue to more sophisticated properties such as the rate of decay of correlations or large deviations with respect to this measure μ. In a fundamental paper [Yo2], Young showed that the existence of such a measure μ and, more significantly, the rate of decay of correlations of μ can be deduced from the "geometry" of f, more specifically from the existence and properties of a "Young tower" or "induced Gibbs-Markov map". The verification of this geometric structure is of course generally highly non-trivial, and over the last ten years a substantial number of papers have been devoted to this goal under various kinds of assumptions and using a variety of techniques [Yo1, Yo2, BLS, ALP, Go, Hol, DHL]. Combining these geometric constructions with the abstract results of Young, and more recent results concerning also other statistical properties such as large deviations [MN, RY], much more significant progress has been made in understanding the stochastic-like behaviour of deterministic dynamical systems

2000 *Mathematics Subject Classification.* 37A05, 37C40, 37D25.

Key words and phrases. Gibbs-Markov structure, decay of correlations, large deviations.

Work carried out at CIRM, ICTP, Imperial College and University of Porto. JFA and JMF were partially supported by FCT through CMUP, by POCI/MAT/61237/2004 and by PTDC/MAT/099493/2008. JMF was partially supported by FCT grant SFRH/BPD/66040/2009.

in the last ten years than had been since the pioneering results on uniformly hyperbolic systems in the 60's and early 70's.

A natural question concerns the limitations of this approach. Might there be large classes of systems, or even specific "pathological" systems, that exhibit certain statistical properties but for which this approach does not and cannot work because such systems just do not admit the required geometrical structures? The main purpose of this paper is to show that in many cases such systems do not exist, and that in fact stochastic-like behaviour such as decay of correlations at certain rates is in itself sufficient to imply the existence of an induced Gibbs-Markov map with the corresponding properties. This geometry is therefore both necessary and sufficient for the statistical properties of the system. We will now give the precise formulation of these results.

1.1. **Main definitions.** We start with the definition of a Gibbs-Markov structure and then give the formal definitions of the notion of decay of correlations and large deviations.

Definition 1.1. We say that f admits a *Gibbs-Markov induced map* if there exists a ball $\Delta_0 \subset M$, a countable partition $\mathcal{P}$ (mod 0) of Δ_0 into topological balls U with smooth boundaries, and a return time function $R : \Delta_0 \to \mathbb{N}$ constant on elements of $\mathcal{P}$ satisfying the following properties:

(1) *Markov:* for each $U \in \mathcal{P}$ and $R = R(U)$, $f^R : U \to \Delta_0$ is a C^{1+} diffeomorphism (and in particular a bijection). Thus the induced map $F : \Delta_0 \to \Delta_0$ given by $F(x) = f^{R(x)}(x)$ is defined almost everywhere and satisfies the classical Markov property.
(2) *Uniform expansion*: there exists $\lambda < 1$ such that for almost all $x \in \Delta_0$ we have $\|DF(x)^{-1}\| \le \lambda$. In particular the *separation time* $s(x, y)$ given by the maximum integer such that $F^i(x)$ and $F^i(y)$ belong to the same element of the partition $\mathcal{P}$ for all $i \le s(x, y)$, is defined and finite for almost all $x, y \in \Delta_0$.
(3) *Bounded distortion:* there exists $K > 0$ such that for any points $x, y \in \Delta_0$ with $s(x, y) < \infty$ we have
$$\left|\frac{\det DF(x)}{\det DF(y)} - 1\right| \le K\lambda^{-s(F(x),F(y))}.$$

We define that "tail" of the return time function at time n as the set
$$\mathscr{R}_n = \{x \in \Delta_0 : \; R(x) > n\}$$
of points whose return time is larger than n, and we say that the return time function is integrable if
$$\int R \, dm < \infty.$$

We remark that the usual definition of Gibbs-Markov does not require each subdomain U to map surjectively onto the entire domain Δ_0, our notion is therefore more restrictive and will allow us to deduce significantly stronger properties.

Definition 1.2 (Expanding measure)**.** We say that a measure μ is *regularly expanding* if

$$\log \|Df^{-1}\| \in L^1 \quad \text{ and } \quad \int \log \|Df^{-1}\| d\mu < 0.$$

A first example of the way in which geometric structure is related to statistical properties is given by the relation between the above two definitions. Indeed, it is shown in [ADL] that for large classes of maps including multidimensional maps with "non-degenerate" critical points the two structures are completely equivalent in the sense that f admits a Gibbs-Markov induced map if and only if it admits a regularly expanding absolutely continuous invariant probability measure. In this paper we develop these general philosophy further by considering more refined statistical properties.

Definition 1.3 (Decay of correlations)**.** Let $\mathcal{B}_1, \mathcal{B}_2$ denote Banach spaces of real valued measurable functions defined on M. We denote the *correlation* of non-zero functions $\varphi \in \mathcal{B}_1$ and $\psi \in \mathcal{B}_2$ with respect to a measure μ as

$$\mathrm{Cor}_\mu(\varphi, \psi \circ f^n) := \frac{1}{\|\varphi\|_{\mathcal{B}_1} \|\psi\|_{\mathcal{B}_2}} \left| \int \varphi \, (\psi \circ f^n) \, d\mu - \int \varphi \, d\mu \int \psi \, d\mu \right|.$$

We say that we have *decay of correlations*, with respect to the measure μ, for observables in $\mathcal{B}_1$ *against* observables in $\mathcal{B}_2$ if, for every $\varphi \in \mathcal{B}_1$ and every $\psi \in \mathcal{B}_2$ we have

$$\mathrm{Cor}_\mu(\varphi, \psi \circ f^n) \to 0, \quad \text{ as } n \to \infty.$$

We will use the notation $\lesssim$ to mean $\leq$ up to multiplication by a constant depending only on the map f. We say that the decay of correlations is *exponential*, *stretched exponential*, or *polynomial* if it is $\lesssim \varepsilon^{-\tau n}$, $\lesssim \varepsilon^{-\tau n^\theta}$ or $\lesssim n^{-\beta}$ respectively, for constants $\tau, \beta > 0$, $\theta \in (0,1)$ which depend only on f. Most of the time we shall choose $\mathcal{B}_2 = L^p$ for $p = 1$ or $p = \infty$, and $\mathcal{B}_1 = \mathcal{H}_\alpha$ the space of Hölder continuous functions with Hölder constant $\alpha > 0$. Recall that the Hölder norm of an observable $\varphi \in \mathcal{H}_\alpha$ is given by

$$\|\varphi\|_{\mathcal{H}_\alpha} := \|\varphi\|_\infty + \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{d(x,y)^\alpha}.$$

1.2. **Local diffeomorphisms.** We start by stating our results in the setting of C^{1+} local diffeomorphisms.

Theorem A. *Let $f : M \to M$ be a C^{1+} local diffeomorphism and let $\alpha > 0$. Suppose that f admits an ergodic expanding acip μ;*

(1) if there exists $\beta > 1$ such that $Cor_\mu(\varphi, \psi \circ f^n) \lesssim n^{-\beta}$ for every $\varphi \in \mathcal{H}_\alpha$ and $\psi \in L^\infty(\mu)$, then there is a Gibbs-Markov induced map with $m(\mathscr{R}_n) \lesssim n^{-\beta+1}$.

Suppose moreover that $d\mu/dm$ is uniformly bounded away from 0 on its support. Then

(2) if there exist $\tau, \theta > 0$ such that $Cor_\mu(\varphi, \psi \circ f^n) \lesssim e^{-\tau n^\theta}$ for every $\varphi \in \mathcal{H}_\alpha$ and $\psi \in L^\infty(\mu)$, then there is a Gibbs-Markov induced map with $m(\mathscr{R}_n) \lesssim e^{-\tau' n^{\theta'}}$ for some $\tau' > 0$ and $\theta' = \theta/(\theta + 2)$.

Combining these results with those of Young [Yo2] we conclude that the rate of decay of correlations is polynomial (resp. stretched exponential) if and only if there exists a Gibbs-Markov induced map with polynomial (resp. stretched exponential) tail. The specific exponents which appear are close to optimal but not quite optimal converses of those in [Yo2]. In the stretched exponential case Young shows that $m(\mathscr{R}_n) \lesssim e^{-\tau' n^{\theta'}}$ implies $\text{Cor}_\mu(\varphi, \psi \circ f^n) \lesssim e^{-\tau n^{\theta}}$ for some $\tau < \tau'$ and every $\theta < \theta'$, later improved in Gouëzel's PhD thesis where he proved that it is possible to take $\tau = \tau'$ and where he showed that this result is optimal. In the polynomial case Young shows that $m(\mathscr{R}_n) \lesssim n^{-\beta}$ implies $\text{Cor}_\mu(\varphi, \psi \circ f^n) \lesssim n^{-\beta+1}$ and Sarig [Sar] and Gouëzel [Go2] have shown that these results are optimal.

Question. *Is it possible to improve the estimates above so that $Cor_\mu(\varphi, \psi \circ f^n) \lesssim n^{-\beta+1}$ implies $m(\mathscr{R}_n) \lesssim n^{-\beta}$ and $Cor_\mu(\varphi, \psi \circ f^n) \lesssim e^{-\tau n^{\theta}}$ implies $m(\mathscr{R}_n) \lesssim e^{-\tau' n^{\theta}}$*

We remark also that the additional assumption on the density of μ for the (stretched) exponential case is due to the use of different technique for constructing the induced map, as we shall explain in more detail below. It holds in various known examples such as when the map is "locally eventually onto", i.e. every open set of positive μ measure covers the support of μ in a finite number of iterates.

We obtain similar results in the exponential case if we assume that the correlation decay is uniformly summable against all L^1 observables.

Theorem B. *Let $f : M \to M$ be a C^{1+} local diffeomorphism. Suppose that f admits an ergodic expanding acip μ with $d\mu/dm$ uniformly bounded away from 0 on its support. Suppose that there exists $\xi(n)$ with $\sum_{n=0}^{\infty} \xi(n) < \infty$ such that for all $\varphi \in \mathcal{H}_\alpha$ and $\psi \in L^1(\mu)$ we have $Cor_\mu(\varphi, \psi \circ f^n) \leq \xi(n)$. Then there exists a Gibbs-Markov induced map with $m(\mathscr{R}_n) \lesssim e^{-\tau' n}$ for some $\tau' > 0$. In particular exponential decay holds against L^∞ observables.*

Notice that from [Yo2] we get that exponential decay of the return time function implies exponential decay of correlation against L^∞. Theorem B thus implies that summable decay against L^1 implies exponential decay against L^∞.

Also in this case, the statement is not quite a direct converse of the results of [Yo2]. An "if and only if" statement could be obtained either by relaxing the assumptions on the decay of correlations against all L^1 functions in the Theorem, or by showing that this assumption actually holds whenever there is a Gibbs-Markov induced map with exponential tails.

Question. *Suppose there is a Gibbs-Markov induced map with $m(\mathscr{R}_n) \lesssim e^{-\tau' n}$ for some $\tau' > 0$. Is there $\xi(n)$ with $\sum_{n=0}^{\infty} \xi(n) < \infty$ such that $Cor_\mu(\varphi, \psi \circ f^n) \leq \xi(n)$ for every $\varphi \in \mathcal{H}_\alpha$ and $\psi \in L^1(\mu)$?*

We cannot presently give a fully positive answer to this question but we will show below that there are some quite general classes of systems which do exhibit in fact exponential decay against L^1 observables, see Corollary H.

1.3. **Maps with critical/singular sets.** The results stated above are for local diffeomorphisms, and are already relevant and non-trivial in that setting, but there exist many interesting examples which may fail to be local diffeomorphisms due to the presence of critical points (where $\det Df = 0$), singular points (where Df does not exist or $\|Df\| = \infty$) or discontinuities of f. We shall generally denote the collection of all such points as the critical/singular set. Most of the results which deduce statistical information from Gibbs-Markov maps apply equally to systems with a non-empty critical/singular set; in fact this is one of the strengths of this approach, the partition structure of Gibbs-Markov induced maps allows in some sense to avoid bad regions of the phase space. For the converse results, the situation is in principle more complicated because we need to show that a Gibbs-Markov map can still be constructed and that possible accumulation of images or preimages of the critical/singular set do not adversely affect the decay rates of tail of the return times. We shall show that in fact most of the results stated above do essentially apply under some mild assumption on the critical/singular set and on the density of the measure μ.

Definition 1.4. We say that x is a *critical point* if $Df(x)$ is not invertible and a *singular point* if $Df(x)$ does not exist. We let $\mathcal{C}$ denote the set of critical/singular points and let $d(x,\mathcal{C})$ denote the distance between the point $x \in M$ and the set $\mathcal{C}$. We say that a set $\mathcal{C}$ of critical/singular points is *non-degenerate* if there are constants $B, d > 0$ such that for all $\epsilon > 0$

(C0) $\quad m\left(\{x : d(x,\mathcal{C}) \le \epsilon\}\right) \le B\epsilon^d \quad$ (in particular $m(\mathcal{C}) = 0$);

and there exists $\eta > 0$ such that for every $x \in M \setminus \mathcal{C}$ and $v \in T_xM$ with $\|v\| = 1$ we have

(C1) $\quad B^{-1}d(x,\mathcal{C})^{\eta} \le \|Df(x)v\| \le Bd(x,\mathcal{C})^{-\eta}.$

Moreover, for all $x, y \in M \setminus \mathcal{C}$ we have

(C2) $\quad \left|\log\|Df(x)^{-1}\| - \log\|Df(y)^{-1}\|\,\right| \le B\,|\log(d(y,\mathcal{C})) - \log(d(x,\mathcal{C}))|;$

(C3) $\quad |\log|\det Df(x)| - \log|\det Df(y)|\,| \le B\,|\log(d(y,\mathcal{C})) - \log(d(x,\mathcal{C}))|.$

We remark that the conditions (C2) and (C3) imply the corresponding conditions used in [ABV, ALP, Go]. As long as the critical set satisfies the above mild non-degeneracy assumptions, we recover essentially the results stated above for local diffeomorphisms in the polynomial and stretched exponential case.

Theorem C. *Let $f : M \to M$ be a C^{1+} local diffeomorphism outside a nondegenerate critical set $\mathscr{C}$. Suppose that f admits an ergodic expanding acip μ with $d\mu/dm \in L^p(m)$ for some $p > 1$;*

(1) if there exists $\beta > 1$ such that $Cor_\mu(\varphi, \psi \circ f^n) \lesssim n^{-\beta}$ for every $\varphi \in \mathcal{H}_\alpha$ and $\psi \in L^\infty(\mu)$, then for any $\gamma > 0$ there is a Gibbs-Markov induced map such that $m(\mathscr{R}_n) \lesssim n^{-\beta+1+\gamma}$.

Suppose moreover that $d\mu/dm$ is uniformly bounded away from 0 on its support;

(2) if there exist $\tau, \theta > 0$ such that $Cor_\mu(\varphi, \psi \circ f^n) \lesssim e^{-\tau n^\theta}$ for every $\varphi \in \mathcal{H}_\alpha$ and $\psi \in L^\infty(\mu)$, then for any $\gamma > 0$ there is a Gibbs-Markov induced map such that $m(\mathscr{R}_n) \lesssim e^{-\tau' n^{\theta' - \gamma}}$ for $\theta' = \theta/(3\theta + 6)$.

Thus also in the very general setting of maps with critical and singular points we obtain a converse to Young's results and conclude that the rate of decay of correlations is polynomial (resp. stretched exponential) if and only if there exists a Gibbs-Markov induced map with polynomial (resp. stretched exponential) tail.

1.4. **Large deviations.** A key step in our argument is to show that the rate of decay of correlations implies certain large deviation estimates. This is itself a result of independent interest partly also because it is a completely abstract result and we use no additional structure on M or f other than $f : M \to M$ being measurable and *nonsingular* (see Section A.1) with respect to an ergodic probability measure μ on M. In particular, we need no Riemannian structure on M.

Definition 1.5 (Large deviations)**.** Given an ergodic probability measure μ and $\epsilon > 0$ we define the *large deviation* at time n of the time average of the observable φ from the spatial average as

$$\mathrm{LD}_\mu(\varphi, \epsilon, n) := \mu\left(\left|\frac{1}{n}\sum_{i=0}^{n-1}\varphi\circ f^i - \int\varphi d\mu\right| > \epsilon\right).$$

By Birkhoff's ergodic theorem the quantity $\mathrm{LD}_\mu(\varphi, \epsilon, n) \to 0$, as $n \to \infty$, and a relevant question also in this case is the rate of this decay.

Theorem D. *Let $f : M \to M$ preserve an ergodic probability measure μ with respect to which f is nonsingular. Let $\mathcal{B} \subset L^\infty(\mu)$ be a Banach space and $\varphi \in \mathcal{B}$.*

(1) [Me] *Let $\beta > 0$ and suppose that for all $\psi \in L^\infty(\mu)$ we have $Cor_\mu(\varphi, \psi \circ f^n) \lesssim n^{-\beta}$. Then, for every $\epsilon > 0$, there exists $C = C(\varphi, \epsilon) > 0$ such that $LD_\mu(\varphi, \epsilon, n) \le Cn^{-\beta}$.*

(2) Let $\theta, \tau > 0$ and suppose that for all $\psi \in L^\infty(\mu)$ we have $Cor_\mu(\varphi, \psi \circ f^n) \lesssim e^{-\tau n^\theta}$. Then, for every $\epsilon > 0$ there exist $C = C(\varphi, \epsilon) > 0$ and $\tau' = \tau'(\tau, \varphi, \epsilon) > 0$ such that $LD_\mu(\varphi, \epsilon, n) \le Ce^{-\tau' n^{\theta/(\theta+2)}}$.

The polynomial case has been proved in [Me], the stretched exponential case will be proved in Section 2.2 below. In both cases the explicit expressions for the constants, see Propositions 2.3 and 2.4, play a crucial role in the application of these estimates.

For exponential estimates we need to suppose that the decay of correlations is against L^1 observables.

Theorem E. *Let $f : M \to M$ preserve an ergodic probability measure μ with respect to which f is nonsingular. Let $\mathcal{B} \subset L^\infty(\mu)$ be a Banach space and $\varphi \in \mathcal{B}$. Suppose that there exists $\xi(n)$ with $\sum_{n=0}^\infty \xi(n) < \infty$ such that for all $\psi \in L^1(\mu)$ we have $Cor_\mu(\varphi, \psi\circ f^n) \le \xi(n)$. Then there exists $\tau = \tau(\varphi) > 0$ and, for every $\epsilon > 0$, there exists $C = C(\varphi, \epsilon) > 0$ such that $LD_\mu(\varphi, \epsilon, n) \le Ce^{-\tau n}$.*

1.5. **Further statistical properties.** A main motivation for the results stated above, is the remarkable fact that a statistical property such as decay of correlations can have such significant implications for the geometry of the system. On the other hand, one of the main original motivations for studying induced Gibbs-Markov maps is that they imply several statistical properties of the system. Thus combining the results stated above with recent results which have appeared in the last few years, we deduce that the decay of correlations implies several other statistical properties of great interest.

Corollary F. *Let $f : M \to M$ be a C^{1+} local diffeomorphism with an ergodic expanding acip μ, or a C^{1+} local diffeomorphism outside a nondegenerate critical set $\mathcal{C}$ with an ergodic expanding acip μ such that $d\mu/dm \in L^p(m)$, for some $p > 1$. Assuming that $Cor_\mu(\varphi, \psi \circ f^n) \lesssim n^{-\beta}$ for some $\beta > 3$, for all $\varphi \in \mathcal{H}_\alpha$ and $\psi \in L^\infty(\mu)$, then (f, μ) satisfies the Central Limit Theorem, the Almost Sure Invariance Principle, the vector-valued Almost Sure Invariance Principle, the Local Limit Theorem and the Berry-Esseen Theorem.*

All these concepts are formally defined in Appendix B, together with the precise form of these Theorems and Principles which we obtain in our setting including values of constants which appear in the statements. The assumption that $\mathrm{Cor}_\mu(\varphi, \psi \circ f^n) \lesssim n^{-\beta}$ for some $\beta > 3$ implies, by Theorems A and C, that there exists a Gibbs-Markov induced map with $m(\mathscr{R}_n) \lesssim n^{\beta'}$, for some $\beta' > 2$. All the required statistical properties then follow from this by existing results: The Central Limit Theorem [Yo2], the Almost Sure Invariance Principles [MN1, MN2], the Local Limit Theorem and the Berry-Esseen Theorem [Go1].

Interestingly, large deviation properties also follow from the decay of correlations, but they can be proved directly in an abstract measurable setting and do not depend at all on the existence of the induced Gibbs-Markov map.

1.6. **Applications.** We give a selection of specific systems which satisfy the conditions of some of the Theorems given above.

1.6.1. *Large deviations for Viana maps.* An important class of nonuniform expanding dynamical systems (with critical sets) in dimension greater than one was introduced by Viana in [Vi]. This has served as a model for some relevant results on the ergodic properties of non-uniformly expanding maps in higher dimensions; see [Al1, AA, ABV, AV]. This class of maps can be described as follows. Let $a_0 \in (1, 2)$ be such that the critical point $x = 0$ is pre-periodic for the quadratic map $Q(x) = a_0 - x^2$. Let $S^1 = \mathbb{R}/\mathbb{Z}$ and $b : S^1 \to \mathbb{R}$ be a Morse function, for instance, $b(s) = \sin(2\pi s)$. For fixed small $\alpha > 0$, consider the map

$$\begin{array}{rccc} \hat{f}: & S^1 \times \mathbb{R} & \longrightarrow & S^1 \times \mathbb{R} \\ & (s, x) & \longmapsto & \big(\hat{g}(s), \hat{q}(s, x)\big) \end{array}$$

where $\hat{q}(s, x) = a(s) - x^2$ with $a(s) = a_0 + \alpha b(s)$, and $\hat{g}$ is the uniformly expanding map of the circle defined by $\hat{g}(s) = ds$ (mod $\mathbb{Z}$) for some large integer d. In fact, d was chosen greater or equal to 16 in [Vi], but recent results in [BST] showed that some estimates in [Vi] can be improved and $d = 2$ is enough. It is easy to check that for $\alpha > 0$ small enough there is an interval $I \subset (-2, 2)$ for which $\hat{f}(S^1 \times I)$ is contained in the interior of $S^1 \times I$. Thus, any map f sufficiently close to $\hat{f}$ in the C^0 topology has $S^1 \times I$ as a forward invariant region. We consider from here on these maps restricted to $S^1 \times I$ and we call any such map a *Viana map.* It was shown in [Al1, AV] that Viana maps have a unique ergodic expanding acip μ.

Corollary G. *Let f be a Viana map and let μ be its unique expanding acip. Then, for every $\varphi \in \mathcal{H}_\alpha$ there exists $\tau = \tau(\varphi) > 0$ and, for every $\epsilon > 0$, there exists $C = C(\varphi, \epsilon) > 0$ such that $LD_\mu(\varphi, \epsilon, n) \leq Ce^{-\tau n^{1/5}}$.*

As observed for example in [ALP], Viana maps satisfy the non-degeneracy conditions on the critical set. Moreover, it is proved in [Go] that every Viana map exhibits stretched exponential decay of correlations, with $\theta = 1/2$, for Hölder continuous functions against $L^\infty(\mu)$ functions. The theorem is then a direct application of part (2) of Theorem D.

We emphasize that there have been several recent results concerning large deviations for nonuniformly expanding maps, see [AP, MN, RY, Me], they all either do not apply to this class of maps or give weaker estimates such as polynomial large deviations. This is therefore a new result. All the other statistical properties as per Corollary F also hold, but these all already follow explicitly or implicitly from the existing results in [Me, MN, Go].

1.6.2. *Systems with spectral gap.* Let M be a measurable space (at this stage M needs not to be a Riemannian manifold) endowed with a reference probability measure m on a σ-algebra $\mathcal{M}$, and let $f : M \to M$ be a measurable map. Consider the usual Perron-Frobenius operator $P_m : L^1(m) \to L^1(m)$ as in Appendix A. Assume that there is a seminorm $|\cdot|_\mathcal{B}$ on $L^1(m)$ such that:

(1) $\mathcal{B} = \{\varphi \in L^1(m) : |\varphi|_\mathcal{B} < \infty\}$ is a Banach space with the norm

$$\|\cdot\|_\mathcal{B} = |\cdot|_\mathcal{B} + \|\cdot\|_{L^1(m)};$$

(2) $\mathcal{B}$ is adapted to $L^1(m)$: the inclusion $\mathcal{B} \hookrightarrow L^1(m)$ is compact;
(3) $P_m(\mathcal{B}) \subset \mathcal{B}$ and $P_m|_\mathcal{B}$ is bounded with respect to the norm $\|\cdot\|_\mathcal{B}$;
(4) Lasota-Yorke inequality holds: there are $n_0 \geq 1$, $0 < \alpha < 1$ and $\beta > 0$ such that

$$|P_m^{n_0}\varphi|_\mathcal{B} \leq \alpha|\varphi|_\mathcal{B} + \beta\|\varphi\|_{L^1(m)}, \quad \forall \varphi \in \mathcal{B};$$

(5) $\mathcal{B}$ is a Banach algebra with the norm $\|\cdot\|_\mathcal{B}$; in particular, there is $C > 0$ such that

$$\|\varphi\psi\|_\mathcal{B} \leq C\|\varphi\|_\mathcal{B}\|\psi\|_\mathcal{B}, \quad \forall\, \varphi, \psi \in \mathcal{B};$$

(6) $\mathcal{B}$ is continuously injected in $L^\infty(m)$: there exist a constant $C' > 0$ such that

$$\|\varphi\|_{L^\infty(m)} \leq C'\|\varphi\|_\mathcal{B}, \quad \forall \varphi \in \mathcal{B}.$$

We shall give some explicit examples of maps satisfying these conditions in Appendix C.

Corollary H. *Let $f : M \to M$ verify conditions (1)-(6). Then f admits an invariant expanding acip μ. Assuming moreover that $d\mu/dm$ is uniformly bounded away from zero, then*

(1) f exhibits exponential (and thus in particular, summable) decay of correlations for observables in $\mathcal{B}$ against $L^1(\mu)$;
(2) for all $\varphi \in \mathcal{B}$ there exists $\tau = \tau(\varphi) > 0$ and, for every $\epsilon > 0$, there exists $C = C(\varphi, \epsilon) > 0$ such that $LD_\mu(\varphi, \epsilon, n) \leq Ce^{-\tau n}$.

The proof the first part of this result is relatively standard and we include it in the Appendix in Section C.4. The large deviation estimate in the second part is, as far as we know, a new result, and follows from the first part and a direct application of Theorem D.

Suppose now that we are in the setting of Theorem A or of Theorem C. In this case we get, in addition to the properties (1) and (2) from Corollary H, also the geometric structure of a Gibbs-Markov induced map.

Corollary I. *Let $f : M \to M$ be a C^{1+} local diffeomorphism with an ergodic expanding acip μ, or a C^{1+} local diffeomorphism outside a nondegenerate critical set $\mathcal{C}$ with an ergodic expanding acip μ such that $d\mu/dm \in L^p(m)$, for some $p > 1$. Assuming moreover that $d\mu/dm$ is uniformly bounded away from zero and f verifies conditions (1)-(6), then there is a Gibbs-Markov induced map with $m(\mathscr{R}_n) \lesssim e^{-\tau' n^{1/10}}$, for some $\tau' > 0$.*

This result follows from the exponential decay of correlations in part (1) of Corollary H and part (2) of Theorem C. Corollary F also applies and yields several other statistical properties, but these are probably all already known or at least deducible from the existing literature.

1.6.3. *Non-summable decay for intermittent maps.* Finally we give an application of our results to show that one-dimensional intermittent maps *cannot* exhibit summable decay of correlations agains L^1 functions. Let $f : S^1 \to S^1$ be a C^{1+} local diffeomorphism of the circle satisfying $f'(x) > 1$ for all $x \neq 0$ and such that

$$f(x) \approx x + |x|^{1+\gamma}$$

in some neighbourhood of 0, for some $\gamma \in (0, 1)$. We remark that the notation $\approx$ is used here to indicate the fact that f in a neighbourhood of 0 is equal to $x + |x|^{1+\gamma}$ plus higher order terms and the first and second derivative of the higher order terms are still of higher order.

This is a very well known and well studied class of maps, see e.g. [Pia, PY, Hu, Hol, Sar, PS], first introduced in [PM]. They are well known to have a unique expanding acip μ. Their decay of correlations has been studied in detail and been shown to be at least polynomial for several classes of observables in several papers, we mention for example [LSV] for C^1 observables, in [Yo2] for Hölder continuous observables.

Corollary J. *Suppose there exists $\xi(n)$ such that $Cor_\mu(\varphi, \psi \circ f^n) \leq \xi(n)$ for all $\varphi \in \mathcal{H}_\alpha$ and $\psi \in L^1(\mu)$. Then $\sum_{n=0}^{\infty} \xi(n) = \infty$.*

This follows by contradiction from Theorem B. Indeed, this states that summable decay of correlations against all L^1 functions implies the existence of a Gibbs-Markov induced map with exponential tail of the return times. By [Yo2] this implies exponential decay of correlations for all Hölder continuous observables. However, it is proved in [Hu], see also [Sar], that the decay of correlations cannot be faster than polynomial: there exist Lipschitz functions $\varphi, \psi : S^1 \to \mathbb{R}$ such that $\mathrm{Cor}_\mu(\varphi, \psi \circ f^n) \geq Cn^{1-1/\gamma}$. This gives rise to a contradiction and thus Theorem J holds.

1.7. **Strategy and overview.** In Section 2 we prove Theorems D and E, namely the fact that decay of correlations imply large deviations. These are abstract results of an essentially probabilistic nature and can be formulated in terms of bounds on sums of random variables. For the polynomial case we follow [Me]. For the other cases we apply a result of Azuma and Hoeffding (see Appendix A) on large deviations for a sequence of martingale differences. In the exponential case we will need to use that $(P_\mu^n \varphi)_n$ is summable in $L^\infty(\mu)$ for every $\varphi \in L^\infty(\mu)$, where P_μ is the Perron-Frobenius operator, and we can show this under the assumption of summable decay of correlation against $L^1(\mu)$ functions.

In Section 3 we prove Theorems A and B. These follow by applying first Theorem D to get the large deviation estimates. We then formulate and prove Proposition 3.1 where we show that the large deviation estimates imply the existence of the induced Gibbs-Markov map. This is relatively straightforward since in the case of C^{1+} local diffeomorphisms, the function $\log \|Df^{-1}\|$ is Hölder continuous and therefore, from Theorem D satisfies large deviations either with a polynomial or stretched exponential rate or, from Theorem E, with an exponential rate. We show that such large deviation rates for $\log \|Df^{-1}\|$ imply the assumptions of the constructions of Gibbs-Markov induced maps in [ALP, Go] which therefore yield the desired result.

The situation in the presence of critical points or singularities is significantly more complicated. We still eventually show that the assumptions of [ALP, Go] are satisfied, but in this case we need large deviation estimates for both functions $\log \|Df^{-1}\|$ and $-\log d(x, \mathcal{C})$, where $d(x, \mathcal{C})$ denotes the distance to the critical/singular set, neither of which in this case are Hölder continuous. In Section 4 we assume large deviation estimates (polynomial, stretched exponential, and exponential) for these two functions and show how to obtain the Gibbs-Markov maps with the required tail estimates, and thus in particular deduce the proof of Theorem C.

In Section 5, we use an approximation argument to obtain large deviation estimates for the two particular functions we are interested in, even though they are not Hölder continuous, using the fact that we have the estimates for Hölder continuous functions. Technically, it is exactly at this point that we lose the exponential estimates and are thus not able to prove a version of Theorem B for systems with critical or singular points.

In Appendix A, we give standard definitions and notation concerning Perron-Frobenius operators and martingales, and state the two main probabilistic theorems which we apply in the paper. In Appendix B we give precise statements of the results in Corollary F. In Appendix C we give several classes of piecewise expanding maps which satisfy the assumptions of Corollary H above.

We conclude this introduction with some brief remarks concerning the assumption that $d\mu/dm$ is bounded away from zero on its support, which appears in the statement of some of the results, specifically when dealing with stretched exponential and exponential estimates. This is due to some subtle differences between the construction of induced Markov maps in [ALP] where polynomial estimates are obtained, and [Go], where stretched exponential and exponential (as well as polynomial) estimates are obtained. Both papers work with similar sets of assumptions but the construction of [Go] is in some sense more "global", thus requiring an assumption on the density $d\mu/dm$ on all of its support. On the other hand, it is possible to prove that the density $d\mu/dm$ is necessarily bounded away from zero in some small ball, and this is sufficient for the construction of [ALP], which is more "local". It is not therefore clear at this point whether this assumption is merely technical.

2. Decay of correlations imply large deviations

In this section we prove Theorem D. Assume that $f : M \to M$ is measurable and nonsingular with respect to an ergodic acip μ defined on a σ-algebra $\mathcal{M}$ of M, and let $\mathcal{B} \subset L^\infty(\mu)$ be a Banach space. Let $\varphi \in \mathcal{B}$ and suppose without loss of generality that $\int \varphi d\mu = 0$. For $n \in \mathbb{N}$ we write

$$S_n = \sum_{i=0}^{n-1} \varphi \circ f^i. \tag{2.1}$$

We are therefore interested in an upper bound for $\mu(|S_n| > \epsilon n)$. The idea of the proof of Theorem D is to write S_n as the sum of martingale differences plus some error terms that can be controlled by means of the assumption on the rate of decay of correlations. Then, everything boils down to bound the sum of martingale differences using for which we use existing results from the literature, see below.

Several standard notions that we will use in this section are collected for convenience in Appendix A. In particular, we shall use repeatedly properties (P1)-(P5) about Perron-Frobenius and Koopman operators

$$P_\mu : L^1(\mu) \to L^1(\mu) \quad \text{and} \quad U_\mu : L^\infty(\mu) \to L^\infty(\mu).$$

For notational simplicity we shall omit the measure μ in the notation for these operators and spaces. Also, we denote by $\|\cdot\|_p$ the usual norm in $L^p(\mu)$ for $1 \le p \le \infty$. We define for $j = 1, \ldots, n$

$$\mathcal{F}_j = f^{-(n-j)}\mathcal{M}. \tag{2.2}$$

Observe that the measurability of f does indeed imply that $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \subset \mathcal{F}_n$. Then let

$$X_j := \varphi \circ f^{n-j}$$

Notice that the measurability of f implies that each X_j is measurable with respect to $\mathcal{F}_j$ and therefore $\{\mathcal{F}_j\}_{j=1}^n$ indeed forms a filtration as defined in Appendix A. For every $k \in \mathbb{N}$ let

$$\chi^{(k)} := \sum_{j=1}^{k} P^j \varphi \quad \text{and} \quad \xi^{(k)} := \varphi + \chi^{(k)} - \chi^{(k)} \circ f - P^k \varphi, \tag{2.3}$$

and, for every $j = 1, ..., n$,

$$Z_j^{(k)} := \xi^{(k)} \circ f^{n-j}. \tag{2.4}$$

It is straightforward to check that

$$X_j = Z_j^{(k)} + (\chi^{(k)} \circ f^{n-j+1} - \chi^{(k)} \circ f^{n-j}) + (P^k \varphi) \circ f^{n-j}, \tag{2.5}$$

and therefore

$$S_n = \sum_{j=1}^{n} X_j = \sum_{j=1}^{n} Z_j^{(k)} + \chi^{(k)} \circ f^n - \chi^{(k)} + \sum_{j=1}^{n} P^k\varphi \circ f^{n-j}. \tag{2.6}$$

We emphasize that this equality holds for every k. At the moment k is a free parameter, but we shall eventually choose k as a function of n in order to get the final estimates. The terms above will be used in the polynomial and stretched exponential case. For the exponential case we use a similar decomposition essentially taking $k = \infty$. Then we write

$$\chi := \sum_{i=1}^{\infty} P^i\varphi \quad \text{and} \quad \xi := \varphi + \chi - \chi \circ f,$$

and, for every $j = 1, ..., n$,

$$Z_j := \xi \circ f^{n-j} \tag{2.7}$$

We remark that we will show in the exponential case that χ is well defined and in fact lies in L^∞. It is straightforward to check that

$$S_n = \sum_{j=1}^{n} Z_j + \chi \circ f^n - \chi. \tag{2.8}$$

Lemma 2.1. *$\{Z_j^{(k)}\}_{j=1}^n$ is a sequence of martingale differences.*

Proof. Clearly, $Z_j^{(k)}$ is measurable with respect to $\mathcal{F}_j$, for all $j = 1, \dots, n$. By property (P1) and the invariance of μ we have

$$\begin{aligned}
\mathbb{E}(Z_1^{(k)}) &= \int \varphi \circ f^{n-1} d\mu + \int \chi^{(k)} \circ f^{n-1} d\mu - \int \chi^{(k)} \circ f^n d\mu - \int P^k\varphi \circ f^{n-1} d\mu \\
&= \int \varphi d\mu + \int \chi^{(k)} d\mu - \int \chi^{(k)} d\mu - \int P^k \varphi d\mu = 0.
\end{aligned}$$

Hence, it remains to show that $\mathbb{E}(Z_j^{(k)} | \mathcal{F}_{j-1}) = 0$ for every $j = 2, .., n$. Using (P3) we have

$$\begin{aligned}
P\xi^{(k)} &= P\varphi + P\chi^{(k)} - PU\chi^{(k)} - P^{k+1}\varphi \\
&= P\varphi + P\chi^{(k)} - \chi^{(k)} - P^{k+1}\varphi \\
&= P\varphi + \left(\sum_{n=1}^{k} P^{n+1}\varphi - \sum_{n=1}^{k} P^n\varphi \right) - P^{k+1}\varphi \\
&= P\varphi + (P^{k+1}\varphi - P\varphi) - P^{k+1}\varphi = 0
\end{aligned} \tag{2.9}$$

By property (P4) we have $\mathbb{E}(\cdot | f^{-(i+1)}(\mathcal{M})) = U^{i+1}P^{i+1}$, then using property (P3) and (2.9) it follows that for all $i = 0, \dots, n-2$,

$$\mathbb{E}(Z_{n-i}^{(k)} | \mathcal{F}_{n-i-1}) = \mathbb{E}(\xi^{(k)} \circ f^i | f^{-(i+1)}(\mathcal{M})) = U^{i+1}P^{i+1}U^i\xi^{(k)} = U^{i+1}P\xi^{(k)} = 0, \tag{2.10}$$

which completes the proof that $\{Z_j^{(k)}\}_{j=1}^n$ is a sequence of martingale differences. □

Lemma 2.2. *For any $j \in \mathbb{N}$, $q \geq 1$ and $\psi = sgn(P^j\varphi)$ we have*
$$\|P^j\varphi\|_q \leq Cor_\mu(\varphi, \psi \circ f^j)^{1/q}\|\varphi\|_{\mathcal{B}}^{1/q}\|\varphi\|_\infty^{1-1/q}.$$

Proof. We start by writing
$$\begin{aligned}\|P^j\varphi\|_q &= \left(\int |P^j\varphi|^q \, d\mu\right)^{1/q} \\ &\leq \left(\|P^j\varphi\|_\infty^{q-1}\int |P^j\varphi| \, d\mu\right)^{1/q} \\ &= \left(\|P^j\varphi\|_\infty^{q-1}\|P^j\varphi\|_1\right)^{1/q}. \end{aligned} \tag{2.11}$$
We use property (P5) to get
$$\|P^j\varphi\|_\infty^{q-1} \leq \|\varphi\|_\infty^{q-1}.$$
Then, taking $\psi = \mathrm{sgn}(P^j\varphi)$, using property (P2) and our assumptions on polynomial decay of correlations we have
$$\|P^j\varphi\|_1 = \int \left|P^j\varphi\right| d\mu = \int (P^j\varphi)\psi d\mu = \int \varphi(\psi \circ f^j)d\mu = \|\varphi\|_{\mathcal{B}}\|\psi\|_\infty \mathrm{Cor}_\mu(\varphi, \psi \circ f^j)$$
Thus, substituting into (2.11) and using that $\|\psi\|_\infty = 1$ we get
$$\|P^j\varphi\|_q = \left(\|P^j\varphi\|_\infty^{q-1}\|P^j\varphi\|_1\right)^{1/q} \leq C^{1/q}\|\varphi\|_{\mathcal{B}}^{1/q}\|\varphi\|_\infty^{1-1/q}j^{-\beta/q}.$$
□

2.1. **Polynomial case.** In this section we recall the precise form of the result in [Me, Theorem 1.2 and Lemma 2.1] which implies the statement in the first part of Theorem D but gives more explicit forms of the constants, which will be required below.

Proposition 2.3. *Let $\beta, C > 0$ be such that for all $\psi \in L^\infty$ we have*
$$Cor_\mu(\varphi, \psi \circ f^n) \leq Cn^{-\beta}.$$
Then there exists a constant $C' > 0$, depending only on C, such that for every $\epsilon > 0$ and $q > \max\{1, \beta\}$ we have
$$LD_\mu(\varphi, \epsilon, n) \leq C'\|\varphi\|_{\mathcal{B}}\|\varphi\|_\infty^{2q-1}\epsilon^{-2q}\, n^{-\beta}.$$

We remark that the proof of this result, and of the previous related results in [MN], use Rio's inequality on the sum of random variables, see [Ri, MPU].

2.2. **Stretched exponential case.**

Proposition 2.4. *Let $C, \tau, \theta > 0$ be such that for all $\psi \in L^\infty$ we have*
$$Cor_\mu(\varphi, \psi \circ f^n) \leq Ce^{-\tau n^\theta}.$$
Then, for every $\epsilon > 0$ and $\tau' = \min\{\tau, \epsilon^2/(162\|\varphi\|_\infty^2)\}$ we have
$$LD_\mu(\varphi, \epsilon, n) \leq \left(2 + \frac{C\|\varphi\|_{\mathcal{B}}}{\epsilon}\right)e^{-\tau' n^{\theta/(\theta+2)}}.$$

Proof. From (2.6) we can bound $\mu\left(n^{-1}|S_n|>\epsilon\right)$ by

$$\mu\left(\frac{1}{n}\left|\sum_{j=1}^{n} Z_j^{(k)}\right|>\frac{\epsilon}{3}\right)+\mu\left(\frac{1}{n}|\chi^{(k)}\circ f^n-\chi^{(k)}|>\frac{\epsilon}{3}\right)+\mu\left(\frac{1}{n}\left|\sum_{j=1}^{n} P^k\varphi\circ f^{n-j}\right|>\frac{\epsilon}{3}\right). \tag{2.12}$$

We shall estimate each of the three terms in (2.12) separately and by distinct arguments. We start with a preliminary remark which will be used for both the first and the second terms. Since P is defined with respect to the invariant measure μ, by property (P5) we have that $\|P\varphi\|_\infty\le\|\varphi\|_\infty$ and therefore we get $\|\chi^{(k)}\|_\infty\le k\|\varphi\|_\infty$ which immediately implies

$$\|\chi^{(k)}\circ f^n-\chi^{(k)}\|_\infty\le 2k\|\varphi\|_\infty. \tag{2.13}$$

From the definition of $Z_j^{(k)}$ and using (2.13), we have for $k>2$

$$\|Z_j^{(k)}\|_\infty\le\|\varphi\|_\infty+2k\|\varphi\|_\infty+\|\varphi\|_\infty\le 2(k+1)\|\varphi\|_\infty\le 3k\|\varphi\|_\infty. \tag{2.14}$$

By Lemma 2.1 we know that the $Z_j^{(k)}$ form a sequence of martingale differences. Then, letting $b=\epsilon/3$ and $a=3k\|\varphi\|_\infty$ and applying the Azuma-Hoeffding inequality thus gives

$$\mu\left(\frac{1}{n}\left|\sum_{j=1}^{n} Z_j^{(k)}\right|>\frac{\epsilon}{3}\right)\le 2\exp\left\{-\frac{n\epsilon^2}{162k^2\|\varphi\|_\infty^2}\right\}. \tag{2.15}$$

To estimate the third term in (2.12) we use Chebyshev-Markov's inequality and the invariance of μ to get

$$\begin{aligned}\mu\left(\frac{1}{n}\left|\sum_{j=1}^{n} P^k\varphi\circ f^{n-j}\right|>\frac{\epsilon}{3}\right)&\le\frac{3}{\epsilon n}\int\left|\sum_{j=1}^{n} P^k\varphi\circ f^{n-j}\right|d\mu\\&\le\frac{3}{\epsilon n}\sum_{j=1}^{n}\int\left|P^k\varphi\circ f^{n-j}\right|d\mu\\&\le\frac{3}{\epsilon}\int\left|P^k\varphi\right|d\mu\\&\le\frac{3}{\epsilon}C\|\varphi\|_{\mathcal{B}}e^{-\tau k^\theta}.\end{aligned} \tag{2.16}$$

For the last inequality we have used a simple application of Lemma 2.2 with $q=1$ and our assumptions on the stretched exponential decay of correlations. Notice that the estimates obtained in (2.15) and (2.16) involve k. At this point we set

$$k=k(n):=n^{1/(\theta+2)}.$$

Then, for all sufficiently large n, we have from (2.13) that the condition in the second term of (2.12) is never satisfied and so the term vanishes. Therefore substituting (2.15) and (2.16) and the formula for $k(n)$ into (2.12) we get

$$\mu\left(\frac{1}{n}|S_n|>\epsilon\right)\le 2\exp\left\{-\frac{\epsilon^2}{168\|\varphi\|_\infty^2}\,n^{\frac{\theta}{\theta+2}}\right\}+\frac{C\|\varphi\|_{\mathcal{B}}}{\epsilon}\,\exp\left\{-\tau n^{\frac{\theta}{\theta+2}}\right\}.$$

This completes the proof of Proposition 2.4. □

2.3. **Exponential case.** The actual technical condition needed to get exponential large deviations is given in the following

Proposition 2.5. *Let $\varphi \in L^\infty$ and suppose that*

$$\sum_{n=0}^{\infty} P^n\varphi \in L^\infty.$$

Then for every $\epsilon > 0$ there exists $C' = C'(\varphi, \epsilon) > 0$ such that

$$LD_\mu(\varphi, \epsilon, n) \le C' e^{-\tau n},$$

where $\tau = 1/8(\|\varphi\|_\infty + 2\|\sum P^n\varphi\|_\infty)^2$.

Proof. We show that $\{Z_j\}_{j=1}^n$ as defined in (2.7) is a finite sequence of martingale differences with respect to the filtration $\{\mathcal{F}_j\}_{j=1}^n$, where $\mathcal{F}_j = f^{-(n-j)}\mathcal{M}$, as in (2.2). Indeed, as before, we also have that Z_j is measurable with respect to $\mathcal{F}_j$ for all $j = 1, \dots, n$ and

$$\mathbb{E}(Z_1) = \int \varphi \circ f^{n-1} d\mu + \int \chi \circ f^{n-1} d\mu - \int \chi \circ f^n d\mu = \int \varphi d\mu + \int \chi d\mu - \int \chi d\mu = 0.$$

Furthermore

$$P\xi = P\varphi + P\chi - PU\chi = P\varphi + (P\chi - \chi) = P\varphi - P\varphi = 0,$$

which allows us to conclude that for all $i = 0, \dots, n-2$,

$$\mathbb{E}(Z_{n-i} | \mathcal{F}_{n-i-1}) = \mathbb{E}(\xi \circ f^i | f^{-(i+1)}(\mathcal{M})) = U^{i+1}P^{i+1}U^i\xi = U^{i+1}P\xi = 0,$$

where we used property (P3) and the fact that property (P4) implies that

$$\mathbb{E}(\cdot | f^{-(i+1)}(\mathcal{M})) = U^{i+1}P^{i+1}.$$

Additionally, from the definition of Z_j we have, for all $j = 1, \dots, n$,

$$\|Z_j\|_\infty \le \|\varphi\|_\infty + 2\|\chi\|_\infty. \tag{2.17}$$

and therefore, by the Azuma-Hoeffding inequality we get

$$\mu\left(\frac{1}{n}\left|\sum_{j=1}^n Z_j\right| > \frac{\epsilon}{2}\right) \le 2\exp\left\{-\frac{\epsilon^2}{8(\|\varphi\|_\infty + 2\|\chi\|_\infty)^2}\, n\right\}.$$

Thus, for all sufficiently large values of n, in particular for $n \ge N$ where $2/N\|\chi\|_\infty \le \epsilon/2$ we have

$$\begin{aligned}
\mu\left(\frac{1}{n}|S_n| > \epsilon\right) &\le \mu\left(\frac{1}{n}\left|\sum_{j=1}^n Z_j\right| + \frac{2}{n}\|\chi\|_\infty > \epsilon\right) \\
&\le \mu\left(\frac{1}{n}\left|\sum_{j=1}^n Z_j\right| > \frac{\epsilon}{2}\right) \le 2\exp\left\{-\frac{\epsilon^2}{8(\|\varphi\|_\infty + 2\|\chi\|_\infty)^2}\, n\right\}
\end{aligned}$$

□

To complete the proof of Theorem B it therefore just remains to show that the assumption of Proposition 2.5 is satisfied.

Lemma 2.6. *Let $\mathcal{B} \subset L^\infty$ be a Banach space and $\varphi \in \mathcal{B}$ with $\int \varphi d\mu = 0$. If there is $\xi(n)$ with $\sum_{n=0}^{\infty} \xi(n) < \infty$ and $Cor_\mu(\varphi, \psi \circ f^n) \leq \xi(n)$ for all $\psi \in L^1$, then*

$$\sum_{n=0}^{\infty} P^n \varphi \in L^\infty.$$

Proof. By Riesz' representation theorem we may identify L^∞ with the dual of L^1 by associating to $\varphi \in L^\infty$ the linear functional $\ell_\varphi : L^1 \to \mathbb{R}$ defined by $\ell_\varphi(\psi) = \int \varphi\psi d\mu$. Since $\|\varphi\|_\infty = \|\ell_\varphi\|$, we have for all $n \geq 0$

$$\begin{aligned}
\|P^n\varphi\|_\infty &= \sup_{\psi \in L^1} \frac{|\int (P^n\varphi)\psi d\mu|}{\|\psi\|_1} \\
&= \sup_{\psi \in L^1} \frac{|\int \varphi(\psi \circ f^n) d\mu|}{\|\psi\|_1} \\
&= \frac{\|\varphi\|_{\mathcal{B}} \|\psi\|_1 \mathrm{Cor}_\mu(\varphi, \psi \circ f^n)}{\|\psi\|_1} \\
&\leq \|\varphi\|_{\mathcal{B}} \xi(n).
\end{aligned}$$

Therefore

$$\left\| \sum_{n=0}^{\infty} P^n \varphi \right\|_\infty \leq \sum_{n=0}^{\infty} \|P^n\varphi\|_\infty \leq \|\varphi\|_{\mathcal{B}} \sum_{n=0}^{\infty} \xi(n) < \infty.$$

□

3. Gibbs-Markov structures for local diffeomorphisms

In this section we prove Theorems A and B. We consider the function

$$\phi(x) := \log \|Df(x)^{-1}\|$$

and note that in the case of C^{1+} local diffeomorphisms, ϕ is Hölder continuous. From the assumptions of Theorem A and and B and the conclusions of Theorems D and E we therefore have

$$LD_\mu(\phi, \epsilon, n) = \mathcal{O}(n^{-\beta}) \quad \text{and} \quad LD_\mu(\phi, \epsilon, n) = \mathcal{O}(e^{-\tau n^\theta}) \tag{3.1}$$

in the polynomial case and in the stretched and exponential cases respectively ($\theta = 1$ in the exponential case). Theorems A and B then follow directly from

Theorem 3.1. *Let f be a C^{1+} local diffeomorphism with an ergodic expanding acip μ;*

(1) if there exists $\beta > 1$ such that for small $\epsilon > 0$ we have $LD_\mu(\phi, \epsilon, n) \lesssim n^{-\beta}$, then there is a Gibbs-Markov induced map with $m(\mathscr{R}_n) \lesssim n^{-\beta+1}$.

Suppose moreover that $d\mu/dm$ is uniformly bounded away from 0 on its support. Then

(2) if there exist $\tau, \theta > 0$ such that for small $\epsilon > 0$ we have $LD_\mu(\phi, \epsilon, n) \lesssim e^{-\tau n^\theta}$, then there is a Gibbs-Markov induced map with $m(\mathscr{R}_n) \lesssim e^{-\tau' n^\theta}$, for some $\tau' > 0$.

Notice that the second part of the theorem applies in particular if $\theta = 1$, i.e. in the exponential case. Notice also that the large deviation rates are not assumed to be uniform in ϵ. To prove this theorem we first state a general result which will also be useful in the case of maps with critical/singular sets. Suppose we are given an arbitrary function $\varphi \in L^1$. Define

$$\tilde{S}_n(x) = \tilde{S}_n\varphi(x) := \left| \frac{1}{n} \sum_{i=0}^{n-1} \varphi\left(f^i(x)\right) - \int \varphi d\mu \right|.$$

Then $\tilde{S}_n(x) \to 0$ for μ almost every x. Notice that the large deviation estimates are precisely bounds on the rate of decay of the tail $\mu\{\tilde{S}_n > \epsilon\}$. For $\epsilon > 0$ define

$$N_\epsilon(x) := \min\{N : \tilde{S}_n \le \epsilon \quad \forall\, n \ge N\}. \tag{3.2}$$

Lemma 3.2. *Let $A \subseteq M$ be such that $d\mu/dm > c$ on A for some $c > 0$. Suppose that given $\varphi \in L^1$ and $\epsilon > 0$ there exists $\xi : \mathbb{N} \to \mathbb{R}^+$ such that $LD_\mu(\varphi, \epsilon, n) \le \xi(n)$. Then for every $n \ge 1$ we have*

$$m(\{N_\epsilon > n\} \cap A) \le \frac{1}{c} \sum_{\ell \ge n} \xi(\ell).$$

Proof. For $\epsilon > 0$ we have

$$\{N_\epsilon > n\} \subset M \setminus \bigcap_{\ell \ge n} \{\tilde{S}_\ell \le \epsilon\} \subset \bigcup_{\ell \ge n} \{\tilde{S}_n > \epsilon\}.$$

The assumption on the density gives $m(B) \le \|dm/d\mu\|_\infty \mu(B) \le \mu(B)/c$ for any measurable set $B \subset A$, and therefore

$$m(\{N_\epsilon > n\} \cap A) \le \frac{1}{c}\mu(\{N_\epsilon > n\} \cap A) \le \frac{1}{c}\mu(\bigcup_{\ell \ge n}\{\tilde{S}_\ell \ge \epsilon\}) \le \frac{1}{c}\sum_{\ell \ge n} \xi(n).$$

The last inequality uses the assumption on the large deviation rate function which gives $\mu\{\tilde{S}_n \ge \epsilon\} \le \xi(n)$. $\square$

Proof of Theorem 3.1. By the expansivity assumption on μ and a straightforward application of Birkhoff's ergodic theorem we have that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi(f^j(x)) = \int \phi d\mu =: \lambda < 0 \tag{3.3}$$

is satisfied μ almost everywhere. Thus we have that

$$\mathcal{E}(x) := \min\left\{ N : \frac{1}{n} \sum_{i=0}^{n-1} \phi(f^j(x)) \le \lambda/2 \ \ \forall n \ge N \right\}.$$

is defined and finite almost everywhere in M. Notice that using the notation in (3.2) for $\varphi = \phi$ and $\epsilon = \lambda/2$ we have that

$$\{\mathcal{E} > n\} \subseteq \{N_\epsilon > n\}.$$

In [ALP, Go] induced Markov maps are constructed and it is shown that tails of return times have the same rate of decay (polynomial, stretched or exponential) as the rate of decay of the Lebesgue measure of $m\{\mathcal{E} > n\}$. The conclusion therefore follows from an application of Lemma 3.2, substituting the corresponding polynomial or (stretched) exponential bounds. We just need to specify the set A on which the density of μ is bounded below.

For the polynomial case we take advantage of a result of [ADL] where it is shown that there exists a ball $\Delta_0 \subset \mathrm{supp}(\mu)$ centred at a point p whose preimages are dense in the support of μ, such that the density of μ with respect to Lebesgue is uniformly bounded below on Δ_0. This is sufficient for the purposes of applying the construction of [ALP] which indeed only requires the existence of such a ball and where the required tail estimates are then formulated in terms of the decay of $m(\{\mathcal{E} > n\} \cap \Delta_0)$. In the stretched and exponential case we apply instead the arguments of [Go] which rely on somewhat more global assumptions and therefore require a control of the density on the entire support of μ. For this reason we need to include the boundedness from below of the density as part of our assumptions. Theorem 3.1 is now a direct consequence of the above where we let $A = \Delta_0$ in the polynomial case, or $A = \mathrm{supp}(\mu)$ in the other cases. □

4. Gibbs-Markov structures for maps with critical/singular sets

In this section we consider maps with critical/singular sets and prove Theorem C. We shall follow a similar strategy used in the proof of Theorem A and once again we aim to apply the construction and estimates of [ALP, Go]. A main difference here is that the function $\log \|Df^{-1}\|$ is not necessarily Hölder continuous and therefore we cannot apply directly the results of Theorem D which give bounds on the large deviation rates. Moreover, we also need to consider an additional function related to the recurrence to the critical/singular set. We let

$$\phi_1(x) = \log \|Df^{-1}\| \quad \text{and} \quad \phi_2(x) = \phi_2^{(\delta)}(x) = \begin{cases} -\log d(x, \mathcal{C}) & \text{if } d(x,\mathcal{C}) < \delta\ , \\ \frac{\log \delta}{\delta}(d(x,\mathcal{C}) - 2\delta) & \text{if } \delta \le d(x,\mathcal{C}) < 2\delta\ , \\ 0 & \text{if } d(x,\mathcal{C}) \ge 2\delta, \end{cases}$$

where $\delta > 0$ is a small constant to be fixed later. We remark that $\phi_2(x) = -\log d(x,\mathcal{C})$ in the δ neighbourhood and $\phi_2(x) = 0$ outside a 2δ neighbourhood of the critical set $\mathcal{C}$. The definition in the remaining region is motivated by the requirement that the function be Hölder continuous except at the critical/singular set. We do need some large deviation estimates for these functions as we had in (3.1) for the local diffeomorphism case. These are provided in the following

Proposition 4.1. *Let $f : M \to M$ be a C^{1+} local diffeomorphism outside a nondegenerate critical set $\mathscr{C}$. Suppose that f admits an ergodic expanding acip μ with $d\mu/dm \in L^p(m)$ for some $p > 1$;*

(1) if there exists $\beta > 1$ such that $Cor_\mu(\varphi, \psi \circ f^n) \lesssim n^{-\beta}$ for every $\varphi \in \mathcal{H}$ and $\psi \in L^\infty$, then for every $\gamma > 0$ there is $C' > 0$ such that $LD_\mu(\phi_i, \epsilon, n) \leq C' n^{-\beta+\gamma}$, for $i = 1, 2$.

Suppose moreover that $d\mu/dm$ is uniformly bounded away from 0 on its support;

(2) if there exist $\tau, \theta > 0$ such that $Cor_\mu(\varphi, \psi \circ f^n) \lesssim e^{-\tau n^\theta}$ for every $\varphi \in \mathcal{H}$ and $\psi \in L^\infty$, then there exists $\zeta > 0$ such that for any $\gamma > 0$ and $\epsilon > 0$ sufficiently small there is $C' > 0$ such that $LD_\mu(\phi_i, \epsilon, n) \leq C' e^{-\zeta n^{\theta/(3\theta+6)-\gamma}}$ for $i = 1, 2$.

The proof of Proposition 4.1 is relatively technical and we postpone it to the following section. Assuming the conclusions of this proposition for the moment, Theorem B follows from

Theorem 4.2. *Let $f : M \to M$ be a C^{1+} local diffeomorphism outside a nondegenerate critical set $\mathscr{C}$. Suppose that f admits an ergodic expanding acip μ with $d\mu/dm$ with $d\mu/dm \in L^p(m)$ for some $p > 1$. Then $\phi_i \in L^1(\mu)$ for $i = 1, 2$. Moreover,*

(1) if there exists $\beta > 1$ such that for small $\epsilon > 0$ we have $LD_\mu(\phi_i, \epsilon, n) \lesssim n^{-\beta}$ for $i = 1, 2$, then there is a Gibbs-Markov induced map with $m(\mathscr{R}_n) \lesssim n^{-\beta+1}$.

Suppose moreover that $d\mu/dm$ is uniformly bounded away from 0 on its support;

(2) if there exist $\tau, \theta > 0$ such that for small $\epsilon > 0$ we have $LD_\mu(\phi_i, \epsilon, n) \lesssim e^{-\tau n^\theta}$ for $i = 1, 2$, then there is a Gibbs-Markov induced map with $m(\mathscr{R}_n) \lesssim e^{-\tau' n^\theta}$, for some $\tau' > 0$.

Notice that the second part of the theorem applies in particular if $\theta = 1$, i.e. in the exponential case. Notice also that the large deviation rates are not assumed to be uniform in ϵ. We begin by introducing the natural auxiliary function

$$\phi_0(x) := -\log d(x, \mathcal{C}).$$

Then, for $i = 0, 1, 2$ and $k > 0$ we let

$$A_{i,k} := \{x : \phi_i(x) \geq k\}$$

Lemma 4.3. *There exists $\zeta > 0$ such that for all $k > 0$ and for all $i = 0, 1, 2$ we have*

$$(1)\ \mu(A_{i,k}) \lesssim e^{-\zeta k}; \qquad (2)\ \phi_i \in L^1(\mu); \qquad \textit{and} \qquad (3)\ \int \phi_2^{(\delta)} d\mu \to 0 \ \textit{as}\ \delta \to 0.$$

Proof. Recall that we have assumed that $d\mu/dm \in L^p(m)$ for some $p > 1$. We define $q > 1$ by the usual condition $1/p + 1/q = 1$. Then, by Hölder's inequality, we have

$$\mu(A_{0,k}) = \int \mathbb{1}_{A_{0,k}} \frac{d\mu}{dm} dm \leq \|\mathbb{1}_{A_{0,k}}\|_q \left\| \frac{d\mu}{dm} \right\|_p \leq m(A_{0,k})^{1/q} \left\| \frac{d\mu}{dm} \right\|_p \lesssim m(A_{0,k})^{1/q}$$

and thus (1) for $i = 0$ follows directly from condition (C0). For $i = 2$ we also have the result since $\phi_2(x) = \phi_0(x)$ as long as $d(x, \mathcal{C}) \leq \delta$ or, equivalently, $k \geq -\log \delta$. For $i = 1$ we use condition (C1) which implies that there exists a constant $\tilde{B} > 0$ such that for every $x \in M \setminus \mathcal{C}$ we have

$$-\tilde{B} + \eta \log d(x, \mathcal{C}) \le \phi_1(x) \le \tilde{B} - \eta \log d(x, \mathcal{C}). \tag{4.1}$$

Therefore there exists some constant $\tilde{\eta} > 0$ such that $\{\phi_1 \ge k\} \subseteq \{\phi_0 > \tilde{\eta} k\}$ which then clearly gives the conclusion for ϕ_1 and thus completes the proof of (1). The integrability of ϕ_i in (2) now follows easily from the fact that for $i = 0, 1, 2$

$$\int \phi_i d\mu \le \sum_{n=1}^{\infty} \mu(A_{i,n})$$

and using (1). Finally, to prove (3) we let $k_1 = -\log \delta$, $k_2 = -\log 2\delta$ and write

$$\int \phi_2^{(\delta)} d\mu = \int_{A_{0,k_1}} \phi_2^{(\delta)} d\mu + \int_{A_{0,k_2} \setminus A_{0,k_1}} \phi_2^{(\delta)} d\mu + \int_{M \setminus A_{0,k_2}} \phi_2^{(\delta)} d\mu$$

Since $\phi_2^{(\delta)}(x) = 0$ for $x \in M \setminus A_{0,k_2}$, the third term vanishes. For the first term notice that $\phi_2^{(\delta)}(x) = \phi_0(x)$ for $x \in A_{0,k_1}$. Since $\phi_0 \in L^1(\mu)$ and $\mu(A_{0,k_1}) \to 0$ as $\delta \to 0$, it follows that

$$\int_{A_{0,k_1}} \phi_0 d\mu \to 0 \quad \text{as} \quad \delta \to 0.$$

Finally, for the middle term we have $\phi_2^{(\delta)}(x) \le -\log \delta$ for $x \in A_{0,k_2 \setminus A_{0,k_1}}$, and so

$$\int_{A_{0,k_2 \setminus A_{0,k_1}}} \phi_2^{(\delta)} d\mu \le (-\log \delta) \mu(A_{0,k_2}) \lesssim (-\log \delta) e^{-\zeta k_2} \le (-\log \delta)(2\delta)^{\zeta}$$

which clearly tends to zero as $\delta \to 0$. □

Proof of Theorem 4.2. We follow a similar strategy as in the proof of Theorem 3.1, applying the results of [ALP, Go]. We consider as before the tail $\{\mathcal{E}(x) > n\}$ of the expansion time related to the function ϕ_1 but also need to consider an analogous term related to the function ϕ_2. More precisely we need to show that for $\epsilon > 0$ sufficiently small, there exists $\delta > 0$ such that

$$\limsup_{n \to +\infty} \frac{1}{n} \sum_{j=0}^{n-1} -\log d_\delta(f^j(x), \mathcal{C}) \le \epsilon.$$

We note that it is sufficient to have this for some $\epsilon > 0$ depending only on the map, see e.g. [Al2, Remark 3.8]. In fact, fixing such an ϵ, from Lemma 4.3 we can choose $\delta > 0$ sufficiently small so that

$$\limsup_{n \to +\infty} \frac{1}{n} \sum_{j=0}^{n-1} -\log d_\delta(f^j(x), \mathcal{C}) \le \lim_{n \to +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \phi_2^{(\delta)}(f^j(x)) = \int \phi_2^{(\delta)} d\mu \le \epsilon. \tag{4.2}$$

We introduce the *recurrence time* function

$$\mathcal{R}_{\epsilon,\delta}(x) = \min \left\{ N \ge 1 : \frac{1}{n} \sum_{i=0}^{n-1} -\log d_\delta(f^j(x), \mathcal{C}) \le 2\epsilon, \quad \forall n \ge N \right\}$$

which is defined and finite μ almost everywhere in M. Using again the notation in (3.2) for $\varphi = \phi_2^{(\delta)}$ we have

$$\{\mathcal{R}_{\epsilon,\delta} > n\} \subseteq \{N_\epsilon > n\}.$$

In [ALP, Go] induced Markov maps are constructed and it is shown that tails of return times have the same rate of decay (polynomial, stretched or exponential) as the rate of decay of the Lebesgue measure of

$$\{x : \mathcal{E}(x) > n \ \text{ or } \ \mathcal{R}_{\epsilon,\delta}(x) > n\}.$$

The conclusion therefore follows from an application of Lemma 3.2 and Proposition4.1, substituting the corresponding polynomial or (stretched) exponential bounds. We note that here we take A equal to the whole support of μ since we have the density uniformly bounded below by assumption. □

5. Large deviations for the special non Hölder observables

In this section we prove Proposition 4.1. Our strategy is to approximate ϕ_1 and ϕ_2 by "truncated" functions which are Hölder continuous. For all $k > 0$ and $i = 1, 2$, let

$$\phi_{i,k}(x) := \begin{cases} \phi_i(x) & \text{if } x \in M \setminus A_{i,k}; \\ k & \text{if } x \in A_{i,k}. \end{cases}$$

Then we can write, for $i = 1, 2$, $n \in \mathbb{N}$ and $k \in \mathbb{N}$ sufficiently large so that $\left|\int(\phi_i - \phi_{i,k})d\mu\right| < \epsilon/2$,

$$\begin{aligned} \mu\left(\frac{1}{n}|\tilde{S}_n\phi_i(x)| > \epsilon\right) &\le \mu\left(\left\{\frac{1}{n}|\tilde{S}_n\phi_i(x)| > \epsilon\right\} \setminus \bigcup_{j=0}^{n-1} f^{-j}A_{i,k}\right) + \mu\left(\bigcup_{j=0}^{n-1} f^{-j}(A_{i,k})\right) \\ &\le \mu\left(\frac{1}{n}|\tilde{S}_n\phi_{i,k}(x)| + \left|\int(\phi_i - \phi_{i,k})\,d\mu\right| > \epsilon\right) + \sum_{j=0}^{n-1}\mu\left(f^{-j}(A_{i,k})\right) \\ &\le \mu\left(\frac{1}{n}|\tilde{S}_n\phi_{i,k}(x)| > \epsilon/2\right) + n\mu(A_{i,k}). \end{aligned} \tag{5.1}$$

The invariance of the measure μ is used in the last step. The second term in (5.1) is easily bounded by Lemma 4.3.

To bound the first term of (5.1), notice that $\phi_{i,k}$ is Hölder continuous with exponent α for every $\alpha \in (0, 1]$. Therefore we shall use our assumptions which apply to Hölder continuous observables, in particular we will apply the conclusions of Propositions 2.3 and 2.4 with $\mathcal{B} = \mathcal{H}_\alpha$. For this we need to obtain bounds for the L^∞ and Hölder norms of the functions $\phi_{i,k}$.

Lemma 5.1. *For any $0 < \alpha \le 1$ and $i = 1, 2$ we have $\|\phi_{i,k}\|_{\mathcal{H}_\alpha} \lesssim \alpha^{-1}ke^{\alpha k}$.*

Proof. By the definition of $\phi_{i,k}$ we have $\|\phi_{i,k}\|_\infty \le k$ for $i = 1, 2$. Given $x, y \in M \setminus \mathcal{C}$ and assuming without loss of generality that $d(y, \mathcal{C}) \ge d(x, \mathcal{C})$ we have

$$\frac{|\log d(y,\mathcal{C}) - \log d(x,\mathcal{C})\,|}{d(x,y)^\alpha} \le \frac{\left|\log\left(1 + \frac{d(y,\mathcal{C})-d(x,\mathcal{C})}{d(x,\mathcal{C})}\right)\right|}{d(x,\mathcal{C})^\alpha \left(\frac{d(x,y)}{d(x,\mathcal{C})}\right)^\alpha} \le \frac{\left|\log\left(1 + \frac{d(x,y)}{d(x,\mathcal{C})}\right)\right|}{\left(\frac{d(x,y)}{d(x,\mathcal{C})}\right)^\alpha} d(x,\mathcal{C})^{-\alpha}. \quad (5.2)$$

Notice that the function $z^{-\alpha}\log(1+z)$ is bounded above with a global maximum z_0 satisfying $\log(1+z_0) = z_0\alpha^{-1}(1+z_0)^{-1}$. Substituting this back into the function we get $z_0^{-\alpha}\log(1+z_0) = \alpha^{-1}z_0^{1-\alpha}(1+z_0)$ which is bounded by $1/\alpha$. Using this bound in (5.2) we get, for x such that $d(x,\mathcal{C}) \ge e^{-k}$,

$$\frac{|\log d(y,\mathcal{C}) - \log d(x,\mathcal{C})|}{d(x,y)^\alpha} \le \frac{1}{\alpha} d(x,\mathcal{C})^{-\alpha} \le \frac{1}{\alpha} e^{\alpha k}.$$

From (C2) in the nondegeneracy conditions and using that $k + \alpha^{-1}e^{\alpha k} \lesssim \alpha^{-1}ke^{\alpha k}$ we thus obtain the required bound for $\phi_{1,k}$.

For $\phi_{2,k}$ we just need to consider the extra term corresponding to the region where both $d(x,\mathcal{C})$ and $d(y,\mathcal{C})$ belong to $(\delta, 2\delta)$. Here we have

$$\frac{|\phi_{2,k}(x) - \phi_{2,k}(y)|}{d(x,y)^\alpha} \le \frac{\log\delta}{\delta}\frac{|d(x,\mathcal{C}) - d(y,\mathcal{C})|}{d(x,y)^\alpha} \le \frac{\log\delta}{\delta} d(x,y)^{1-\alpha} \lesssim \frac{\log\delta}{\delta}.$$

Keeping in mind that δ is fixed, this completes the proof for $\phi_{2,k}$. □

Proof of Proposition 4.1. We are now ready to estimate the first term in (5.1) and thus complete the estimates required to proof the Proposition. From this point onwards, all estimates will apply equally to $\phi_{1,k}$ and $\phi_{2,k}$. Thus, to simplify the notation we shall just write ϕ_k.

We consider first the polynomial case. Substituting the estimates of Lemma 5.1 into the results of Proposition 2.3 we get

$$\mu\left(\frac{1}{n}|\tilde{S}_n\phi_k(x)| > \epsilon\right) \lesssim \|\phi_k\|_{\mathcal{H}_\alpha}\|\phi_k\|_\infty^{2q-1}\epsilon^{-2q}n^{-\beta} \lesssim \alpha^{-1}k^{2q}e^{\alpha k}\epsilon^{-2q}\,n^{-\beta}, \quad (5.3)$$

Using Lemma 4.3 and substituting (5.3) into (5.1) gives

$$\mu\left(\frac{1}{n}|\tilde{S}_n\phi(x)| > \epsilon\right) \lesssim \alpha^{-1}k^{2q}e^{\alpha k}\epsilon^{-2q}n^{-\beta} + ne^{-\zeta k} \le \epsilon^{-2q}(\alpha^{-1}k^{2q}e^{\alpha k}n^{-\beta} + ne^{-\zeta k}). \quad (5.4)$$

We now complete the estimate by choosing k appropriately and taking advantage of the fact that we can also choose α arbitrarily small. Indeed, if $\varphi \in \mathcal{H}_{\alpha'}$ then $\varphi \in \mathcal{H}_\alpha$ for all $\alpha \in (0, \alpha')$. We aim to obtain an upper bound of the order of $n^{-\beta+\gamma}$ and thus require that the two inequalities

$$ne^{-\zeta k} \lesssim n^{-\beta+\gamma} \quad \text{and} \quad \alpha^{-1}k^{2q}e^{\alpha k} \lesssim n^\gamma$$

are simultaneously satisfied. We will show that this can be achieved by fixing a sufficiently small α and then choosing k, n sufficiently large. First observe that

$$k \ge \frac{\beta + 1 - \gamma}{\zeta}\log n \quad \Rightarrow \quad ne^{-\zeta k} \le n^{-\beta+\gamma}$$

and

$$\alpha k + 2q \log k \leq \log \alpha + \gamma \log n \quad \Rightarrow \quad \alpha^{-1} k^{2q} e^{\alpha k} \leq n^{\gamma}.$$

Now for any fixed α and $k = k(\alpha)$ sufficiently large, we have $\alpha k + 2q \log k \leq 2\alpha k$; also for $n = n(\alpha)$ sufficiently large we have $\frac{\gamma}{2} \log n \leq \log \alpha + \gamma \log n$. Therefore we can write the one-sided implication

$$k \leq \frac{\gamma}{4\alpha} \log n \quad \Rightarrow \quad \frac{1}{\alpha} k^{2q} e^{\alpha k} \leq n^{\gamma}$$

Thus it is enough to show that for α sufficiently small we have

$$\frac{\beta + 1 - \gamma}{\zeta} \log n \leq \frac{\gamma}{4\alpha} \log n.$$

This is clearly true and in fact we can choose the explicit value

$$\alpha = \frac{\gamma\zeta}{4(\beta + 1 - \gamma)}.$$

This completes the proof in the polynomial case.

We now consider the stretched exponential case. Substituting the estimates of Proposition 2.4 and Lemma 4.3 into (5.1) we get

$$\mu\left(\frac{1}{n}|\tilde{S}_n \phi_i(x)| > \epsilon\right) \lesssim \|\phi_{i,k}\|_{\mathcal{H}_\alpha} \epsilon^{-1} e^{-\tau' n^{\theta'}} + n e^{-\zeta k} \tag{5.5}$$

where $\theta' = \theta/(\theta + 2)$ and $\tau' = \min\{\tau, \epsilon^2/(162\|\phi_{i,k}\|_\infty^2)\}$. Notice that taking k sufficiently large we have in fact $\tau' = \epsilon^2/(162k^2)$, and therefore, using the bound on the Hölder norm from Lemma 5.1 and substituting into (5.5) we have

$$\mu\left(\frac{1}{n}|\tilde{S}_n \phi_i(x)| > \epsilon\right) \lesssim k e^{\alpha k} e^{-\epsilon^2 n^{\theta'}/(162k^2)} + n e^{-\zeta k}.$$

We recall once again that the constant implicit in the inequality $\lesssim$ is allowed to depend on ϵ and on α, even though α plays no special role in the stretched exponential case. It is now just a question of making a convenient choice of $k = k(n)$. In this case we choose $k = n^{\frac{\theta'}{3} - \gamma}$ and get

$$k e^{\alpha k} e^{-\epsilon^2 n^{\theta'}/(162k^2)} + n e^{-\zeta k} \leq n e^{(\alpha - \epsilon^2 n^{3\gamma}/162) n^{\frac{\theta'}{3} - \gamma}} + n e^{-\zeta n^{\frac{\theta'}{3} - \gamma}}$$

Now just observe that for any given ϵ, as long as n is sufficiently large we have $\alpha - \epsilon^2 n^{3\gamma}/162 < -\zeta$. Since γ can also be chosen arbitrarily small, we obtain the proof of Proposition 4.1 also in the stretched exponential case. □

Appendix A. Special operators and martingales

A.1. **Perron-Frobenius and Koopman operators.** Let $(M, \mathcal{M}, \mu)$ be a probability measure space and $f\colon M \to M$ a measurable map (not necessarily preserving μ). We say that f is *nonsingular* with respect to μ if $f_*\nu \ll \mu$ whenever $\nu \ll \mu$. Given $\varphi \in L^1(\mu)$, the (signed) measure ν_φ on $\mathcal{M}$, defined for each $A \in \mathcal{M}$ as

$$\nu_\varphi(A) = \int_A \varphi d\mu,$$

is clearly absolutely continuous with respect to μ. Using the nonsingularity of f we define the *Perron-Frobenius operator* $P_\mu : L^1(\mu) \to L^1(\mu)$ by

$$P_\mu \varphi = \frac{df_*\nu_\varphi}{d\mu}.$$

The *Koopman operator* $U_\mu : L^\infty(\mu) \to L^\infty(\mu)$ is defined by

$$U_\mu \varphi = \varphi \circ f.$$

Given $\mathcal{A}$ a sub-σ-algebra of $\mathcal{M}$ and $\varphi \in L^1(\mu)$, the (signed) measure $\nu_\varphi^{\mathcal{A}}$ on $\mathcal{A}$, defined for each $A \in \mathcal{A}$ as

$$\nu_\varphi^{\mathcal{A}}(A) = \int_A \varphi d\mu,$$

is clearly absolutely continuous with respect to $\mu|_{\mathcal{A}}$. We finally define the *conditional expectation* $\mathbb{E}_\mu(\cdot|\mathcal{A}) : L^1(\mu) \to L^1(\mu|_{\mathcal{A}})$ as

$$\mathbb{E}_\mu(\varphi|\mathcal{A}) = \frac{d\nu_\varphi^{\mathcal{A}}}{d\mu|_{\mathcal{A}}}.$$

Observe that $\mathbb{E}_\mu(\varphi|\mathcal{A})$ is the unique $\mathcal{A}$-measurable function such that for each $A \in \mathcal{A}$

$$\int_A \mathbb{E}_\mu(\varphi|\mathcal{A}) d\mu = \int_A \varphi d\mu.$$

Perron-Frobenius and Koopman operators enjoy some well-known properties that we collect in (P1)-(P5) below; see e.g. [GB, Chapter 4]. We observe that in the first two properties we do not need invariance of the measure μ. For all $\varphi \in L^1(\mu)$ we have

(P1) $\int P_\mu \varphi \, d\mu = \int \varphi d\mu$;
(P2) $\int (P_\mu \varphi)\psi d\mu = \int \varphi (U_\mu \psi) d\mu$ for all $\psi \in L^\infty(\mu)$.

Moreover, if μ is f-invariant, then for all $\varphi \in L^1(\mu)$ we have

(P3) $P_\mu U_\mu \varphi = \varphi$;
(P4) $U_\mu^n P_\mu^n \varphi = \mathbb{E}_\mu(\varphi | f^{-n}(\mathcal{M}))$ for all $n \geq 1$;
(P5) $\|P_\mu \varphi\|_p \leq \|\varphi\|_p$ whenever $\varphi \in L^p(\mu)$ for some $1 \leq p \leq \infty$.

A.2. **Filtrations and martingale differences.** Consider a sequence of σ-algebras $\{\mathcal{F}_i\}_{i\in\mathbb{N}}$ which forms a *filtration*, meaning that $\mathcal{F}_i \subset \mathcal{F}_{i+1}$ for all $i \in \mathbb{N}$. We say that a sequence of random variables $\{X_i\}_{i\in\mathbb{N}}$ is *adapted to a filtration* $\{\mathcal{F}_i\}_{i\in\mathbb{N}}$ if each X_i is measurable with respect to $\mathcal{F}_i$. We say that random variables $\{X_i\}_{i\in\mathbb{N}}$ form a sequence of *martingale differences* with respect to a filtration $\{\mathcal{F}_i\}_{i\in\mathbb{N}}$ if the sequence is adapted to the filtration and

$$\mathbb{E}(X_1) = 0, \qquad \mathbb{E}(X_{i+1}|\mathcal{F}_i) = 0, \quad \forall i \geq 1. \tag{A.1}$$

The following result follows from [Az] and [Hoe] and it can be found in the present formulation in [LV, Theorem 3.1].

Theorem A.1 (Azuma-Hoeffding). *Let $\{X_i\}_{i\in\mathbb{N}}$ be a sequence of martingale differences. If there is $a>0$ such that $\|X_i\|_\infty < a$ for all $1\le i\le n$, then for all $b\in\mathbb{R}$ we have*

$$\mu\left(\sum_{i=1}^{n} X_i \ge nb\right) \le e^{-n\frac{b^2}{2a^2}}.$$

Appendix B. Statistical properties

Here we will explain in detail the content of Corollary F. As we define the properties we will also define more precisely the specific form of the result which is contained in the Corollary F.

B.1. **Central Limit Theorem.** As a consequence of Theorem A or C and [Yo2, Theorem 4] we have:

Corollary B.1. *Suppose that $Cor_\mu(\varphi,\psi\circ f^n)\lesssim n^{-\beta}$ for some $\beta>3$ for every $\varphi\in\mathcal{H}_\alpha$ and $\psi\in L^\infty(\mu)$. Let $\varphi\in\mathcal{H}_\alpha$ be such that $\int\varphi d\mu=0$. Then*

$$\sigma^2=\lim_{n\to\infty}\frac{1}{n}\int\left(\sum_{i=0}^{n-1}\varphi\circ f^i\right)^2 d\mu\ge 0$$

is well defined and in case $\sigma^2>0$ we have for all $a\in\mathbb{R}$

$$\mu\left(\left\{x:\frac{1}{\sqrt{n}}\sum_{i=0}^{n-1}\varphi\circ f^i(x)\le a\right\}\right)\to\int_{-\infty}^{a}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}dx,\ \text{as } n\to\infty.$$

Moreover, $\sigma^2=0$ if and only if φ is a coboundary *($\varphi\neq\psi\circ f-\psi$ for any $\psi\in L^2$).*

This means that, essentially, the fact that the system looses memory sufficiently fast, alone, is enough to guarantee that the deviation of time averages from the spatial average, when properly normalised, is asymptotically normally distributed.

B.2. **Local Limit Theorem.** A function $\varphi: M\to\mathbb{R}$ is said to be *periodic* if there exist $\rho\in\mathbb{R}$, a measuruble $\psi: M\to\mathbb{R}$, $\lambda>0$ and $q: M\to\mathbb{Z}$, such that $\varphi=\rho+\psi-\psi\circ f+\lambda q$ almost everywhere. Otherwise, it is said to be *aperiodic*.

Putting together Theorem A or C and [Go1, Theorem 1.2] we easily get:

Corollary B.2. *Suppose that $Cor_\mu(\varphi,\psi\circ f^n)\lesssim n^{-\beta}$ for some $\beta>3$ for every $\varphi\in\mathcal{H}_\alpha$ and $\psi\in L^\infty(\mu)$. Let $\varphi\in\mathcal{H}_\alpha$ be such that $\int\varphi d\mu=0$ and σ^2 given by Corollary* B.1*. If φ is aperiodic, which implies that $\sigma^2>0$, then for any bounded interval $J\subset\mathbb{R}$, for any real sequence $\{k_n\}_{n\in\mathbb{N}}$ with $k_n/n\to\kappa\in\mathbb{R}$, for any $u\in\mathcal{H}_\alpha$, for any measurable $v: M\to\mathbb{R}$ we have*

$$\sqrt{n}\mu\left(\left\{x\in M:\ \sum_{i=0}^{n-1}\varphi\circ f^i(x)\in J+k_n+u(x)+v(f^nx)\right\}\right)\to m(J)\frac{e^{-\frac{\kappa^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}.$$

This result can be seen as saying that as long as the system has sufficiently fast decay of correlations, the normalised deviation of time averages from the spatial average also behaves locally as Gaussian random variable.

B.3. **Berry-Esseen inequalities.** If f admits a Gibbs-Markov induced map of base Δ_0 and return time function R, then for any $\varphi : M \to \mathbb{R}$ define $\varphi_{\Delta_0} : \Delta_0 \to \mathbb{R}$ by

$$\varphi_{\Delta_0}(x) = \sum_{i=0}^{R(x)-1} \varphi(f^i x).$$

Combining Theorem A or C and [Go1, Theorem 1.3] we obtain:

Corollary B.3. *Suppose that* $Cor_\mu(\varphi, \psi \circ f^n) \lesssim n^{-\beta}$ *for some* $\beta > 3$ *for every* $\varphi \in \mathcal{H}_\alpha$ *and* $\psi \in L^\infty(\mu)$*, which implies the existence of a a Gibbs-Markov induced map of base* Δ_0 *and return time function* R *such that* $m(\mathscr{R}_n) \lesssim n^{-\beta'}$*, with* $\beta' > 2$*. Let* $\varphi \in \mathcal{H}_\alpha$ *be such that* $\int \varphi d\mu = 0$ *and* σ^2 *be given by Corollary* B.1*. Assume that* $\sigma^2 > 0$ *and that there exists* $0 < \delta \leq 1$ *such that* $\int |\varphi_{\Delta_0}|^2 \mathbf{1}_{|\varphi_{\Delta_0}|>z} d\mu \lesssim z^{-\delta}$*, for large* z*. If* $\delta = 1$*, assume also that* $\int |\varphi_{\Delta_0}|^3 \mathbf{1}_{|\varphi_{\Delta_0}|\leq z} d\mu$ *is bounded. Then there exists* $C > 0$ *such that for all* $n \in \mathbb{N}$ *and* $a \in \mathbb{R}$ *we have*

$$\left| \mu\left(\left\{ x : \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \varphi \circ f^i(x) \leq a \right\} \right) - \int_{-\infty}^{a} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \right| \leq \frac{C}{n^{\delta/2}}.$$

The above Corollary shows that sufficiently fast loss of memory not only implies that the normalised deviation of time averages from the spatial average converges in distribution to the normal distribution as it allows to obtain bounds for the speed of convergence.

B.4. **Almost Sure Invariance Principle.** Given $d \geq 1$ and a Hölder continuous $\varphi\colon M \to \mathbb{R}^d$ with 0 mean, we denote $S_n = \sum_{i=0}^{n-1} \varphi \circ f^i$, for each $n \geq 1$. We say that the sequence $\{S_n\}_n$ satisfies a *d-dimensional almost sure invariance principle* (*ASIP*) if there exists $\lambda > 0$ and a probability space supporting a sequence of random variables $\{S_n^*\}_n$ and a d-dimensional Brownian motion $W(t)$ such that

(1) $\{S_n\}_n$ and $\{S_n^*\}_n$ are equally distributed;
(2) $S_n^* = W(n) + O(n^{1/2-\lambda})$, as $n \to \infty$, almost everywhere.

The ASIP is said to be *nondegenerate* if the Brownian motion $W(t)$ has nonsingular covariance matrix Σ.

Putting together Theorem A or C and [MN1, Theorem 1.6] we easily get:

Corollary B.4. *Suppose that* $Cor_\mu(\varphi, \psi \circ f^n) \lesssim n^{-\beta}$ *for some* $\beta > 3$ *for every* $\varphi \in \mathcal{H}_\alpha$ *and* $\psi \in L^\infty(\mu)$*. Let* $\varphi\colon M \to \mathbb{R}^d$ *be a mean* 0 *Hölder continuous observable and* S_n *its time* n *partial sum as above. Then there exists* $\lambda > 0$ *such that*

$$S_n = W(n) + O(n^{1/2-\lambda}),\ a.e.$$

Satisfying an ASIP is a strong statistical property that a stochastic process may have and which implies other limiting laws such as Central Limit Theorem, Functional Central

Limit Theorem, Law of the Iterated Logarithm. As we can see from the result above, it can be verified simply by checking if the system has sufficiently fast decay of correlations.

Appendix C. Piecewise expanding maps

In Theorem B we consider decay of correlations for observables in a Banach space $\mathcal{B}$ against observables in L^1. In Theorem H we show that this holds for systems satisfying some general conditions on the Perron-Frobenius operator. In this appendix we give more explicit examples of dynamical systems satisfying these conditions. As a consequence, we obtain also exponential large deviations for all these systems. We describe here three classes of examples where each one strictly generalizes the previous one.

C.1. **One-dimensional maps.** The first example is given by C^1 piecewise uniformly expanding maps f on the countable partition $\mathfrak{A}$ of the unit interval $M = [0, 1]$, and verifying the Adler property

$$\sup_{A\in\mathfrak{A}} \sup_{x\in A} \frac{|f''(x)|}{(f'x)^2} < \infty.$$

In this case the Lasota-Yorke inequality holds by taking $\mathcal{B}$ as the space of functions φ on the interval with bounded total variation $V_{[0,1]}\varphi$. The corresponding Banach norm will be given by the sum of $V_{[0,1]}\varphi$ plus the $L^1(m)$ norm of φ and this norm is adapted to $L^1(m)$; moreover the Banach space just constructed is an algebra.

Finally, whenever the images under f of the elements in $\mathfrak{A}$ coincide with the whole space $[0, 1]$ (Markovian case), the density of the acip is bounded from below by a strictly positive constant; see e.g. [Br]. In the general non-Markovian situation the positivity of the density will follow whenever the support of the density will be the whole interval (we use here a result by Kowalski [Ko] and Keller [Ke1] which states that if an invariant density ρ is lower semicontinuous, then it admits a constant $a > 0$ such that $\rho|_{\mathrm{supp}_\rho} \geq a$).

C.2. **Markov maps.** The second example generalizes the previous one by relaxing the Markov property. Suppose $\mathfrak{A}$ is a measurable partition of M (not necessarily a Riemannian manifold) endowed with a probability measure m on a σ-algebra $\mathcal{M}$. Let $f : M \to M$ be a measurable map such that

$$f(A) \in \sigma(\mathfrak{A}) \pmod{m}, \quad \text{for all } A \in \mathfrak{A},$$

where $\sigma(\mathfrak{A})$ stands for the σ-algebra generated by $\mathfrak{A}$. We also suppose that $\mathfrak{A}$ generates $\mathcal{M}$ under f in the sense that $\sigma(\bigvee_{n=0}^{\infty} f^{-n}(\mathfrak{A})) = \mathcal{M}$. Assume moreover that $f|_A$ is invertible and nonsingular for all $A \in \mathfrak{A}$. This allows us to define for each $A \in \bigvee_{j=0}^{n-1} f^{-j}(\mathfrak{A})$ the inverse branches $g_{A,n} : f^n(A) \to A$ and the Radon-Nykodym derivaties $\rho_{A,n} = dm \circ g_{A,n}/dm$. We assume the following properties:

(1) mixing: $\forall A, B \in \mathfrak{A}\ \exists n_0 \geq 0 :\ f^n(A) \supset B,\ \forall n \geq n_0$;
(2) big images: $\inf_{A\in\mathfrak{A}} m(fA) > 0$;

(3) bounded distortion: $\exists C>0\ \forall n\geq 1\ \forall A\in\bigvee_{j=0}^{n-1} f^{-j}(\mathfrak{A})\ \forall x,y\in f^n(A)$

$$\left|\frac{\rho_{A,n}(x)}{\rho_{A,n}(y)}-1\right|\leq C\theta^{s(x,y)},$$

where θ is some real number in $(0,1)$ and $s(x,y)$ is the separation time defined as in Definition 1.1. For these systems we consider the functional space of piecewise Lipschitz functions defined in this way: $\varphi: M\to\mathbb{R}$ is Lipschitz on the set $A\subset M$ if the following seminorm is finite

$$D_A\varphi\equiv\sup_{x,y\in A}\frac{|\varphi(x)-\varphi(y)|}{\theta^{s(x,y)}}<\infty.$$

Letting $\mathfrak{B}$ be the partition such that $\sigma(f(\mathfrak{A}))=\sigma(\mathfrak{B})$, we define $D_{\mathfrak{B}}\varphi=\sup_{A\in\mathfrak{B}}D_A\varphi$. Finally we define

$$\mathcal{L}=\{\varphi\in L^\infty(m): D_{\mathfrak{B}}\varphi<\infty\},$$

equipped with the norm

$$\|\varphi\|_{\mathcal{L}}:=\|\varphi\|_{L^\infty(m)}+D_{\mathfrak{B}}\varphi.$$

This norm is adapted to $L^1(m)$.

On the space $\mathcal{L}$ the Perron-Frobenius operator satisfies the Lasota-Yorke inequality and the density of the invariant measure will be m-almost everywhere bounded away from zero; see [AD].

C.3. **Multidimensional maps.** The third interesting example is given by multidimensional piecewise uniformly expanding maps for which we will use the space of quasi-Hölder functions described below. We emphasize that this class of maps generalizes, i.e. contains, those defined above. We follow here the definition proposed by Saussol [Sau]; these maps have also been investigated by Blank [Bl], Buzzi [Bu], Buzzi and Keller [BK] and Tsuji [Ts]; the situation where the expansion is not anymore uniform has been investigated in the paper [HV].

Let $M\subset\mathbb{R}^N$ be a compact subset with $\overline{\operatorname{int} M}=M$ and $f: M\to M$. For $A\subset M$ and $\varepsilon>0$ we put $B_\varepsilon(A)=\{x\in\mathbb{R}^N: d(x,A)\leq\varepsilon\}$, where d be the Euclidean distance in $\mathbb{R}^N$. Assume that there exist at most countably many disjoint open sets U_i such that $m(M\setminus\bigcup_{i=1}U_i)=0$, where m denotes Lebesgue measure in the Borel sets of $\mathbb{R}^N$. Assume moreover that there are open sets $\widetilde{U}_i\supset\overline{U_i}$ and $C^{1+\alpha}$ maps $f_i:\widetilde{U}_i\to\mathbb{R}^N$ such that $f_i|_{U_i}=f|_{U_i}$ for each i. Suppose that there are constants $c,\varepsilon_1>0$ and $0<\alpha<1$ such that the following hold:

(1) $f_i(\widetilde{U}_i)\supset B_{\varepsilon_1}(f(U_i))$ for each i;

(2) for each i and $x,y\in f(U_i)$ with $d(x,y)\leq\varepsilon_1$,

$$\left|\det Df_i^{-1}(x)-\det Df_i^{-1}(y)\right|\leq c|\det Df_i^{-1}(x)|d(x,y)^\alpha;$$

(3) there exists $s=s(f)<1$ such that

$$\sup_i\sup_{x\in f_i(\widetilde{U}_i)}\left\|Df_i^{-1}(x)\right\|<s;$$

(4) each ∂U_i is a codimension one embedded compact C^1 submanifold and

$$s^\alpha + \frac{4s}{1-s} Y(f) \frac{\gamma_{N-1}}{\gamma_N} < 1, \tag{C.1}$$

where $Y(f) = \sup_x \sum_i \#\{\text{smooth pieces intersecting } \partial U_i \text{ containing } x\}$ and γ_N is the volume of the unit ball in $\mathbb{R}^N$.

According to [Sau], condition (C.1) can be weakened. We nevertheless keep that condition which is particularly simple to handle with when the boundaries of the U_i are smooth. Given a Borel set $\Omega \subset M$, we define the oscillation of $\varphi \in L^1(m)$ over Ω as

$$\operatorname{osc}(\varphi, \Omega) := \operatorname{essup} \varphi|_\Omega - \operatorname{essinf} \varphi|_\Omega.$$

Letting $B_\epsilon(x)$ denote the ball of radius ϵ around the point x, we get a measurable function $x \to \operatorname{osc}(\varphi,\ B_\epsilon(x))$. Given $0 < \alpha < 1$ and $\varepsilon_0 > 0$, we define the α-seminorm of φ as

$$|\varphi|_\alpha = \sup_{0<\epsilon\leq\epsilon_0} \epsilon^{-\alpha} \int_{\mathbb{R}^N} \operatorname{osc}(\varphi, B_\epsilon(x)) dm(x). \tag{C.2}$$

We consider the space of the functions with bounded α-seminorm

$$V_\alpha = \left\{\varphi \in L^1(m) : |\varphi|_\alpha < \infty\right\}, \tag{C.3}$$

and equip V_α with the norm

$$\|\cdot\|_\alpha = \|\cdot\|_{L^1(m)} + |\cdot|_\alpha. \tag{C.4}$$

We remark that this space does not depend on the choice of ϵ_0 and V_α is a Banach space endowed with the norm $\|\cdot\|_\alpha$. Moreover, according to Theorem 1.13 in [Ke2], the unit ball in V_α is compact in $L^1(\mu)$.

The assumptions (1)-(4) above allow us to get a Lasota-Yorke inequality when the Perron-Frobenius operator is applied to functions belonging to the space V_α; see [Bl] and [Ke2] for the introduction of such a space in the theory of dynamical systems.

C.4. **Decay of correlations.** Here we prove the part of Corollary H which is still left to prove, namely that f exhibits exponential decay of correlations for observables in $\mathcal{B}$ against $L^1(\mu)$.

It is well known that under conditions (1)-(4) in Section 1.6.2, the Ionescu-Tulcea-Marinescu theorem [IM] asserts that the operator P_m is quasi-compact and this implies the existence of an invariant probability measure μ for the map f which is absolutely continuous with respect to m on M and with density $h \in \mathcal{B}$ (see also [HH] for a simple proof the ITM Theorem with estimate of the essential spectral radius). The measure μ has a finite number of ergodic components, and it is the "unique greatest" in the sense that any other measure absolutely continuous with respect to m is absolutely continuous with respect to μ. Moreover, M is partitioned μ mod 0 into a finite number of measurable sets upon which a certain power of f is mixing. Since we are mostly interested in the rate of decay of correlations, we will suppose that M is the only mixing component for f.

The iterates of the Perron-Frobenius operator enjoy the following spectral decomposition:

$$P_m^n = \Pi + Q^n, \tag{C.5}$$

where Π projects $\varphi \in \mathcal{B}$ into the fixed points of P_m,

$$\Pi(\varphi) = h \int \varphi dm, \tag{C.6}$$

and the linear operator Q verifies

$$\|Q^n(\varphi)\|_{\mathcal{B}} \leq C'' q^n \|\varphi\|_{\mathcal{B}}, \tag{C.7}$$

where $C'' > 0$ and $0 < q < 1$ are constants depending on f.

Now, take $\varphi \in \mathcal{B}$ and assume with no loss of generality that $\int \varphi d\mu = 0$, or equivalently $\int \varphi h dm = 0$, where $h = d\mu/dm$. Since $h \in \mathcal{B}$, by property (5) in Section 1.6.2 we have that $\varphi h \in \mathcal{B}$. Therefore, using (C.5), (C.6), (C.7) and property (6) in Section 1.6.2, for any $\psi \in L^1(m)$ we have

$$\begin{aligned}\left|\int \varphi(\psi \circ f^n) d\mu\right| &= \left|\int \psi P_m^n(\varphi h) dm\right| \\ &\leq \left|\int \psi Q^n(\varphi h) dm\right| \\ &\leq C' \|\psi\|_{L^1(m)} \|Q^n(\varphi h)\|_{\mathcal{B}} \\ &\leq C' C'' q^n ||\psi||_{L^1(m)} ||h\varphi||_{\mathcal{B}}.\end{aligned}$$

Recalling that by assumption there is $c > 0$ such that $h \geq c$, it then follows that

$$\|\psi\|_{L^1(m)} = \int \frac{|\psi|}{h} d\mu \leq \frac{1}{c} \|\psi\|_{L^1(\mu)}.$$

Thus we have

$$\mathrm{Cor}_\mu(\varphi, \psi \circ f^n) \leq \frac{1}{c} C' C'' \|h\|_{\mathcal{B}}\, q^n.$$

To finish, we just need to observe that $L^p(\mu) \subset L^p(m)$ for any $1 \leq p \leq \infty$, as long as $d\mu/dm$ is uniformly bounded away from zero.

Acknowledgments. We thank the referee for their comments and in particular for pointing out Corollary F. We wish to thank Ian Melbourne for very valuable observations and relevant suggestions. We are grateful to Neil Dobbs, Carlangelo Liverani, Matt Nicol and Mike Todd for fruitful conversations and comments.

References

[AD] J Aaronson, M Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps*, Stochastic Dyn., **1**, (2001), 193-237 C.2

[Al1] J.F. Alves, *SRB measures for non-hyperbolic systems with multidimensional expansion*, Ann. Scient. Éc. Norm. Sup. (4) **33** (2000), no. 1, 1-32. 1.6.1

[Al2] J.F. Alves, *Strong statistical stability of non-uniformly expanding maps,* Nonlinearity **17** (2004) 1193-1215. 4

[AA] J. F. Alves, V. Araújo, *Random perturbations of nonuniformly expanding maps.* Astérisque **286** (2003), 25-62. 1.6.1

[ABV] J.F. Alves, C. Bonatti, M. Viana, *SRB measures for partially hyperbolic systems whose central direction is mostly expanding,* Invent. Math. **140** (2000) 351-398. 1.3, 1.6.1

[ADL] J.F. Alves, C.L. Dias, S. Luzzatto, *Geometry of expanding absolutely continuous invariant measures and the liftability problem*, preprint 2009. 1.1, 3

[ALP] J.F. Alves, S. Luzzatto, V. Pinheiro, *Markov structures and decay of correlations for non-uniformly expanding dynamical systems,* Ann. Inst. H. Poincaré Anal. Non Linéaire **22** (2005) 817-839. 1, 1.3, 1.6.1, 1.7, 3, 4, 4, 4

[AV] J. F. Alves, M. Viana, *Statistical stability for robust classes of maps with non-uniform expansion.* Ergodic Theory & Dynam. Systems **22** (2002), no. 1, 1-32. 1.6.1

[AP] V. Araújo, M.J. Pacifico, *Large deviations for non-uniformly expanding maps.* J. Stat. Phys. **125** (2006) 415–457. 1.6.1

[Az] K. Azuma, *Weighted sums of certain dependent random variables.* Tôhoku Math. J. **19** (1967) 357–367. A.2

[Bl] M. Blank, *Stochastic properties of deterministic dynamical systems*, Sov. Sci. Rev. C Maths/Phys., **6**, (1987), 243-271 C.3, C.3

[Br] A. Broise, *Transformations dilatantes de l'intervalle et théorèmes limites*, Asterisque, **238**, (1996), 5-110. C.1

[BLS] H. Bruin, S. Luzzatto. S. van Strien, *Decay of correlations in one-dimensional dynamics,* Ann. Sci. Ecole Norm. Sup. **36** (2003) 621-646. 1

[Bu] J. Buzzi, *Absolutely continuous invariant probability measures for arbitrary expanding piecewise* **R***-analytic mappings of the plane*, Ergodic Theory Dynam. Systems, **20**, (2000), 697-708 C.3

[BK] J. Buzzi, G. Keller, *Zeta functions and transfer operators for multidimensional piecewise affine and expanding maps*, Ergodic Theory Dynam. Systems, **21**, (2001), 689-716 C.3

[BST] J. Buzzi, O. Sester and M. Tsujii, *Weakly expanding skew-products of quadratic maps.* Ergod. Th. & Dynam. Sys. **23** (2003), 1401–1414. 1.6.1

[CCRV] J.-R. Chazottes, P. Collet, F. redig, E. Verbitskiy, *A concentration inequality for interval maps with an indifferent fixed point* Ergodic Theory Dynam. Systems **29** (2009) 1097–1117

[DHL] K. Díaz-Ordaz, M. Holland, S. Luzzatto, *Statistical properties of one-dimensional maps with critical points and singularities*, Stoch. Dyn. **6** (2006), 423-458. 1

[GB] P. Góra, A. Boyarsky, *Laws of Chaos. Invariant measures and Dynamical Systems in one dimension.*, Birkhäuser, 1997. A.1

[Go1] S. Gouëzel, *Berry-Esseen theorem and local limit theorm for nonuniformly expanding maps,* Ann. I.H. Poincaré, **41** (2005) 997–1024. 1.5, B.2, B.3

[Go] S. Gouëzel, *Decay of correlations for nonuniformly expanding systems,* Bull. Soc. Math. France **134** (2006) 1–31. 1, 1.3, 1.6.1, 1.7, 3, 4, 4, 4

[Go2] S. Gouëzel, *Sharp polynomial estimates for the decay of correlations,* Israel Journal of Mathematics, **139(1)** (2004) 29-65. 1.2

[HH] H. Hennion, L. Hervé, Limit theorems for Markov chains and Stochastic Properties of Dynamical Systems by Quasicompactness, *Letc. Notes Math.*, **1766**, (2001), Springer-Verlag C.4

[Hoe] W. Hoeffding, *Probability inequalities for sums of bounded random variables.* J. Amer. Statist. Assoc. **58** (1963), 13–30. A.2

[Hol] M. Holland, *Slowly mixing systems and intermittency maps*, Ergodic Theory Dynam. Systems **25** (2005), 133-159. 1, 1.6.3

[Hu] H. Hu, *Decay of correlations for piecewise smooth maps with indifferent fixed points.* Ergodic Theory Dynam. Systems **24** (2004), no. 2, 495–524. 1.6.3, 1.6.3

[HV] H. Hu, S. Vaienti, *Absolutely continuos invariant measures for non-uniformly expanding maps*, Ergodic Theory Dynam. Systems , **29**, (2009), 1185-1215 C.3

[IM] C. Ionescu-Tulcea, G. Marinescu, *Théorie ergodique pour des classes d'opératations non complétement continues*, Ann. Math., **52**, (1950), 140-147 C.4

[Ke1] G. Keller, *Piecewise monotonic transformations and exactness*, Collection : Seminar on Probability, Rennes, 1978 (French) ; Exp. **6**, 32 pp. Univ. Rennes, (1978) C.1

[Ke2] G. Keller, *Generalized bounded variation and applications to piecewise monotonic transformations*, Z. Wahr. verw. Geb., **69**, (1985), 461-478 C.3

[Ko] Z.S. Kowalski, *Invariant measure for piecewise monotonic trasformations has a lower bound on its support*, Bull Acad. Pol. Sci. Math., **27**, (1979), 53-57 C.1

[LV] E. Lesigne, D. Volný. *Large deviations for martingales.* Stoch. Proc. Applns. **96** (2001) 143–159. A.2

[LSV] C. Liverani, B. Saussol, S. Vaienti, *A probabilistic approach to intermittency.* Ergodic Theory Dynam. Systems **19** (1999), no. 3, 671–685. 1.6.3

[Me] I. Melbourne. *Large and moderate deviations for slowly mixing dynamical systems*, Proc. Amer. Math. Soc. 137 (2009) 1735-1741 1, 1.4, 1.6.1, 1.7, 2.1

[MN1] I. Melbourne, M. Nicol. *Almost Sure Invariance Principle for Nonuniformly Hyperbolic Systems*, Comm. Math. Phys. 260 (2005) 131-1456 1.5, B.4

[MN] I. Melbourne, M. Nicol. *Large deviations for nonuniformly hyperbolic systems,* Trans. Amer. Math. Soc. 360 (2008) 6661-6676. 1, 1.6.1, 2.1

[MN2] I. Melbourne, M. Nicol. *A vector-valued almost sure invariance principle for hyperbolic dynamical systems,* The Annals of Probability, **37(2)** (2009) 478-505. 1.5

[MPU] F. Merlevéde, M. Peligrad, S. Utev. *Recent advances in invariance principles for stationary sequences.* Probab. Surv. **3** (2006) 1–36 (electronic). 2.1

[Pia] G.Pianigiani, *First return map and invariant measures.* Israel J. Math. **35** (1980), 32–48 1.6.3

[PM] Y. Pomeau, P. Manneville, *Intermittent transition to turbulence in dissipative dynamical systems.* Comm. Math. Phys. **74** (1980), no. 2, 189–197. 1.6.3

[PS] M. Pollicott, R. Sharp, *Large deviations for intermittent maps.* Nonlinearity **22** (2009), 2079–2092 1.6.3

[PY] Statistical Properties of maps with indifferent fixed points, *Comm. Math. Phys.* **217** (2001), 503–520. 1.6.3

[RY] L. Rey-Bellet, L.S. Young, *Large deviations in non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems **28** (2008) 587–612. 1, 1.6.1

[Ri] E. Rio, *Théorie asymptotique des processus aléatoires faiblement dépendants.* Mathématiques & Applications (Berlin) [Mathematics & Applications] 31, Springer-Verlag, Berlin, 2000. 2.1

[Sar] O. Sarig, *Subexponential decay of correlations.* Invent. Math. **150** (2002), no. 3, 629–653. 1.2, 1.6.3, 1.6.3

[Sau] B. Saussol, *Absolutely continuous invariant measures for multidimensional expanding maps*, Israel J. Math., **116**, (2000), 223-248. C.3, C.3

[Ts] M. Tsujii, *Absolutely continuous invariant measures for expanding piecewise linear maps*, Invent. Math., **143**, (2001), 349-373 C.3

[Vi] M. Viana, *Multidimensional nonhyperbolic attractors.* Inst. Hautes Études Sci. Publ. Math. No. **85** (1997), 63–96. 1.6.1

[Yo1] L.S. Young, *Statistical properties of dynamical systems with some hyperbolicity,* Ann. of Math. (2) **147** (1998) 585–650. 1

[Yo2] L.S. Young, *Recurrence times and rates of mixing,* Israel J. Math. **110** (1999) 153–188. 1, 1.2, 1.2, 1.5, 1.6.3, 1.6.3, B.1

José F. Alves, Departamento de Matematica, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal
E-mail address: jfalves@fc.up.pt
URL: http://www.fc.up.pt/cmup/jfalves

Jorge M. Freitas, Departamento de Matematica, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal
E-mail address: jmfreita@fc.up.pt
URL: http://www.fc.up.pt/pessoas/jmfreita

Stefano Luzzatto, Mathematics Department, Imperial College, 180 Queen's Gate, London SW7, UK
E-mail address: luzzatto@ictp.it
URL: http://www.ictp.it/∼luzzatto
Current address: Abdus Salam International Centre for Theoretical Physics, Strada Costiera 11, 34151 Trieste, Italy.

Sandro Vaienti, UMR-6207 Centre de Physique Théorique, CNRS, Universités d'Aix-Marseille I, II, Université du Sud, Toulon-Var and FRUMAM, Fédéderation de Recherche des Unités de Mathématiques de Marseille, CPT, Luminy Case 907, F-13288 Marseille Cedex 9, France
E-mail address: vaienti@cpt.univ-mrs.fr