Almost Sure Invariance Principle for Nonuniformly Hyperbolic Systems

Ian Melbourne
Department of Maths and Stats
University of Surrey
Guildford GU2 7XH, UK

Matthew Nicol
Department of Maths
University of Houston
Houston TX 77204-3008, USA

23 September, 2004. Revised 8 February, 2005.

Abstract

We prove an almost sure invariance principle that is valid for general classes of nonuniformly expanding and nonuniformly hyperbolic dynamical systems. Discrete time systems and flows are covered by this result. In particular, the result applies to the planar periodic Lorentz flow with finite horizon.

Statistical limit laws such as the central limit theorem, the law of the iterated logarithm, and their functional versions, are immediate consequences.

1 Introduction

Statistical properties of uniformly expanding maps and uniformly hyperbolic (Axiom A) diffeomorphisms are by now classical. Hölder observations satisfy exponential decay of correlations and the central limit theorem (CLT), see for example Bowen [8], Ratner [34], Ruelle [35], Parry and Pollicott [31]. Furthermore, Denker and Philipp [17] proved an almost sure invariance principle (ASIP) for Hölder observations. Immediate consequences of the ASIP are the CLT, the law of the iterated logarithm (LIL), and their functional versions, see [32].

Many proofs of the CLT for dynamical systems use directly the martingale approximation method of Gordin [20], see [25, 26, 30]. The ASIP can often also be obtained in this way see [16, 19, 30, 39] and indeed this method yields a better error estimate in the ASIP than the usual one, see Field *et al.* [19]. However, it should be emphasised that the martingale approximation of Gordin [20] leads directly only to a *reverse* martingale increment sequence and so the ASIP is obtained in backwards time in the first instance. This is not an issue for distributional results such as the CLT, but the ASIP in [16, 19, 30] uses explicitly the fact that the class of systems being studied is closed under time-reversal. To obtain *forward* martingale approximations, it is necessary to use more sophisticated versions of Gordin's approach [32].

Recently, there has been an explosion of interest in nonuniformly expanding maps and nonuniformly hyperbolic diffeomorphisms (possibly with singularities). We refer to the articles of Young [40, 41] as well as Aaronson [1], Baladi [4, 5], Gouëzel [22], Viana [38] and references therein. In particular, decay of correlations and the CLT are studied extensively in these references. However, such classes of dynamical systems are intrinsically time-orientation specific, and largely for this reason the ASIP has not previously been proved. Similarly, the LIL was previously unproved for such systems.

In this paper, we establish the ASIP, and hence the (functional) LIL, for nonuniformly expanding/hyperbolic systems. Both discrete time systems and flows are covered by our results.

Remark 1.1 We note that [33] attempted to apply the approach in [19] to nonuniformly expanding systems. However, it appears that the time-orientation issue discussed above was overlooked in [33], and that this is a gap. Hence it seems necessary to find an alternative approach to the one in [19], and that is what is done in the current paper.

Precise formulations are given in the body of the paper, but here is an outline of our main result, and the strategy behind its proof, for a nonuniformly expanding map $T : M \to M$ where (M, d) is a metric space. By standard methods, T can be modelled by a discrete-time suspension over a *Gibbs-Markov* map [1] $f : Y \to Y$ with return time function $R : Y \to \mathbb{Z}^+$. (Roughly speaking, a Gibbs-Markov map is like a uniformly expanding map with possibly countably many inverse branches.) There exists a unique ergodic T-invariant probability measure equivalent to Lebesgue, and the following result is formulated with this measure in mind.

Theorem 1.2 *Let $T : M \to M$ be a nonuniformly expanding map. Assume moreover that $R \in L^{2+\delta}(Y)$. Let $\phi : M \to \mathbb{R}$ be a mean zero Hölder observation. Then ϕ satisfies the ASIP. That is, there exists $\epsilon > 0$, a sequence of random variables $\{S_n\}$ and a Brownian motion W with variance $\sigma^2 \geq 0$ such that $\{\sum_{j=0}^{N-1} \phi \circ T^j\} =_d \{S_N\}$, and*

$$S_N = W(N) + O(N^{\frac{1}{2}-\epsilon}) \quad \text{as } N \to \infty,$$

almost everywhere.

Using a method due to Hofbauer and Keller [24] which exploits a result of Philipp and Stout [32, Theorem 7.1], we obtain the ASIP (in the correct time direction but without the improved error term) for $f : Y \to Y$ and a class of "weighted Lipschitz" observations. Theorem 1.2 then follows directly by Melbourne and Török [29]. (We note that the method in [29] has independently been used by Gouëzel [21] to obtain a simplified derivation of the CLT and stable laws.)

A precise version of Theorem 1.2 is stated and proved in Section 2(e). The ASIP for nonuniformly hyperbolic maps extends easily to a class of nonuniformly expanding semiflows, see Section 2(e).

Our results for nonuniformly hyperbolic diffeomorphisms and nonuniformly hyperbolic flows are completely analogous, but the set-up is more technical and we postpone further details until Section 3.

Planar periodic Lorentz gas The planar periodic Lorentz gas is a class of examples introduced by Sinaĭ [36]. See [15] for a survey of results about Lorentz gases. The Lorentz flow is a billiard flow on $\mathbb{T}^2 - \Omega$ where Ω is a disjoint union of convex regions with C^3 boundaries. (The phase-space of the flow is three-dimensional; planar position and direction.) The flow has a natural global cross-section $M = \partial\Omega \times [-\pi/2, \pi/2]$ corresponding to collisions and the Poincaré map $T : M \to M$ is called the billiard map. Bunimovich, Sinaĭ and Chernov [11] proved the central limit theorem and weak invariance principle for such maps.

Denote the return time function by $h : M \to \mathbb{R}^+$. The Lorentz flow satisfies the *finite horizon* condition if h is uniformly bounded. The central limit theorem and weak invariance principle was proved by [11] for Lorentz flows satisfying the finite horizon condition.

Theorem 1.3 *Suppose that T_t is a planar periodic Lorentz gas.*

(i) The billiard map satisfies the ASIP for Hölder observations.

(ii) If the finite horizon condition holds, then the Lorentz flow satisfies the ASIP for Hölder observations.

In Section 2, we prove the ASIP for nonuniformly expanding maps and semiflows. In Section 3, we prove the analogous results for systems that are nonuniformly hyperbolic in the sense of Young [40]. In Section 4, we list numerous examples in the literature for which our results apply. In particular, we prove Theorem 1.3. The results of [32, 29] required in this paper are reproduced as appendices.

2 Nonuniformly expanding systems

In this section, we prove the ASIP for nonuniformly expanding systems. The first step is to prove the ASIP for Gibbs-Markov maps. Such maps are reviewed in Subsection (a) and a class of "weighted Lipschitz" observations is introduced in Subsection (b). The ASIP for Gibbs-Markov maps is proved in Subsection (c) using an approach of Hofbauer and Keller [24]. In Subsection (d), we obtain the ASIP for Young towers [41] as an application of [29]. In Subsection (e), we prove the ASIP for nonuniformly expanding maps and semiflows.

(a) Gibbs-Markov maps

Let (Λ, m) be a Lebesgue space with a countable measurable partition α. Without loss, we suppose that all partition elements $a \in \alpha$ have $m(a) > 0$. Recall that a measure-preserving transformation $f : \Lambda \to \Lambda$ is a *Markov map* if $f(a)$ is a union of elements of α and $f|_a$ is injective for all $a \in \alpha$. Define α' to be the coarsest partition of Λ such that fa is a union of atoms in α' for all $a \in \alpha$. (So α' is a coarser partition than α.) If $a_0, \ldots, a_{n-1} \in \alpha$, we define the n-cylinder $[a_0, \ldots, a_{n-1}] = \cap_{i=0}^{n-1} f^{-i} a_i$. It is assumed that f and α separate points in Λ (if $x, y \in \Lambda$ and $x \neq y$, then for n large enough there exist distinct n-cylinders that contain x and y).

Let $0 < \beta < 1$. We define a metric d_β on Λ by $d_\beta(x, y) = \beta^{s(x,y)}$ where $s(x, y)$ is the greatest integer $n \geq 0$ such that x, y lie in the same n-cylinder. Define $g = Jf^{-1} = \frac{dm}{d(m \circ f)}$ and set $g_k = g\, g \circ f \cdots g \circ f^{k-1}$.

The map $f : \Lambda \to \Lambda$ is a *Gibbs-Markov map* if it satisfies the additional properties:

(i) *Big images property:* There exists $c > 0$ such that $m(fa) \geq c$ for all $a \in \alpha$.

(ii) *Distortion:* $\log g|_a$ is Lipschitz with respect to d_β for all $a \in \alpha'$.

It follows from assumptions (i) and (ii) that there exists a constant $D \geq 1$ such that for all x, y lying in a common k-cylinder $[a_0, \ldots, a_{k-1}]$,

$$\Big|\frac{g_k(x)}{g_k(y)} - 1\Big| \leq D d_\beta(f^k x, f^k y) \quad \text{and} \quad D^{-1} \leq \frac{m[a_0, \ldots, a_{k-1}]}{g_k(x)} \leq D. \tag{2.1}$$

(b) Weighted Lipschitz observations

Let $p \geq 1$. We fix a sequence of weights $R(a) > 0$ satisfying $|R|_p = (\sum_{a \in \alpha} m(a) R(a)^p)^{1/p} < \infty$.

Given $v : \Lambda \to \mathbb{R}$ continuous, we set $v_a = v|_a$ and define $|v|_\beta$ to be the Lipschitz constant of v with respect to the metric d_β. Let

$$\|v\|_\infty = \sup_{a \in \alpha} |v_a|_\infty / R(a), \qquad \|v\|_\beta = \sup_{a \in \alpha} |v_a|_\beta / R(a).$$

Let $\mathcal{B}$ consist of the space of weighted Lipschitz functions with $\|v\| = \|v\|_\infty + \|v\|_\beta < \infty$. Note in particular that $R \in \mathcal{B}$ and $\|R\| = 1$. We have the embeddings

$$\mathrm{Lip} \subset \mathcal{B} \subset L^p \subset L^1,$$

where Lip is the space of (globally) Lipschitz functions.

The transfer (Perron-Frobenius) operator $P : L^1 \to L^1$ maps $v \in L^1$ to Pv where $\int_\Lambda Pv\, w\, dm = \int_\Lambda v\, w \circ f\, dm$ for all $w \in L^\infty$, and is given by $(Pv)(x) = \sum_{fy=x} g(y)v(y)$. Note that $|P|_1 = 1$.

Proposition 2.1 *Let $a \in \alpha$ be an n-cylinder and suppose that $v : a \to \mathbb{R}$ is Lipschitz. Then $|v|_\infty \le \frac{1}{m(a)} \int_a |v|\, dm + \beta^n |v|_\beta$.*

Proof For $x \in a$,

$$|v(x)| \le \tfrac{1}{m(a)} \textstyle\int_a |v|\, dm + |v(x) - \tfrac{1}{m(a)} \int_a v\, dm| \le \tfrac{1}{m(a)} \int_a |v|\, dm + |v|_\beta \operatorname{diam}(a).$$

The result follows since $\operatorname{diam}(a) = \beta^n$. ∎

Lemma 2.2 *The transfer operator P restricts to an operator $P : \mathcal{B} \to \mathcal{B}$ and there exists a constant $C \ge 1$ such that*

$$\|P^n v\| \le C(|v|_1 + \beta^n \|v\|_\beta),$$

for all $v \in \mathcal{B}$ and $n \ge 1$. Moreover, $P(\mathcal{B}) \subset \mathrm{Lip}$.

Proof We prove the estimate on $\|P^n v\|$. The remaining statements of the lemma are evident from the proof.

Note that $(P^n v)(x) = \sum_{f^n y = x} g_n(y)v(y)$. Since the n-cylinders $[a_0, \dots, a_{n-1}]$ form a partition and each n-cylinder contains precisely one preimage y_a, we have

$$\begin{aligned} |(P^n v)(x)| &\le \sum_{a=[a_0,\dots,a_{n-1}]} g_n(y_a)|v(y_a)| \le D \sum_{a=[a_0,\dots,a_{n-1}]} m(a)|v_a|_\infty \\ &\le D \sum_{a=[a_0,\dots,a_{n-1}]} \Big[\int_a |v|\, dm + m(a)\beta^n |v_a|_\beta \Big] \\ &\le D \sum_{a=[a_0,\dots,a_{n-1}]} \Big[\int_a |v|\, dm + \beta^n m(a) R(a_0) \|v\|_\beta \Big] \end{aligned}$$

where we have used Proposition 2.1 and estimate (2.1). Hence $|P^n v|_\infty \le D\Big[|v|_1 + \beta^n |R|_1 \|v\|_\beta\Big]$. Similarly,

$$\begin{aligned} |(P^n v)(x) - (P^n v)(x')| \le &\sum_{a=[a_0,\dots,a_{n-1}]} |g_n(y_a) - g_n(y'_a)||v(y_a)| \\ &+ \sum_{a=[a_0,\dots,a_{n-1}]} |g_n(y'_a)||v(y_a) - v(y'_a)|. \end{aligned}$$

Each term in the first summation can be estimated by

$$\begin{aligned} D|g_n(y_a')|d_\beta(f^n y_a, f^n y_a')|v_a|_\infty &\le D^2 m(a) d_\beta(x,x')\big[\tfrac{1}{m(a)}\textstyle\int_a |v_a|\,dm + \beta^n|v_a|_\beta\big] \\ &\le D^2\Big[\textstyle\int_a |v|\,dm + \beta^n m(a)R(a_0)\|v\|_\beta\Big] d_\beta(x,x), \end{aligned}$$

so the first summation is bounded by $D^2\big[|v|_1 + \beta^n|R|_1\|v\|_\beta\big]d_\beta(x,x')$. Each term in the second summation can be estimated by

$$Dm(a)|v_a|_\beta d_\beta(y_a,y_a') \le Dm(a)R(a_0)\|v\|_\beta \beta^n d_\beta(x,x'),$$

so the second summation is bounded by $\beta^n D|R|_1\|v\|_\beta d_\beta(x,x')$. The result follows. ∎

We have the following standard consequences of Lemma 2.2.

Corollary 2.3 *Let $p,q \ge 1$ with $\frac{1}{p}+\frac{1}{q}=1$. Assume that $f:\Lambda\to\Lambda$ is mixing and that $R \in L^p$. Then there exist constants $C \ge 1$ and $\tau \in (0,1)$ such that*

(a) $\|P^n v - \int_\Lambda v\,dm\| \le C\tau^n\|v\|$ for all $v \in \mathcal{B}$ and $n \ge 1$.

(b) $|\int_\Lambda v\,(w\circ f^n)\,dm - \int_\Lambda v\,dm\int_\Lambda w\,dm| \le C\tau^n\|v\||w|_q$ for all $v\in\mathcal{B}$, $w \in L^q$, $n\ge 1$.

(c) If $R\in L^2$, then for any $v\in\mathcal{B}$ with $\int_\Lambda v\,dm = 0$, the series

$$\sigma^2 = \int_\Lambda v^2\,dm + 2\sum_{k=1}^\infty \int_\Lambda v\,(v\circ f^k)\,dm,$$

is absolutely convergent, and $\int_\Lambda v_N^2\,dm = \sigma^2 N + O(1)$ as $N\to\infty$, where $v_N = \sum_{j=0}^{N-1} v\circ f^j$. Moreover, $\sigma = 0$ if and only if there exists a Lipschitz function $w:\Lambda\to\mathbb{R}$ such that $v = w\circ f - w$.

Proof Most of this result is completely standard, but we include the details for completeness. By an Arzela-Ascoli argument, the unit ball in $\mathcal{B}$ is compact in L^1. This combined with Lemma 2.2 implies, by Hennion [23], that the essential spectral radius of $P:\mathcal{B}\to\mathcal{B}$ is bounded above by $\beta<1$. There is a simple eigenvalue at 1 with eigenspace consisting of constant functions, but the mixing assumption guarantees that there are no further eigenvalues on the unit circle. Now choose $\tau\in(\beta,1)$ such that all eigenvalues of P other than 1 lie strictly inside the disk of radius τ. Part (a) follows for such a choice of τ.

To prove part (b), compute that

$$\begin{aligned} |\textstyle\int_\Lambda v\,(w\circ f^n)\,dm - \int_\Lambda v\,dm\int_\Lambda w\,dm| &= |\textstyle\int_\Lambda (P^n v - \int_\Lambda v)\,w\,dm| \le |P^n v - \int_\Lambda v|_p|w|_q \\ &\le \|P^n v - \textstyle\int_\Lambda v\||w|_q \le C\tau^n\|v\||w|_q. \end{aligned}$$

It follows from (b) that $|\int_\Lambda v\,(v\circ f^k)\,dm| \le C\tau^n\|v\||v|_2$ and so the series for σ^2 converges absolutely. Moreover

$$\begin{aligned}\int_\Lambda v_N^2\,dm &= N\int_\Lambda v^2\,dm + 2\sum_{0\le i<j\le N-1}\int_\Lambda v\,(v\circ f^{j-i})\,dm\\ &= N\int_\Lambda v^2\,dm + 2\sum_{k=1}^{N}(N-k)\int_\Lambda v\,(v\circ f^k)\,dm\\ &= N\sigma^2 - 2\sum_{k=1}^{N}k\int_\Lambda v\,(v\circ f^k)\,dm - 2\sum_{k=N+1}^{\infty}N\int_\Lambda v\,(v\circ f^k)\,dm\\ &= N\sigma^2 + O(1),\end{aligned}$$

proving (c).

The criterion for $\sigma = 0$ follows as in [19, 28]. If $v = w\circ f - w$, then $v_N = w\circ f^N - w$ so it is clear that $\sigma = 0$. To prove the converse, define $w = \sum_{j=1}^\infty P^j v$. This series converges in $\mathcal{B}$ by (b) and is Lipschitz by Lemma 2.2. Write $v = \widehat{v} + w\circ f - w$. Then it is easily seen that $\widehat{v}$ has the same variance as v and that $P\widehat{v} = 0$. Hence $\sigma^2 = \int_\Lambda \widehat{v}^2\,dm$, so if $\sigma = 0$, then $\widehat{v} = 0$ and $v = w\circ f - w$. ∎

(c) ASIP for Gibbs-Markov maps

Let α_0^{k-1} denote the partition into length k cylinders $a = [a_0,\dots,a_{k-1}]$.

Lemma 2.4 *Assume that $f:\Lambda\to\Lambda$ is mixing and that $R\in L^{2+\delta}$ for some $\delta > 0$. Let $v\in\mathcal{B}$ with $\int_\Lambda v\,dm = 0$. Then*

(a) $\sum_{a\in\alpha_0^{k-1}}\int_a |v - \frac{1}{m(a)}\int_a v\,dm|^{2+\delta}\,dm \le \big(\|v\|_\beta\,|R|_{2+\delta}\,\beta^k\big)^{2+\delta}$.

(b) $\big|m(a\cap f^{-(N+k)}(b)) - m(a)m(b)\big| \le C\tau^N m(a)m(b)^{1/2}$ *for all $a\in\alpha_0^{k-1}$ and all measurable sets b.*

Proof Note that $|v - \frac{1}{m(a)}\int_a v\,dm| \le |v_a|_\beta\,\mathrm{diam}(a) \le \|v\|_\beta R(a_0)\beta^k$. Part (a) follows immediately.

We argue as in Aaronson & Denker [2] to establish (b). Let $v_{a,k} = P^k\chi_a$. By definition, $v_{a,k} = \sum_{f^k y = x} g_k(y)\chi_a(y) = g_k(y_a)$ where y_a is the unique point in a such that $f^k y_a = x$. Hence by (2.1),

$$|v_{a,k}(x) - v_{a,k}(x')| \le D|g_k(y_a)|d_\beta(x,x') \le D^2 m(a)d_\beta(x,x').$$

It follows that $\|v_{a,k}\| \le Em(a)$ where $E = D^2 + D$.

Using this estimate and Corollary 2.3(b), we compute that

$$\begin{aligned}\big|m(a\cap f^{-(N+k)}b)-m(a)m(b)\big| &= \big|\textstyle\int P^k\chi_a\,\chi_b\circ f^N-\int P^k\chi_a\int\chi_b\big| \\ &= \big|\textstyle\int v_{a,k}\,\chi_b\circ f^N-\int v_{a,k}\int\chi_b\big| \le C\tau^N\|v_{a,k}\|\,|\chi_b|_2\le CE\tau^N m(a)m(b)^{1/2},\end{aligned}$$

as required. ∎

Corollary 2.5 *Let $f:\Lambda\to\Lambda$ be an ergodic Gibbs-Markov map. Define the Banach space $\mathcal{B}$ corresponding to weights $R\in L^{2+\delta}$ for some $\delta>0$. Suppose that $v\in\mathcal{B}$ and $\int_\Lambda v\,dm=0$. Define σ^2 as in Corollary 2.3 and assume that $\sigma^2>0$. Then $v_N=\sum_{j=0}^{N-1}v\circ f^j$ satisfies the ASIP.*

Proof We verify the hypotheses of Philipp & Stout [32, Theorem 7.1]. For convenience, we have translated this theorem into dynamical systems terminology in the appendix, see Theorem A.1. Condition (i) of Theorem A.1 is automatic since $\mathcal{B}\subset L^{2+\delta}$ and condition (ii) follows from Corollary 2.3(c). Conditions (iii) and (iv) follow from parts (a) and (b) of Lemma 2.4. ∎

(d) ASIP for tower maps

Suppose that (Λ,m) is a probability space and that $f:\Lambda\to\Lambda$ is a measure-preserving transformation. Let $R:\Lambda\to\mathbb{Z}^+$ be a measurable function (called a *return time function* with $R\in L^1(\Lambda)$. Define the suspension

$$\Delta=\{(x,\ell)\in\Lambda\times\mathbb{N}:0\le\ell\le R(x)\}/\sim,$$

where $(x,R(x))\sim(f(x),0)$. Define $F:\Delta\to\Delta$ by $F(x,\ell)=(x,\ell+1)$ computed subject to identifications. Note in particular that $F(x,0)=(f(x),0)$. An F-invariant probability measure on Δ is given by $m^R=m\times l/\overline{R}$ where $\overline{R}=\int_\Lambda R\,dm$ and l is counting measure on $\mathbb{N}$.

Let $\{\Delta_{j,0}\}$ be a countable measurable partition of Λ such that f and $\{\Delta_{j,0}\}$ separate points in Λ, and for each j, $R_j=R|_{\Delta_{j,0}}$ is constant and $f:\Delta_{j,0}\to\Lambda$ is a measurable isomorphism. For each j and $0\le\ell<R_j$, let $\Delta_{j,\ell}=\Delta_{j,0}\times\{\ell\}$. This defines a partition $\{\Delta_{j,\ell}\}$ of Δ.

A *separation time* function $s:\Delta\times\Delta\to\mathbb{N}$ is defined as follows: If x,y lie in distinct partition elements, then $s(x,y)=0$. If $x,y\in\Delta_{j,0}$ for some j, then $s(x,y)$ is the greatest integer $n\ge0$ such that f^kx and f^ky lie in the same partition element of Λ for $k=0,\dots,n$. If $x,y\in\Delta_{j,\ell}$, then write $x=F^\ell x_0$, $y=F^\ell y_0$ where $x_0,y_0\in\Delta_{j,0}$ and define $s(x,y)=s(x_0,y_0)$. For $\theta\in(0,1)$, we define a metric d_θ on Δ by setting $d_\theta(x,y)=\theta^{s(x,y)}$.

Definition 2.6 The suspension $F : \Delta \to \Delta$ is called a *Young tower* if $f : \Lambda \to \Lambda$ is a Gibbs-Markov map with respect to the partition $\alpha = \{\Delta_{j,0}\}$.

Remark 2.7 The big images condition for f to be a Gibbs-Markov map is automatically satisfied in the strong sense that $f(a) = \Lambda$ for each each $a \in \alpha$. Hence, $F : \Delta \to \Delta$ is a Young tower provided the distortion condition holds: there exist constants $\theta \in (0,1)$ and $C \geq 1$ such that for each j the Jacobian $g_j = Jf|_{\Delta_{j,0}} : \Delta_{j,0} \to \Lambda$ satisfies $|\log g_j(x) - \log g_j(y)| \leq C d_\theta(x,y)$ for all $x, y \in \Delta_{j,0}$.

Theorem 2.8 *Let $F : \Delta \to \Delta$ be a Young tower defined as a suspension over $f : \Lambda \to \Lambda$ with return time function R. Assume that $R \in L^{2+\delta}(\Lambda)$. Let $\phi : \Delta \to \mathbb{R}$ be a mean zero observation and assume that ϕ is Lipschitz with respect to d_θ. Then $\phi_N = \sum_{j=0}^{N-1} \phi \circ F^j$ satisfies the ASIP.*

Proof Define a mean zero observation $\Phi : \Lambda \to \mathbb{R}$ by setting $\Phi(x) = \sum_{j=0}^{R(x)-1} \phi(x,j)$. Since ϕ is Lipschitz, it is immediate that Φ lies in the space $\mathcal{B}$ of weighted Lipschitz observations. Since $R \in L^{2+\delta}(\Lambda)$, it follows from Corollary 2.5 that $\Phi_N = \sum_{j=0}^{N-1} \Phi \circ f^j$ satisfies the ASIP on Λ.

Note that $R - \overline{R}$ also satisfies the hypotheses of Corollary 2.5, and so the ASIP, and hence the LIL, applies. Therefore, it is certainly the case that $\sum_{j=0}^{N-1} R \circ f^j = N\overline{R} + o(N^{1-\delta})$ almost everywhere. The result follows from [29, Theorem 4.2], see Corollary B.2. ∎

(e) ASIP for nonuniformly expanding systems

Let (M,d) be a locally compact separable bounded metric space with Borel probability measure η and let $T : M \to M$ be a nonsingular transformation for which η is ergodic. Let $Y \subset M$ be a measurable subset with $\eta(Y) > 0$. We suppose that there is an at most countable measurable partition $\{Y_j\}$ with $\eta(Y_j) > 0$, and that there exist integers $R_j \geq 1$, and constants $\lambda > 1$; $C, D > 0$ and $\gamma \in (0,1)$ such that for all j,

(1) $T^{R_j} : Y_j \to Y$ is a (measure-theoretic) bijection.

(2) $d(T^{R_j}x, T^{R_j}y) \geq \lambda d(x,y)$ for all $x, y \in Y_j$.

(3) $d(T^k x, T^k y) \leq C d(T^{R_j}x, T^{R_j}y)$ for all $x, y \in Y_j$, $k < R_j$.

(4) $g_j = \frac{d(\eta|_{Y_j} \circ (T^{R_j})^{-1})}{d\eta|_Y}$ satisfies $|\log g_j(x) - \log g_j(y)| \leq D d(x,y)^\gamma$ for almost all $x, y \in Y$.

(5) $\sum_j R_j \eta(Y_j) < \infty$.

We say that a dynamical system T satisfying (1)–(5) is *nonuniformly expanding*.

Define the *return time function* $R : Y \to \mathbb{Z}^+$ by $R|_{Y_j} \equiv R_j$. Condition (5) says that $\int_Y R\, d\eta < \infty$. The map $f : Y \to Y$ given by $f(y) = T^{R(y)}(y)$ is the corresponding *induced map*. It can be shown (see Young [41, Theorem 1]) that there is a unique invariant probability measure m on M that is equivalent to η.

We can now state and prove a precise version of Theorem 1.2.

Theorem 2.9 *Let $T : M \to M$ be a nonuniformly expanding map satisfying (1)–(5) above. Assume moreover that the return time function R lies in $L^{2+\delta}(Y)$. Let $\phi : M \to \mathbb{R}$ be a mean zero Hölder observation. Then ϕ satisfies the ASIP.*

Proof Let $\Delta = \{(y, \ell) : y \in Y,\ \ell = 0, \dots, R(y) - 1\}$, so Δ is the disjoint union of R_j copies of each Y_j. Define a measure μ on Δ by setting $\mu|_{Y_j \times \{\ell\}} = m|_{Y_j}/\bar{R}$. Define $F : \Delta \to \Delta$ by setting $F(y, \ell) = (y, \ell + 1)$ for $0 \le \ell < R(y) - 1$ and $F(y, R(y) - 1) = (fy, 0)$. Define the separation time $s : \Delta \times \Delta \to \mathbb{N}$ as in the previous section.

By shrinking γ if necessary, we may suppose that ϕ is γ-Hölder for the same γ that appears in condition (4). Define the metric d_θ on Δ with $\theta = 1/\lambda^\gamma$. It follows from condition (2) that $d(x, y) \le \operatorname{diam}(Y)/\lambda^{s(x,y)}$ for all $(x, y) \in \Delta$. Hence f and $\{Y_j\}$ separate points in Y and the required distortion condition on g_j is immediate, so Δ is a Young tower with $\Lambda = Y$ and $\Delta_{j,0} = Y_j$.

If x, y lie in the same partition element of Δ_ℓ, then write $x = F^\ell x_0$, $y = F^\ell y_0$ so $d(fx_0, fy_0) \le \operatorname{diam}(Y)/\lambda^{s(x,y)}$. By condition (3),

$$d(x, y) \le C d(fx_0, fy_0) \le C \operatorname{diam}(Y)/\lambda^{s(x,y)} = C \operatorname{diam}(Y)[d_\theta(x, y)]^{1/\gamma}.$$

Hence, there is a constant $C' \ge 1$ such that $d(x, y) \le C'[d_\theta(x, y)]^{1/\gamma}$ for all $x, y \in \Delta$.

Define the projection $\pi : \Delta \to M$ by $\pi(y, \ell) = T^\ell y$. Then π is a measure-preserving isomorphism and it follows as above that $d(\pi(x), \pi(y))^\gamma \le C'' d_\theta(x, y)$, for all $x, y \in \Delta$. In particular, since $\phi : (M, d) \to \mathbb{R}$ is γ-Hölder, it follows that $\phi \circ \pi : (\Delta, d_\theta) \to \mathbb{R}$ is Lipschitz. By Theorem 2.8, the ASIP holds for $\phi \circ \pi$ on Δ. Since π is a measure-preserving map semiconjugacy, the ASIP holds for ϕ on M. ∎

Remark 2.10 As already pointed out in [21], the CLT for nonuniformly expanding maps holds under slightly weaker hypotheses using [29, Theorem 1.1]. Instead of requiring that $R \in L^{2+\delta}$, it suffices that $R \in L^2$.

Remark 2.11 The ASIP is said to be *degenerate* if $\sigma^2 = 0$. It follows from previous work in connection with the CLT [40, 41] that the ASIPs obtained in this paper are degenerate if and only if $\phi = \psi \circ T - \psi$ where $\psi \in L^2(M)$. Moreover, by a Livšic regularity result of Bruin *et al.* [9], such an L^2 function ψ has a version that is Hölder on $\cup_{j=0}^{\ell} T^j Y$ for each fixed ℓ. (It is easy to construct examples where the ASIP is degenerate but ψ does not have a version that is continuous on the whole of M.) In particular, if T has a periodic point $x \in Y$ of period k and $\sum_{j=0}^{k-1} \phi(T^j y) \neq 0$, then the ASIP is nondegenerate.

Nonuniformly expanding semiflows We continue to assume that $T : M \to M$ is a nonuniformly expanding map satisfying conditions (1)–(5). Suppose that $h : M \to \mathbb{R}^+$ lies in $L^1(M)$. Regarding h as a *roof function*, we form the *suspension* $M^h = \{(x,u) \in M \times [0,\infty) : 0 \le u \le h(x)\}/\sim$ where $(x,h(x)) \sim (Tx,0)$. The suspension semiflow $T_t : M^h \to M^h$ is given by $T_t(x,u) = x(u+t)$ computed modulo identifications. We call $T_t : M^h \to M^h$ a *nonuniformly expanding semiflow.* We say that an observation $\psi : M^h \to \mathbb{R}$ is Hölder if ψ is bounded and $\sup_{(x,u)\neq(y,u)} |\psi(x,u) - \psi(y,u)|/d(x,y) < \infty$.

Corollary 2.12 *Let $T_t : M^h \to M^h$ be a nonuniformly expanding semiflow. Assume moreover that the return time function R lies in $L^{2+\delta}(Y)$ and that the roof function $h : M \to \mathbb{R}^+$ is Hölder. Let $\psi : M^h \to \mathbb{R}$ be a mean zero Hölder observation. Then ψ satisfies the ASIP. That is, there exists $\epsilon > 0$, a family of random variables $\{S_t\}$ and a Brownian motion W with variance $\sigma^2 \ge 0$ such that $\{\int_0^t \psi \circ T_s \, ds\} =_d \{S_t\}$, and $S_t = W(t) + O(t^{\frac{1}{2}-\epsilon})$ as $t \to \infty$, almost everywhere.*

Proof According to [29, Theorem 4.2] (Theorem B.1), it suffices that (i) $h \in L^{2+\delta}(Y)$, (ii) $\phi(x) = \int_0^{h(x)} \psi(x,u)du$ satisfies the ASIP on Y, and (iii) h satisfies the ASIP on Y. Hence, the result is immediate from Theorem 2.9. ∎

Remark 2.13 We have not striven for greatest generality in the statements of Theorem 2.9 and Corollary 2.12. However, it is clear from the proof that in Theorem 2.9 we can relax the assumption that ϕ is Hölder. It is sufficient that ϕ is such that $\Phi(x) = \sum_{\ell=0}^{R(x)-1} \phi(T^\ell x)$ lies in the space of weighted Lipschitz observations in Subsection (b) for an appropriate choice of weight function. Taking the weight function to be the return time function, it suffices that ϕ is Hölder on $T^\ell Y_j$ for all $j \ge 1$, $0 \le \ell < R(j) - 1$, with L^∞ norm and Hölder constant independent of j, ℓ.

Similarly, the hypotheses that ψ and h are Hölder can be weakened in Corollary 2.12. For example, provided ψ is Hölder, it suffices that h is Hölder on $T^\ell Y_j$ for all $j \ge 1$, $0 \le \ell < R(j) - 1$, with L^∞ norm and Hölder constant independent of j, ℓ.

3 Nonuniformly hyperbolic systems

In this section, we show how to prove the ASIP for Lipschitz observations of a dynamical system that is *nonuniformly hyperbolic in the sense of Young [40]*. Instead of using the original set up, we make four assumptions (A1)–(A4) that are distilled from those in [40]. In doing so, we bypass the differential structure, and certain conclusions in [40] become assumptions here, particularly (A4) below.

Let $T : M \to M$ be a diffeomorphism (possibly with singularities) defined on a Riemannian manifold (M, d). We assume from the start that T preserves a "nice" probability measure m (one of the conclusions in Young [40] is that m is a SRB measure). Assumption (A4) contains the properties of m that we require for the ASIP.

We fix a subset $\Lambda \subset M$ and a family of subsets of M that we call "stable disks" $\{W^s\}$ that are disjoint and cover Λ. If x lies in a stable disk, we label the disk $W^s(x)$.

(A1) There is a partition $\{\Lambda_j\}$ of Λ and integers $R_j \geq 1$ such that for all $x \in \Lambda_j$ we have $T^{R_j}(W^s(x)) \subset W^s(T^{R_j}x)$.

Define the return time function $R : \Lambda \to \mathbb{Z}^+$ by $R|_{\Lambda_j} = R_j$ and the induced map $f : \Lambda \to \Lambda$ by $f(x) = T^{R(x)}(x)$. Form the discrete suspension map $F : \Delta \to \Delta$ where $F(x, \ell) = (x, \ell + 1)$ for $\ell < R(x) - 1$ and $F(x, R(x) - 1) = (fx, 0)$. We define a separation time $s : \Lambda \times \Lambda \to \mathbb{N}$ by defining $s(x, x')$ to be the greatest integer $n \geq 0$ such that $f^k x, f^k x'$ lie in the same partition element of Λ for $k = 0, \ldots, n$. (If x, x' do not lie in the same partition element, then we take $s(x, x') = 0$.) For general points $p = (x, \ell), p' = (x', \ell') \in \Delta$, define $s(p, q) = s(x, x')$ if $\ell = \ell'$ and $s(p, q) = 0$ otherwise. This defines a separation time $s : \Delta \times \Delta \to \mathbb{N}$. We have the projection $\pi : \Delta \to M$ given by $\pi(x, \ell) = T^\ell x$ and satisfying $\pi T = F\pi$.

(A2) There is a distinguished subset or "unstable leaf" $W^u \subset \Lambda$ such that each stable disk intersects W^u in precisely one point, and there exist constants $C \geq 1$, $\alpha \in (0, 1)$ such that

(i) $d(T^n x, T^n y) \leq C\alpha^n$, for all $y \in W^s(x)$, all $n \geq 0$, and

(ii) $d(T^n x, T^n y) \leq C\alpha^{s(x,y)}$ for all $x, y \in W^u$ and all $0 \leq n < R$.

Remark 3.1 We note that Young [40] uses a separation time s_0 defined in terms of the underlying diffeomorphism $T : M \to M$ whereas our separation time s is defined in terms of the induced map $f : \Lambda \to \Lambda$. In particular, [40, conditions (iii) and (iv), p. 589] guarantee that $s_0 \geq s$ and moreover that $s_0 - (R - 1) \geq s$. Hence [40, assumption (P4)(a)] ($d(T^n x, T^n y) \leq C\alpha^{s_0(x,y)-n}$ for $0 \leq n < s_0(x, y)$) implies our assumption (A2)(ii).

There is also a separation time in [40] that is denoted s. This is different from our separation time and plays no role in this paper.

Let $\overline{\Lambda} = \Lambda/\sim$ where $x \sim x'$ if $x \in W^s(x')$. Similarly, define the partition $\{\overline{\Lambda}_j\}$ of $\overline{\Lambda}$. We obtain a well-defined return time function $R : \overline{\Lambda} \to \mathbb{Z}^+$ and induced map $f : \overline{\Lambda} \to \overline{\Lambda}$. Let $F : \overline{\Delta} \to \overline{\Delta}$ denote the corresponding suspension map. We note that this can be viewed as the quotient of $F : \Delta \to \Delta$ where (x, ℓ) is identified with (x', ℓ') if $\ell = \ell'$ and $x' \in W^s(x)$. Let $\overline{\pi} : \Delta \to \overline{\Delta}$ denote the natural projection.

The separation time on Δ drops down to a separation time on $\overline{\Delta}$ (and agrees with the natural separation time defined using $f : \overline{\Lambda} \to \overline{\Lambda}$ and the partition $\{\overline{\Lambda}_j\}$).

(A3) The map $f : \overline{\Lambda} \to \overline{\Lambda}$ and partition $\{\overline{\Lambda}_j\}$ separate points in $\overline{\Lambda}$.

It follows that $d_\theta(p, q) = \theta^{s(p,q)}$ defines a metric on $\overline{\Delta}$ for each $\theta \in (0, 1)$.

(A4) There exist F-invariant probability measures $\widetilde{m}$ on Δ and $\overline{m}$ on $\overline{\Delta}$ such that

- (i) $\pi : \Delta \to M$ and $\overline{\pi} : \Delta \to \overline{\Delta}$ are measure-preserving (π takes $\widetilde{m}$ to m and $\overline{\pi}$ takes $\widetilde{m}$ to $\overline{m}$); and
- (ii) $F : \overline{\Delta} \to \overline{\Delta}$ is a Young tower (in the sense of section 2(d)).

We say that an observation $\psi : \Delta \to \mathbb{R}$ *depends only on future coordinates* if $\psi(p) = \psi(q)$ whenever $p \sim q$ where $\sim$ is the equivalence relation on Δ arising from quotienting along stable disks. Such an observation drops down to an observation $\psi : \overline{\Delta} \to \mathbb{R}$. The following result shows that any Hölder observation on M is related to a Lipschitz observation on $\overline{\Delta}$ (cf. [37, 8]).

Lemma 3.2 *Suppose that $\phi : M \to \mathbb{R}$ is γ-Hölder with respect to the metric d. Then there exist functions $\psi, \chi : \Delta \to \mathbb{R}$ such that*

(i) $\phi \circ \pi = \psi + \chi - \chi \circ F$,

(ii) χ is bounded,

(iii) ψ depends only on future coordinates,

(iv) $\psi : \overline{\Delta} \to \mathbb{R}$ is Lipschitz with respect to the metric d_θ, for $\theta = \alpha^{\gamma/2}$.

Proof Given $p = (x, \ell) \in \Delta$, define $\widehat{p} = (\widehat{x}, \ell)$ where $\widehat{x}$ is the unique point in $W^s(x) \cap W^u$ (see (A2)). Define

$$\chi(p) = \sum_{j=0}^{\infty} \phi(\pi F^j p) - \phi(\pi F^j \widehat{p}).$$

Note that $\pi F^j p = T^j \pi p = T^{j+\ell} x$ and similarly $\pi F^j \widehat{p} = T^{j+\ell}\widehat{x}$. Since x and $\widehat{x}$ lie in the same stable disk W^s, it follows from (A2)(i) that

$$\begin{aligned}|\chi(p)| &\le \sum_{j=0}^{\infty} |\phi(\pi F^j p) - \phi(\pi F^j \widehat{p}) \le |\phi|_\gamma \sum_{j=0}^{\infty} d(T^{j+\ell}x, T^{j+\ell}\widehat{x})^\gamma \\ &\le |\phi|_\gamma C^\gamma \sum_{j=0}^{\infty} \alpha^{j\gamma} = |\phi|_\gamma C^\gamma (1-\alpha^\gamma)^{-1}.\end{aligned}$$

Define $\psi = \phi \circ \pi - \chi + \chi \circ F$. Then $\psi(p) = \sum_{j=0}^{\infty} \phi(\pi F^j \widehat{p}) - \phi(\pi F^j \widehat{Fp})$ depends only upon future coordinates. It remains to check that ψ is Lipschitz with respect to the metric d_θ. In fact, we prove that ψ is Lipschitz with respect to $d_{\theta^{1/2}}$ where $\theta = \alpha^\gamma$.

For any $N \ge 1$, $p, q \in \Delta$,

$$\begin{aligned}|\psi(p) - \psi(q)| \le \sum_{j=0}^{N} |\phi(\pi F^j \widehat{p}) - \phi(\pi F^j \widehat{q})| + \sum_{j=0}^{N-1} |\phi(\pi F^j \widehat{Fp}) - \phi(\pi F^j \widehat{Fq})| \\ + \sum_{j=N+1}^{\infty} |\phi(\pi F^j \widehat{p}) - \phi(\pi F^{j-1} \widehat{Fp})| + \sum_{j=N+1}^{\infty} |\phi(\pi F^j \widehat{q}) - \phi(\pi F^{j-1} \widehat{Fq})|.\end{aligned} \tag{3.1}$$

Suppose that $d_\theta(p,q) = d_\theta(\widehat{p},\widehat{q}) \approx \theta^{2N}$. We show that each of these four terms is bounded by $\theta^N \approx d_{\theta^{1/2}}(p,q)$ up to a constant.

Starting with the third term in (3.1), we note that $F\widehat{p} = \widehat{Fp}$ unless $p = (x, R(x)-1)$, in which case $F\widehat{p} = (f\widehat{x}, 0)$ and $\widehat{Fp} = (\widehat{fx}, 0)$. Then $\pi F^j \widehat{p} = T^{j-1}(f\widehat{x})$ and $\pi F^{j-1}\widehat{Fp} = T^{j-1}(\widehat{fx})$. Since $f\widehat{x}$ and $\widehat{fx}$ lie in the same stable disk W^s, we have $|\phi(\pi F^j \widehat{p}) - \phi(\pi F^{j-1}\widehat{Fp})| \le |\phi|_\gamma C^\gamma \alpha^{(j-1)\gamma}$ so that $\sum_{j=N+1}^{\infty} |\phi(\pi F^j \widehat{p}) - \phi(\pi F^{j-1}\widehat{Fp})| \le C'\theta^N$ as required. Similarly for the fourth term in (3.1).

Next, we consider the first term in (3.1). By assumption, $s(p,q) \approx 2N$ so separation does not takes place during the calculation. Write $p = (x,\ell)$, $q = (y,\ell)$. Then $\pi F^j \widehat{p} = T^{j+\ell}\widehat{x} = T^L f^J \widehat{x}$ where $J \le j$ and $L < R(f^J\widehat{x})$. Similarly, $\pi F^j \widehat{q} = T^L f^J \widehat{y}$. Hence by (A2)(ii),

$$\begin{aligned}|\phi(\pi F^j \widehat{p}) - \phi(\pi F^j \widehat{q})| &\le |\phi|_\gamma d\big(T^L f^J \widehat{x}, T^L f^J \widehat{y}\big)^\gamma \le |\phi|_\gamma C^\gamma \alpha^{s(f^J\widehat{x}, f^J\widehat{y})\gamma} \\ &= |\phi|_\gamma C^\gamma \alpha^{[s(\widehat{x},\widehat{y})-J]\gamma} \le |\phi|_\gamma C^\gamma \alpha^{[s(\widehat{x},\widehat{y})-j]\gamma} \approx |\phi|_\gamma C^\gamma \theta^{2N-j},\end{aligned}$$

so that $\sum_{j=0}^{N} |\phi(\pi F^j \widehat{p}) - \phi(\pi F^j \widehat{q})| \le C'\theta^N$ as required. Similarly for the second term in (3.1). ∎

Remark 3.3 Although Lemma 3.2 is modelled on the treatments in [8, 31], we have not defined a metric on Δ and hence the usual regularity statement about χ is missing.

Theorem 3.4 *Suppose that $T : M \to M$ satisfies (A1)–(A4) and assume that $R \in L^{2+\delta}(\Lambda)$ for some $\delta > 0$. Let $\phi : M \to \mathbb{R}$ be a mean zero Hölder observation. Then ϕ satisfies the ASIP.*

Proof Since $\pi : \Delta \to M$ is measure preserving, it suffices to prove the ASIP for the lift $\widetilde{\phi} = \phi \circ \pi : \Delta \to \mathbb{R}$. By Lemma 3.2, there exists $\psi : \Delta \to \mathbb{R}$ depending only on future coordinates such that $\widetilde{\phi}_N - \psi_N$ is uniformly bounded, and it suffices to prove the ASIP for ψ. Since the projection $\overline{\pi} : \Delta \to \overline{\Delta}$ is measure preserving, it suffices to prove the ASIP for ψ at the level of $\overline{\Delta}$. Finally, Lemma 3.2 guarantees that $\psi : \overline{\Delta} \to \mathbb{R}$ is Lipschitz with respect to d_θ, so it suffices to prove the ASIP for Lipschitz observations on $\overline{\Delta}$ which is a Young tower by (A4)(ii). Now apply Theorem 2.8. ∎

Nonuniformly hyperbolic flows Given an L^1 roof function $h : M \to \mathbb{R}^+$, we define a suspension flow $T_t : M^h \to M^h$ in the same way that we defined the semiflow in Section 2(e). If $T : M \to M$ satisfies (A1)–(A4), we say that $T_t : M^h \to M^h$ is a *nonuniformly hyperbolic flow.*

Corollary 3.5 *Let $T_t : M^h \to M^h$ be a nonuniformly hyperbolic flow. Assume moreover that the return time function R lies in $L^{2+\delta}(Y)$ and that the roof function $h : M \to \mathbb{R}^+$ is Hölder. Let $\psi : M^h \to \mathbb{R}$ be a mean zero Hölder observation. Then ψ satisfies the ASIP.*

Proof This follows immediately from Theorem 3.4, applying Theorem B.1. ∎

Remark 3.6 The weakened hypotheses mentioned in Remark 2.13 apply equally in the nonuniformly hyperbolic setting.

4 Applications

In this section, we indicate a wide range of applications to which the results in this paper apply.

We begin with nonuniformly expanding systems that can be modelled by a Young tower as in Section 2. In the literature it is standard to speak of *return time asymptotics* in the form $m\{y \in Y : R(y) \geq n\} = O(n^{-\gamma})$. (Recall from Section 2 that Y is the subset used for inducing, equivalently the base of the Young tower.)

Proposition 4.1 *If $m\{R \geq n\} = O(n^{-\gamma})$ for some $\gamma > 2$, then $R \in L^{2+\delta}(Y)$ for $\delta \in (0, \gamma - 2)$.*

Proof This is immediate from the inequality $E[R^{2+\delta}] \leq \sum_{n=0}^{\infty} m\{R^{2+\delta} \geq n\} = \sum_{n=0}^{\infty} m\{R \geq n^{\frac{1}{2+\delta}}\}$. ∎

Many maps satisfy the condition in Proposition 4.1:

(i) the Alves-Viana map [3] $T : S^1 \times I \to S^1 \times I$

$$T(\omega, x) = (16\omega, a - x^2 + \epsilon \sin(2\pi\omega))$$

when 0 is preperiodic for the map $x \mapsto a - x^2$ and ϵ is small enough.

(ii) the Liverani-Saussol-Vaienti (Pomeau-Manneville) maps [27] $T : [0, 1] \to [0, 1]$

$$Tx = \begin{cases} x(1 + 2^\alpha x^\alpha) & 0 \leq x < \frac{1}{2} \\ 2x - 1 & \frac{1}{2} \leq x < 1 \end{cases}$$

for $0 < \alpha < \frac{1}{2}$.

(iii) certain classes of multimodal maps, Bruin *et al.* [10].

(iv) a class of expanding circle maps $T : S^1 \to S^1$ of degree $d > 1$ with a neutral fixed point, Young [41, Section 6]: T is C^1 on S^1 and C^2 on $S^1 - \{0\}$, $T' > 1$ on $S^1 - \{0\}$, $T(0) = 0$, $T'(0) = 1$, and for $x \neq 0$, $-xT''(x) \simeq |x|^\alpha$ for $0 < \alpha < \frac{1}{2}$.

Applying Theorem 1.2, we obtain the ASIP for Hölder observations for the systems in (i)–(iv) above. For example, in (iii) and (iv) we obtain the ASIP under the same conditions for which [10] and [41] obtain the CLT.

Next, we recall examples of nonuniformly hyperbolic systems that have been modelled by towers. Consider the following classes of $C^{1+\epsilon}$ diffeomorphisms treated in Young [40] (see also Baladi [4, §4.3]):

(v) Lozi maps and certain piecewise hyperbolic maps [40, 13].

(vi) a class of Hénon maps [6, 7].

(vii) some partially hyperbolic diffeomorphisms with a mostly contracting direction [12, 18].

In these examples, the return time asymptotics are exponential so certainly $R \in L^{2+\delta}$. By Theorem 3.4, we obtain the ASIP for Hölder observations for the systems in (v)–(viii) above.

Billiard maps and Lorentz flows Finally, we consider the application to the planar periodic Lorentz gas discussed in the introduction. Under the finite horizon condition, Young [40] demonstrated that the billiard map (which is the Poincaré map for the flow) is nonuniformly hyperbolic with exponential return time asymptotics. As a result, Young established exponential decay of correlations for such billiard maps, resolving a long-standing (and controversial) open question. Chernov [14] extended Young's method to obtain the same result for infinite horizons.

For our purposes, the weaker conclusion that $R \in L^{2+\delta}$ is again sufficient. Hence, by the results in [14, 40], the first statement of Theorem 1.3 is an immediate consequence of Theorem 3.4.

For the flow itself, the finite horizon condition is crucial since even the CLT is unlikely in the infinite horizon case. Assuming finite horizons, the roof function h is uniformly bounded and piecewise Hölder. Since h is not uniformly Hölder, Corollary 3.5 does not apply directly, but the result is easily modified as in Remarks 2.13 and 3.6 to include such roof functions. Hence, we obtain the second statement of Theorem 1.3.

A ASIP for functions of mixing sequences

Here is a special case of Philipp & Stout [32, Theorem 7.1] adapted to dynamical systems terminology. The notation is as in Section 2(c).

Theorem A.1 (Philipp & Stout) *Assume that there exists $\delta \in (0, 2]$, $\sigma^2 > 0$ and $C > 0$ such that for all $k, N \geq 1$,*

(i) $v \in L^{2+\delta}(\Lambda)$ *and* $\int_\Lambda v \, dm = 0$,

(ii) $\int_\Lambda v_N^2 \, dm = \sigma^2 N + O(N^{1-\delta/30})$,

(iii) $\sum_{a \in \alpha_0^{k-1}} \int_a |v - \frac{1}{m(a)} \int_a v \, dm|^{2+\delta} dm \leq Ck^{-(2+7/\delta)(2+\delta)}$,

(iv) $\left| m(a \cap f^{-(N+k)}(b)) - m(a)m(b) \right| \leq CN^{-168(1+2/\delta)}$ *for all* $a \in \alpha_0^{k-1}$ *and all measurable sets* b.

Then $v_N = W(N) + O(N^{1/2-\delta/600})$.

B ASIP for suspensions

Suppose that (Λ, m) is a probability space and that $f : \Lambda \to \Lambda$ is a measure-preserving transformation. Let $h : \Lambda \to \mathbb{R}^+$ be a roof function and suppose that $f_t : \Lambda^h \to \Lambda^h$ is the corresponding suspension (semi)flow as in Section 2(e). The following result is a special case of [30, Theorem 4.2].

Theorem B.1 (Melbourne & Török) *Let $\delta > 0$. Suppose that $h \in L^{2+\delta}(\Lambda)$ and that $\sum_{j=0}^{N-1} h \circ f^j = N\overline{h} + o(N^{1-\delta})$ as $N \to \infty$ almost everywhere.*

Suppose that $\psi : \Lambda^h \to \mathbb{R}$ lies in $L^\infty(\Lambda^h)$ and has mean zero. Define $\phi : \Lambda \to \mathbb{R}$ by $\phi(x) = \int_0^{h(x)} \psi(f_t x)$. If ϕ satisfies the ASIP on Λ with variance σ_1^2, then ψ satisfies the ASIP on Λ^h with variance $\sigma^2 = \sigma_1^2/\overline{h}$.

Theorem B.1 is easily modified for discrete suspensions. Let $R : \Lambda \to \mathbb{Z}^+$ be an L^1 return time function and form the discrete suspension map $F : \Delta \to \Delta$ as in Section 2(d).

Corollary B.2 *Let $\delta > 0$. Suppose that $R \in L^{2+\delta}(\Lambda)$ and that $\sum_{j=0}^{N-1} R \circ f^j = N\overline{R} + o(N^{1-\delta})$ as $N \to \infty$ almost everywhere.*

Suppose that $\phi : \Delta \to \mathbb{R}$ lies in $L^\infty(\Delta)$ and has mean zero. Define $\Phi : \Lambda \to \mathbb{R}$ by $\Phi(x) = \sum_{j=0}^{R(x)-1} \phi(f^j x)$. If Φ satisfies the ASIP on Λ with variance σ_1^2, then ϕ satisfies the ASIP on Δ with variance $\sigma^2 = \sigma_1^2/\overline{R}$.

Acknowledgements This research was supported in part by EPSRC Grant GR/S11862/01. IM is greatly indebted to UH for the use of e-mail, given that pine is currently not supported on the University of Surrey network.

References

[1] J. Aaronson. *An Introduction to Infinite Ergodic Theory*. Math. Surveys and Monographs **50**, Amer. Math. Soc., 1997.

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

[3] J. Alves, S. Luzzatto and V. Pinheiro. Markov structures and decay of correlations for non-uniformly expanding dynamical systems. *Ann. Inst. H. Poincaré, Anal. Non Linéaire*. To appear.

[4] V. Baladi. *Positive Transfer Operators and Decay of Correlations*. Advanced Series in Nonlinear Dynamics **16**, World Scientific, Singapore, 2000.

[5] V. Baladi. Decay of correlations. *Smooth Ergodicity Theory and its Applications* (A. Katok et al., ed.), Proc. Symp. Pure Math. **69**, Amer. Math. Soc., 2001, pp. 297–325.

[6] M. Benedicks and L.-S. Young. Absolutely continuous invariant measures and random perturbations for certain one-dimensional maps. *Ergod. Th. & Dynam. Sys.* **12** (1992) 13–37.

[7] M. Benedicks and L.-S. Young. Sinai-Bowen-Ruelle measures for certain Hénon maps. *Invent. Math.* **112** (1993) 541–576.

[8] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*. Lecture Notes in Math. **470**, Springer, Berlin, 1975.

[9] H. Bruin, M. Holland and M. Nicol. Livsic regularity for Markov systems. Preprint.

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

[11] L. A. Bunimovich, Y. G. Sinaĭ and N. I. Chernov. Statistical properties of two-dimensional hyperbolic billiards. *Uspekhi Mat. Nauk* **46** (1991) 43–92.

[12] A. Castro. Backward inducing and exponential decay of correlations for partially hyperbolic attractors with mostly contracting direction. Ph. D. Thesis, IMPA (1998).

[13] N. Chernov. Statistical properties of piecewise smooth hyperbolic systems in high dimensions. *Discrete Contin. Dynam. Systems* **5** (1999) 425–448.

[14] N. Chernov. Decay of correlations and dispersing billiards. *J. Statist. Phys.* **94** (1999) 513–556.

[15] N. Chernov and L. S. Young. Decay of correlations for Lorentz gases and hard balls. *Hard ball systems and the Lorentz gas*. Encyclopaedia Math. Sci. **101**, Springer, Berlin, 2000, pp. 89–120.

[16] J.-P. Conze and S. Le Borgne. Méthode de martingales et flow géodésique sur une surface de courbure constante négative. *Ergod. Th. & Dynam. Sys.* **21** (2001) 421–441.

[17] M. Denker and W. Philipp. Approximation by Brownian motion for Gibbs measures and flows under a function. *Ergod. Th. & Dynam. Sys.* **4** (1984) 541–552.

[18] D. Dolgopyat. On dynamics of mostly contracting diffeomorphisms. *Commun. Math. Phys.* **213** (2000) 181–201.

[19] M. J. Field, I. Melbourne and A. Török. Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions. *Ergod. Th. & Dynam. Sys.* **23** (2003) 87–110.

[20] M. I. Gordin. The central limit theorem for stationary processes. *Soviet Math. Dokl.* **10** (1969) 1174–1176.

[21] S. Gouëzel. Statistical properties of a skew product with a curve of neutral points. Preprint, 2004.

[22] S. Gouëzel. Vitesse de décorrélation et théorèmes limites pour les applications non uniformément dilatantes. Ph. D. Thesis, Ecole Normale Supérieure, 2004.

[23] H. Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc.* **118** (1993) 627–634.

[24] F. Hofbauer and G. Keller. Ergodic properties of invariant measures for piecewise monotonic transformations. *Math. Z.* **180** (1982) 119–140.

[25] G. Keller. Un théorème de la limite centrale pour une classe de transformations monotones per morceaux. *C. R. Acad. Sci. Paris* **291** (1980) 155–158.

[26] C. Liverani. Central limit theorem for deterministic systems. *International Conference on Dynamical Systems* (F. Ledrappier, J. Lewowicz and S. Newhouse, eds.), Pitman Research Notes in Math. **362**, Longman Group Ltd, Harlow, 1996, pp. 56–75.

[27] C. Liverani, B. Saussol and S. Vaienti. A probabilistic approach to intermittency. *Ergodic Theory and Dynamical Systems* **19** (1999) 671–685.

[28] I. Melbourne and M. Nicol. Statistical properties of endomorphisms and compact group extensions. *J. London Math. Soc.* **70** (2004) 427–446.

[29] I. Melbourne and A. Török. Statistical limit theorems for suspension flows. *Israel J. Math.* **194** (2004) 191–210.

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

[31] W. Parry and M. Pollicott. *Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics.* Astérique **187-188**, Société Mathématique de France, Montrouge, 1990.

[32] W. Philipp and W. F. Stout. *Almost Sure Invariance Principles for Partial Sums of Weakly Dependent Random Variables*. Memoirs of the Amer. Math. Soc. **161**, Amer. Math. Soc., Providence, RI, 1975.

[33] M. Pollicott and R. Sharp. Invariance principles for interval maps with an indifferent fixed point. *Commun. Math. Phys.* **229** (2002) 337–346.

[34] M. Ratner. The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature. *Israel J. Math.* **16** (1973) 181–197.

[35] D. Ruelle. *Thermodynamic Formalism*. Encyclopedia of Math. and its Applications **5**, Addison Wesley, Massachusetts, 1978.

[36] Y. G. Sinaĭ. Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards. *Uspehi Mat. Nauk* **25** (1970) 141–192.

[37] Y. G. Sinaĭ. Gibbs measures in ergodic theory. *Russ. Math. Surv.* **27** (1972) 21–70.

[38] M. Viana. *Stochastic dynamics of deterministic systems*. Col. Bras. de Matemática, 1997.

[39] C. P. Walkden. Invariance principles for iterated maps that contract on average. Preprint, 2003.

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

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