PAPER

Martingale approximations and anisotropic Banach spaces with an application to the time-one map of a Lorentz gas

IOP Publishing | London Mathematical Society Nonlinearity

Nonlinearity **33** (2020) 4095–4113 https://doi.org/10.1088/1361-6544/ab7d22

Martingale approximations and anisotropic Banach spaces with an application to the time-one map of a Lorentz gas

Mark Demers[1], Ian Melbourne[2] and Matthew Nicol[3,4]

[1] Department of Mathematics, Fairfield University, Fairfield, CT 06824, United States of America
[2] Mathematics Institute, University of Warwick, Coventry, CV4 7AL, United Kingdom
[3] Department of Mathematics, University of Houston, Houston TX 77204-3008, United States of America

E-mail: nicol@math.uh.edu

Received 1 January 2019, revised 17 February 2020
Accepted for publication 5 March 2020
Published 2 July 2020

Abstract

In this paper, we show how the Gordin martingale approximation method fits into the anisotropic Banach space framework. In particular, for the time-one map of a finite horizon planar periodic Lorentz gas, we prove that Hölder observables satisfy statistical limit laws such as the central limit theorem and associated invariance principles. Previously, these properties were known only for a restricted class of observables, excluding for instance velocity.

Keywords: smooth ergodic theory, probabilistic limit laws, martingale methods

Mathematics Subject Classification numbers: 37D25, 37D50, 37A50, 60F17, 60G42.

1. Introduction

The traditional approach to proving decay of correlations and statistical limit laws for deterministic dynamical systems, following [7, 44, 45] and continuing with Young [50, 51], involves symbolic coding. In particular, by quotienting along stable leaves one passes from an invertible dynamical system to a one-sided shift. Decay of correlations is then a consequence of the contracting properties of the associated transfer operator. In addition, Nagaev perturbation arguments [24, 25] and the martingale approximation method of Gordin [22] are available in this setting, leading to numerous statistical limit laws. These results on decay of correlations and statistical limit laws are then readily passed back to the original dynamical system.

[4] Author to whom any correspondence should be addressed.

A downside to this approach is that geometric and smooth structures associated to the underlying dynamical system are typically destroyed by symbolic coding. In recent years, a method proposed by [6] and developed extensively by numerous authors (for recent articles with up-to-date references see [2, 19]) uses anisotropic Banach spaces of distributions to study the underlying dynamical system directly. In particular, the method does not involve quotienting along stable manifolds. This leads to results on rates of decay of correlations and also to various statistical limit laws via Nagaev perturbation arguments, see especially Gouëzel [23].

However, so far Gordin's martingale approximation argument has been absent from the anisotropic Banach space framework. This is the topic of the current paper. The utility of such an approach is illustrated by the following example.

Example 1.1. The landmark result of Young [50] established exponential decay of correlations for the collision map corresponding to planar periodic dispersing billiards with finite horizon. The method, which involves symbolic coding, also yields the central limit theorem (CLT) for Hölder observables, recovering results of [8].

Turning to the corresponding flow, known as the finite horizon planar periodic Lorentz gas, the CLT follows straightforwardly from the result for billiards [8, 40]. However, decay of correlations for the Lorentz gas and the CLT for the time-one map of the Lorentz gas are much harder. Superpolynomial decay of correlations was established for sufficiently regular observables in [35] (see also [36]) using symbolic coding and Dolgopyat-type estimates [20]. This method also yields the CLT for the time-one map [1, 39], but again only for sufficiently regular observables. Here, 'regular' means smooth along the flow direction, so this excludes many physically relevant observables such as velocity. The rate of decay of correlations was improved to subexponential decay [11] and finally in a recent major breakthrough to exponential decay [3]. Both references handle Hölder observables, suggesting that statistical limit laws such as the CLT for the time-one map should hold for general Hölder observables.

Currently the Nagaev method is unavailable for Lorentz gases, and as a consequence the CLT for the time-one map was previously unavailable except for a restricted class of observables. We show that the Gordin approach is applicable and hence the CLT and related limit laws are indeed satisfied by Hölder observables for these examples. In particular, observables such as velocity are covered for the first time.

In the remainder of the introduction, we describe some of the limit laws that follow from the methods in this paper. For definiteness, we focus on example 1.1. Let X be the three-dimensional phase space corresponding to a finite horizon planar periodic Lorentz gas, with invariant volume measure μ, and let $T: X \to X$ be the time-one map of the Lorentz flow. Let $\phi: X \to \mathbb{R}$ be a Hölder observable with mean zero and define the Birkhoff sum $\phi_n = \sum_{j=0}^{n-1} \phi \circ T^j$. It follows from [3, 11] that we can define

$$\sigma^2 = \lim_{n\to\infty} n^{-1} \int_X \phi_n^2 \mathrm{d}\mu = \sum_{n=-\infty}^{\infty} \int_X \phi\, \phi \circ T^n \mathrm{d}\mu.$$

By [1, theorem B and remark 1.1], typically $\sigma^2 > 0$ (the case $\sigma^2 = 0$ is of infinite codimension). We obtain the following results[5].

CLT: $n^{-1/2}\phi_n \to_d N(0, \sigma^2)$ as $n \to \infty$. That is

$$\lim_{n\to\infty} \mu(x \in X : n^{-1/2}\phi_n(x) \leqslant c) = (2\pi\sigma^2)^{-1/2} \int_{-\infty}^{c} \mathrm{e}^{-y^2/(2\sigma^2)}\, \mathrm{d}y \quad \text{for all } c \in \mathbb{R}.$$

[5] In what follows, $\to_{\mathrm{d}}$ denotes convergence in distribution while $\to_{\mathrm{w}}$ denotes weak convergence.

Weak invariance principle (WIP): define $W_n(t) = n^{-1/2}\phi_{nt}$ for $t = 0, \frac{1}{n}, \frac{2}{n}, \ldots$ and linearly interpolate to obtain $W_n \in C[0,1]$. Then $W_n \to_{\mathrm{w}} W$ where W denotes Brownian motion with variance σ^2.

Moment estimates: for every $p \geqslant 1$ there exists $C_p > 0$ such that $|\phi_n|_p \leqslant C_p n^{1/2}$. Consequently, $\lim_{n\to\infty} n^{-p/2}\|\phi_n\|_p^p = \mathbb{E}|Y|^p$ where $Y =_d N(0,\sigma^2)$.

Homogenization: now suppose that $\phi : X \to \mathbb{R}^k$. We continue to suppose that ϕ is C^η for some $\eta \in (0,1]$ and that $\int_X \phi \mathrm{d}\mu = 0$. Consider the fast-slow system

$$\begin{aligned} x(n+1) &= x(n) + \epsilon^2 a(x(n)) + \epsilon b(x(n))\phi(y(n)), \\ y(n+1) &= Ty(n), \end{aligned} \tag{1.1}$$

where $x(0) = \xi \in \mathbb{R}^d$ and $y(0)$ is drawn randomly from (X,μ). We suppose that $a : \mathbb{R}^d \to \mathbb{R}^d$ lies in $C^{1+\eta}$ and $b : \mathbb{R}^d \to \mathbb{R}^{d\times k}$ lies in $C^{2+\eta}$. Solve (1.1) to obtain

$$x_\epsilon(n) = \xi + \epsilon^2 \sum_{j=0}^{n-1} a(x_\epsilon(j)) + \epsilon \sum_{j=0}^{n-1} b(x_\epsilon(j))\phi(y(j)), \qquad y(n) = T^n y(0),$$

and let $\hat{x}_\epsilon(t) = x_\epsilon([t/\epsilon^2])$. This defines a random process on the probability space (X,μ) depending on $y(0) \in X$. Then $\hat{x}_\epsilon \to_w Z$ as $\epsilon \to 0$, where Z satisfies an Itô stochastic differential equation $\mathrm{d}Z = \tilde{a}(Z)\mathrm{d}t + b(Z)\mathrm{d}W$, $Z(0) = \xi$, where W is a k-dimensional Brownian motion with covariance matrix Σ and

$$\tilde{a}(x) = a(x) + \sum_{\alpha=1}^{d} \sum_{\beta,\gamma=1}^{k} E^{\gamma\beta} \frac{\partial b^\beta}{\partial x_\alpha}(x) b^{\alpha\gamma}(x). \tag{1.2}$$

Here, b^β is the βth column of b and the matrices $\Sigma, E \in \mathbb{R}^{k\times k}$ are given by

$$\Sigma^{\beta\gamma} = \sum_{n=-\infty}^{\infty} \int_X \phi^\beta\, \phi^\gamma \circ T^n \mathrm{d}\mu, \qquad E^{\beta\gamma} = \sum_{n=1}^{\infty} \int_X \phi^\beta\, \phi^\gamma \circ T^n \mathrm{d}\mu.$$

The remainder of this paper is organized as follows. In section 2, we recall background material on martingale-coboundary decompositions and statistical limit laws. In section 3, we state an abstract theorem on obtaining martingale-coboundary decompositions for invertible systems with stable directions. In section 4, we apply our results to the time-one map of the Lorentz gas.

2. Martingale approximations

In this section, we review the approach going back to Gordin [22]. This method yields martingale approximations for observables of dynamical systems leading to various limit theorems. Related references include [4, 5, 17, 18, 26, 31, 47, 48, 49]. Let (X,μ) be a probability space,

and let $T : X \to X$ be an invertible ergodic measure-preserving transformation. Let $\mathcal{F}_0$ be a sub-σ-algebra of the underlying σ-algebra on X such that $T^{-1}\mathcal{F}_0 \subseteq \mathcal{F}_0$. Consider an observable[6] $\phi \in L^1(X)$ with $\int_X \phi \mathrm{d}\mu = 0$.

Definition 2.1. We say that ϕ admits a martingale-coboundary decomposition if

$$\phi = m + \chi \circ T - \chi,$$

where $m, \chi \in L^1(X)$, m is $\mathcal{F}_0$-measurable, and $\mathbb{E}[m|T^{-1}\mathcal{F}_0] = 0$.

The conditions on m in definition 2.1 mean that $\{m \circ T^{-n} : n \in \mathbb{Z}\}$ is a sequence of martingale differences with respect to the filtration $\{T^n\mathcal{F}_0 : n \in \mathbb{Z}\}$.

Proposition 2.2. *Let $\phi \in L^p(X)$ for some $p \geqslant 1$. Suppose that*

$$\sum_{n\geqslant 1} |\mathbb{E}[\phi \circ T^{-n}|\mathcal{F}_0]|_p < \infty, \qquad \sum_{n\geqslant 0} |\mathbb{E}[\phi \circ T^n|\mathcal{F}_0] - \phi \circ T^n|_p < \infty. \tag{2.1}$$

Then ϕ admits a martingale-coboundary decomposition with $m, \chi \in L^p(X)$.

Proof. This is a standard argument [26, 48]. We give the details for completeness. By (2.1),

$$\chi = \sum_{n\geqslant 0} (\mathbb{E}[\phi \circ T^n|\mathcal{F}_0] - \phi \circ T^n) + \sum_{n\geqslant 1} \mathbb{E}[\phi \circ T^{-n}|\mathcal{F}_0]$$

converges in $L^p(X)$. Define $m = \phi + \chi - \chi \circ T \in L^p(X)$. Then

$$m = \sum_{n=-\infty}^{\infty} (g_n - g_n \circ T) = \sum_{n=-\infty}^{\infty} (g_{n+1} - g_n \circ T), \tag{2.2}$$

where $g_n = \mathbb{E}[\phi \circ T^n|\mathcal{F}_0]$.

Clearly, $g_n = \mathbb{E}[\phi \circ T^n|\mathcal{F}_0]$ is $\mathcal{F}_0$-measurable. Also, $g_n \circ T$ is measurable with respect to $T^{-1}\mathcal{F}_0 \subseteq \mathcal{F}_0$. Hence m is $\mathcal{F}_0$-measurable.

Next, note that $g_n \circ T = \mathbb{E}[\phi \circ T^n|\mathcal{F}_0] \circ T = \mathbb{E}[\phi \circ T^{n+1}|T^{-1}\mathcal{F}_0]$. Hence

$$\mathbb{E}[g_n \circ T|T^{-1}\mathcal{F}_0] = \mathbb{E}[\phi \circ T^{n+1}|T^{-1}\mathcal{F}_0] = \mathbb{E}[\mathbb{E}[\phi \circ T^{n+1}|\mathcal{F}_0]|T^{-1}\mathcal{F}_0] = \mathbb{E}[g_{n+1}|T^{-1}\mathcal{F}_0],$$

where we used that $T^{-1}\mathcal{F}_0 \subseteq \mathcal{F}_0$. Substituting into (2.2), we obtain $\mathbb{E}[m|T^{-1}\mathcal{F}_0] = 0$ as required. □

2.1. Central limit theorem and invariance principles

Corollary 2.3. *Assume that $\phi \in L^2(X)$ and conditions (2.1) hold with $p = 2$. Then the CLT and WIP hold with $\sigma^2 = \int_X m^2 \mathrm{d}\mu = \lim_{n\to\infty} n^{-1}|\phi_n|_2^2$.*

Proof. This is a standard application of martingale limit theorems [22]. □

Somewhat surprisingly, by the results of [18], if $\phi \in L^\infty(X)$ and conditions (2.1) hold for $p = 1$, then automatically $m \in L^2(X)$ even though proposition 2.2 only gives $m, \chi \in L^1(X)$. This suffices for the CLT. Related references for this phenomenon whereby m has extra regularity include [29, 31, 34, 43, 46, 49]. In particular, the following result holds:

[6] Most observables in this paper are real-valued, but occasionally in this section we consider observables with values in $\mathbb{R}^k$. We write $L^1(X, \mathbb{R}^k)$ to denote vector-valued observables and write $L^1(X)$ instead of $L^1(X, \mathbb{R})$.

Theorem 2.4. *Assume that $\phi \in L^\infty(X)$ and conditions (2.1) hold with $p = 1$. Then the CLT and WIP hold.*

Proof. The CLT and WIP in reverse time (as $n \to -\infty$) is an immediate consequence of [18, corollary 4]. Passing from reverse time to forward time is standard (see for example [27, section 4.2]). □

Now let ϕ be vector-valued with values in $\mathbb{R}^k$. Define càdlàg processes W_n in $\mathbb{R}^k$ and $\mathbb{W}_n \in \mathbb{R}^{k\times k}$:

$$W_n(t) = n^{-1/2} \sum_{0\leqslant j<nt} \phi \circ T^j, \qquad \mathbb{W}_n^{\beta\gamma}(t) = n^{-1} \sum_{0\leqslant i<j<nt} \phi^\beta \circ T^i \phi^\gamma \circ T^j.$$

Proposition 2.5 (Iterated WIP). *Suppose that T is mixing. Assume that $\phi \in L^2(X, \mathbb{R}^k)$ and conditions (2.1) hold with $p = 2$. Then*

(a) The series $\Sigma^{\beta\gamma} = \sum_{n=-\infty}^{\infty} \int_X \phi^\beta\, \phi^\gamma \circ T^n \mathrm{d}\mu$, $\quad E^{\beta\gamma} = \sum_{n=1}^{\infty} \int_X \phi^\beta\, \phi^\gamma \circ T^n \mathrm{d}\mu$, converge.

(b) $(W_n, \mathbb{W}_n) \to_w (W, \mathbb{W})$, where W is a k-dimensional Brownian motion with covariance matrix Σ and $\mathbb{W}^{\beta\gamma}(t) = \int_0^t W^\beta \mathrm{d}W^\gamma + E^{\beta\gamma} t$.

Proof. By proposition 2.2, ϕ admits a martingale-coboundary decomposition with m, $\chi \in L^2(X, \mathbb{R}^k)$, so the result holds by [27, theorem 4.3]. □

2.1.1. Moments. For optimal moment estimates, the following projective version of conditions (2.1) are better suited:

$$\sum_{n\geqslant 1} n^{-1/2} |\mathbb{E}[\phi \circ T^{-n} | \mathcal{F}_0]|_p < \infty, \qquad \sum_{n\geqslant 0} n^{-1/2} |\mathbb{E}[\phi \circ T^n | \mathcal{F}_0] - \phi \circ T^n|_p < \infty. \quad (2.3)$$

Proposition 2.6. *Assume $\phi \in L^p(X)$ and conditions (2.3) hold for some $p > 2$. Then $|\max_{k\leqslant n} |\phi_k||_p = O(n^{1/2})$.*

If in addition $n^{-1/2}\phi_n \to_{\mathrm{d}} Y$ for some L^p random variable Y, then $\lim_{n\to\infty} n^{-q/2} |\phi_n|_q^q = \mathbb{E}|Y|^q$ for all $q < p$.

Proof. Let $A_n = \sum_{j=1}^n \phi \circ T^{-j}$. Then[7] for $r \geqslant 1$,

$$\sum_{k=0}^{r-1} 2^{-k/2} |\mathbb{E}(A_{2^k} | \mathcal{F}_0)|_p \leqslant \sum_{k=0}^{r-1} 2^{-k/2} \sum_{j=1}^{2^k} \left|\mathbb{E}(\phi \circ T^{-j} | \mathcal{F}_0)\right|_p$$

$$= \sum_{j=1}^{2^{r-1}} \sum_{k=\lceil \log_2 j \rceil}^{r-1} 2^{-k/2} \left|\mathbb{E}(\phi \circ T^{-j} | \mathcal{F}_0)\right|_p \ll \sum_{j=1}^{2^{r-1}} j^{-1/2} \left|\mathbb{E}(\phi \circ T^{-j} | \mathcal{F}_0)\right|_p.$$

By condition (2.3), $\sum_{k=0}^{\infty} 2^{-k/2} |\mathbb{E}(A_{2^k} | \mathcal{F}_0)|_p < \infty$. Similarly, $\sum_{k=1}^{\infty} 2^{-k/2} |A_{2^k} - |\mathbb{E}(A_{2^k} | T^{2^k} \mathcal{F}_0)|_p < \infty$. Recalling that $T^{-1}\mathcal{F}_0 \subseteq \mathcal{F}_0$, it follows from [17, corollary 3.9] that $|\max_{k\leqslant 2^r} |A_k||_p \ll 2^{r/2}$.

[7] We use the notation $A \ll B$ to denote $A \leqslant$ const. B, where the constant is independent of the other parameters present.

For general $n \geqslant 1$ choose $r \geqslant 1$ so that $2^{r-1} < n \leqslant 2^r$. Then

$$\left|\max_{k\leqslant n}|A_k|\right|_p \leqslant \left|\max_{k\leqslant 2^r}|A_k|\right|_p \ll 2^{r/2} \leqslant (2n)^{1/2}.$$

Finally, $\phi_k = (A_n - A_{n-k}) \circ T^n$ so

$$\left|\max_{k\leqslant n}|\phi_k|\right|_p = \left|\max_{k\leqslant n}|A_n - A_{n-k}|\right|_p \leqslant 2\left|\max_{k\leqslant n}|A_k|\right|_p \ll n^{1/2},$$

proving the first statement.

The second statement is an immediate consequence of the first, see for example [41, lemma 2.1(e)]. □

Now let ϕ be vector-valued with values in $\mathbb{R}^k$ and define $S_n^{\beta\gamma} = \sum_{0\leqslant i<j<n}\phi^\beta \circ T^i\, \phi^\gamma \circ T^j$.

Proposition 2.7. *Assume that $\phi \in L^p(X, \mathbb{R}^k)$ and conditions (2.1) hold for some $p \geqslant 4$. Then $\left|\max_{k\leqslant n}|S_k^{\beta\gamma}|\right|_{p/2} = O(n)$.*

Proof. By proposition 2.2, we have a martingale-coboundary decomposition $\phi = m + \chi \circ T - \chi$ with m, $\chi \in L^p(X, \mathbb{R}^k)$. Write

$$S_n^{\beta\gamma} = \sum_{0\leqslant i<j<n} m^\beta \circ T^i\, \phi^\gamma \circ T^j + \sum_{1\leqslant j<n}(\chi^\beta \circ T^j - \chi^\beta)\phi^\gamma \circ T^j = I_n + J_n$$

where $I_n = \sum_{0\leqslant i<j<n} m^\beta \circ T^i m^\gamma \circ T^j$ and

$$J_n = \sum_{0\leqslant i<n-1} m^\beta \circ T^i\,(\chi^\gamma \circ T^n - \chi^\gamma \circ T^{i+1}) + \sum_{1\leqslant j<n}(\chi^\beta \circ T^j - \chi^\beta)\phi^\gamma \circ T^j.$$

Now,

$$\max_{k\leqslant n}|J_k| \leqslant \sum_{0\leqslant i<n-1}|m^\beta| \circ T^i\,(|\chi^\gamma| \circ T^n + |\chi^\gamma| \circ T^{i+1}) + \sum_{1\leqslant j<n}(|\chi^\beta| \circ T^j + |\chi^\beta|)|\phi^\gamma| \circ T^j.$$

Hence $\left|\max_{k\leqslant n}|J_k|\right|_{p/2} \leqslant 2n\left(|m^\beta|_p|\chi^\gamma|_p + |\chi^\beta|_p|\phi^\gamma|_p\right)$.

Next, we recall the identity

$$I_k = I_n - I_{n-k} \circ T^k - (m_n^\beta - m_{n-k}^\beta \circ T^k)(m_{n-k}^\gamma \circ T^k), \quad 0 \leqslant k \leqslant n,$$

where $m_n^\beta = \sum_{i=0}^{n-1} m^\beta \circ T^i$. Set

$$m_n^{\beta,-} = \sum_{1\leqslant i\leqslant n} m^\beta \circ T^{-i}, \qquad I_n^- = \sum_{1\leqslant j<i\leqslant n} m^\beta \circ T^{-i} m^\gamma \circ T^{-j}.$$

Then $m_{n-k}^\beta \circ T^k = m_{n-k}^{\beta,-} \circ T^n$ and $I_{n-k} \circ T^k = I_{n-k}^- \circ T^n$ for all $k \leqslant n$. Hence

$$I_k = \left(I_n^- - I_{n-k}^- - (m_n^{\beta,-} - m_{n-k}^{\beta,-})m_{n-k}^{\gamma,-}\right) \circ T^n$$

and so

$$|\max_{k\leqslant n}|I_k||_{p/2} \leqslant 2|\max_{k\leqslant n}|I_k^-||_{p/2} + 2|\max_{k\leqslant n}|m_k^{\beta,-}||_p|\max_{k\leqslant n}|m_k^{\gamma,-}||_p. \tag{2.4}$$

Now

$$I_k^- = \sum_{i=2}^{k} X_i \quad \text{where } X_i = m^\beta \circ T^{-i}\left(\sum_{j=1}^{i-1} m^\gamma \circ T^{-j}\right) = m^\beta \circ T^{-i} m_{i-1}^{\gamma,-}.$$

Since $\{m \circ T^{-n}; n \geqslant 0\}$ is a sequence of L^p martingale differences, $\{X_i; i \geqslant 1\}$ is a sequence of $L^{p/2}$ martingale differences. By the inequalities of Doob and Burkholder [9],

$$|\max_{k\leqslant n}|I_k^-||_{p/2}^2 \ll \left|\left(\sum_{i=1}^{n} X_i^2\right)^{1/2}\right|_{p/2}^2 = \left|\sum_{i=1}^{n} X_i^2\right|_{p/4}.$$

(The implied constant depends only on p and is in particular independent of n.) Hence, using that $p \geqslant 4$,

$$|\max_{k\leqslant n}|I_k^-||_{p/2}^2 \ll \sum_{i=1}^{n} |X_i^2|_{p/4} = \sum_{i=1}^{n} |X_i|_{p/2}^2 \leqslant |m^\beta|_p^2 \sum_{i=1}^{n} |m_{i-1}^{\gamma,-}|_p^2.$$

Applying Burkholder once more, $\left|\max_{k\leqslant n}|m_k^{\gamma,-}|\right|_p \ll n^{1/2}|m^\gamma|_p$; in particular $|\max_{k\leqslant n}|I_k^-||_{p/2} \ll n|m^\beta|_p|m^\gamma|_p$. Substituting these estimates into (2.4) yields $|\max_{k\leqslant n}|I_k||_{p/2} \ll n|m^\beta|_p|m^\gamma|_p$ and the result follows. □

Remark 2.8. There is an error in [27, proposition 7.1] due to an inaccurate application of a (correct) result of [38]. (The argument in [27] is fine for nonuniformly expanding maps but false for nonuniformly hyperbolic maps since the observable ϕ is not adapted to the filtration for the martingale.)

This error was repeated in the first version of the current paper and was spotted by the referee. As pointed out to us by the referee, the reference [17] can be used for the ordinary moments ϕ_n and this argument is now employed in the proof of proposition 2.6. (Indeed, proposition 2.6 is an improvement on the previous result [38, equation (3.1)] since it is no longer required that $\phi \in L^\infty(X)$.) However, it remains an interesting open problem to obtain optimal control of the iterated moments S_n.

2.1.2. Homogenization. As shown in [27, 28], rough path theory yields homogenization of fast-slow systems (1.1) provided the iterated WIP and suitable iterated moment estimates hold. The iterated moment estimates have been relaxed in [15, 16]. We now apply these results to the fast-slow system (1.1).

Define the càdlàg process $\hat{x}_\epsilon$ and the stochastic process Z as in the introduction. We continue to assume that $a \in C^{1+\eta}$ and $b \in C^{2+\eta}$ for some $\eta > 0$.

Theorem 2.9. *Suppose that T is mixing. Assume that $\phi \in L^p(X, \mathbb{R}^k)$ and conditions (2.1) hold with $p = 4$. Then $\hat{x}_\epsilon \to_w Z$ as $\epsilon \to 0$.*

Proof. The iterated WIP holds by proposition 2.5. By [15, theorem 4.10], it now suffices to show that $|\max_{k\leqslant n}|\phi_k||_{2q} = O(n^{1/2})$ and $|\max_{k\leqslant n}|S_k||_q = O(n)$ for some $q > 1$. This and more follows from propositions 2.6 and 2.7. □

Remark 2.10. The standard WIP and moments are insufficient to determine the limiting stochastic process Z. By rough path theory [21, 32] the iterated process $\mathbb{W}_n$ provides the extra information required to determine limiting stochastic integrals, and thereby the modified drift term (1.2). The iterated moment estimate $S_n^{\beta\gamma}$ provides the required tightness.

Note that $\mathbb{W}_n$ and $S_n^{\beta\gamma}$ involve summation over $i<j$. The behaviour of their symmetrized versions (incorporating $i>j$ terms, equivalently $i\geqslant j$ terms) follows immediately from the ordinary WIP and moment estimate, and hence provides no extra information. (Indeed the symmetrized version of $\mathbb{W}_n^{\beta\gamma}$ is $W_n^\beta W_n^\gamma$ which converges weakly to $W^\beta W^\gamma$).

3. Main abstract theorem

Let $T:X\to X$ be an invertible ergodic measure-preserving transformation on a probability space (X,μ). We suppose that X is covered by a collection $\mathcal{W}^{\mathrm{s}}$ of disjoint measurable subsets, called 'local stable leaves', such that $TW^{\mathrm{s}}(x)\subseteq W^{\mathrm{s}}(Tx)$ for all $x\in X$, where $W^{\mathrm{s}}(x)$ is the partition element containing x.

Let $\mathcal{F}_0$ denote the σ-algebra generated by $\mathcal{W}^{\mathrm{s}}$. Note that $W^{\mathrm{s}}(y)\subseteq T^{-1}W^{\mathrm{s}}(x)$ for all $y\in T^{-1}W^{\mathrm{s}}(x)$, so $T^{-1}W^{\mathrm{s}}(x)$ is a union of elements of $\mathcal{W}^{\mathrm{s}}$. Hence $T^{-1}\mathcal{F}_0\subseteq\mathcal{F}_0$. We denote by $L^\infty(\mathcal{F}_0)$ the set of functions in $L^\infty(X)$ that are $\mathcal{F}_0$-measurable.

Theorem 3.1. *Let $\phi\in L^\infty(X)$ be a mean zero observable. Assume that there exists $\beta>1$ and $C>0$ such that for all $n\geqslant 1$,*

(a) $|\int_X \phi\,\psi\circ T^n\mathrm{d}\mu|\leqslant C|\psi|_\infty n^{-\beta}$ *for all* $\psi\in L^\infty(\mathcal{F}_0)$.
(b) $\int_X \operatorname{diam}(\phi(T^nW^{\mathrm{s}}))\mathrm{d}\mu\leqslant Cn^{-\beta}$.

Then the conditions in (2.1) *are satisfied for all* $1\leqslant p<\beta$, *and the conditions in* (2.3) *are satisfied for all* $1\leqslant p<2\beta$.

Proof. This is a standard argument. We again give the details for completeness.

Let

$$\xi=|\mathbb{E}[\phi|T^{-n}\mathcal{F}_0]|^{p-1}\operatorname{sgn}\mathbb{E}[\phi|T^{-n}\mathcal{F}_0]=\psi\circ T^n,$$

where

$$\psi=|\mathbb{E}[\phi\circ T^{-n}|\mathcal{F}_0]|^{p-1}\operatorname{sgn}\mathbb{E}[\phi\circ T^{-n}|\mathcal{F}_0]\in L^\infty(\mathcal{F}_0),$$

and $|\psi|_\infty\leqslant|\phi|_\infty^{p-1}$. Then

$$\begin{aligned}|\mathbb{E}[\phi\circ T^{-n}|\mathcal{F}_0]|_p^p=|\mathbb{E}[\phi|T^{-n}\mathcal{F}_0]|_p^p&=\int_X\mathbb{E}[\phi|T^{-n}\mathcal{F}_0]\xi\mathrm{d}\mu\\&=\int_X\mathbb{E}[\phi\xi|T^{-n}\mathcal{F}_0]\mathrm{d}\mu=\int_X\phi\,\xi\,d\mu=\int_X\phi\,\psi\circ T^n\mathrm{d}\mu.\end{aligned}$$

By assumption (a),

$$|\mathbb{E}[\phi\circ T^{-n}|\mathcal{F}_0]|_p^p=\left|\int_X\phi\,\psi\circ T^n\mathrm{d}\mu\right|\leqslant C|\psi|_\infty n^{-\beta}\leqslant C|\phi|_\infty^{p-1}n^{-\beta},$$

and the first part of conditions (2.1) and (2.3) follows by taking pth roots and using the restriction on p.

Next, using the pointwise estimate $|\mathbb{E}[\phi|T^n\mathcal{F}_0]-\phi|\leqslant \operatorname{diam}(\phi(T^nW^s))$ and assumption (b),

$$|\mathbb{E}[\phi\circ T^n|\mathcal{F}_0]-\phi\circ T^n|_p^p=|\mathbb{E}[\phi|T^n\mathcal{F}_0]-\phi|_p^p\leqslant|\operatorname{diam}(\phi(T^nW^s))|_p^p$$
$$\leqslant(2|\phi|_\infty)^{p-1}|\operatorname{diam}(\phi(T^nW^s))|_1\leqslant 2^{p-1}C|\phi|_\infty^{p-1}n^{-\beta}.$$

The second part of conditions (2.1) and (2.3) follows. □

In the remainder of this section, we show that the conditions in theorem 3.1 are satisfied in many standard situations. (The verifications below are not needed for our main example in section 4.)

3.1. Verifying condition (b) in theorem 3.1

Suppose that $T:X\to X$ and $\mathcal{W}^s$ are as above. Let $Y\subseteq X$ be a positive measure subset that is a union of local stable leaves in $\mathcal{W}^s$. Define the first return time $R:Y\to\mathbb{Z}^+$ and first return map $F:Y\to Y$,

$$R(y)=\inf\{n\geqslant 1: T^ny\in Y\},\qquad F(y)=T^{R(y)}y.$$

Let h_n be the random variable on X given by $h_n(X)=\#\{0\leqslant j\leqslant n: T^jx\in Y\}$.

Lemma 3.2. *Let $\phi:X\to\mathbb{R}$ be measurable. Suppose that $\mu(y\in Y:R(y)>n)=O(n^{-(\beta+1)})$ for some $\beta>1$ and that there are constants $C\geqslant 1$, $\gamma\in(0,1)$ such that*

$$|\operatorname{diam}(\phi(T^nW^s))|\leqslant C\gamma^{h_n(x)}\quad \text{for all } W^s\in\mathcal{W}^s,\ n\geqslant 1.$$

Then condition (b) in theorem 3.1 holds.

Proof. We have

$$\int_X\operatorname{diam}(\phi(T^nW^s))\,\mathrm{d}\mu\leqslant C\sum_{k=0}^{n+1}\gamma^k\int_X 1_{\{h_n=k\}}\mathrm{d}\mu\leqslant C\sum_{k=1}^{n+1}\gamma^k\int_Y 1_{\{h_n=k\}}R\,\mathrm{d}\mu.$$

If $y\in Y\cap\{h_n=k\}$, then $\sum_{j=0}^{k-1}R\circ F^j>n$, and so $R\circ F^j>\frac{n}{k}$ for some $j=0,\dots,k-1$. Hence

$$\int_Y 1_{\{h_n=k\}}R\,\mathrm{d}\mu\leqslant\sum_{j=0}^{k-1}\int_Y 1_{\left\{R\circ F^j\geqslant\frac{n}{k}\right\}}R\,\mathrm{d}\mu.$$

It follows from the tail assumption on R that there is a constant $C_1>0$ such that $\mu(y\in Y:R(y)>n)\leqslant C_1n^{-(\beta+1)}$ and $\int_Y 1_{\{R>n\}}R\mathrm{d}\mu\leqslant C_1n^{-\beta}$. Write $R=1_{\{R\leqslant n\}}R+1_{\{R>n\}}R$. Then

$$\int_Y 1_{\left\{R\circ F^j\geqslant\frac{n}{k}\right\}}R\,\mathrm{d}\mu\leqslant\int_Y n1_{\left\{R\circ F^j\geqslant\frac{n}{k}\right\}}\mathrm{d}\mu+\int_Y 1_{\{R>n\}}R\,\mathrm{d}\mu=n\mu\left(R\geqslant\frac{n}{k}\right)+\int_Y 1_{\{R>n\}}R\,\mathrm{d}\mu$$
$$\leqslant C_1k^{\beta+1}n^{-\beta}+C_1n^{-\beta}\leqslant 2C_1k^{\beta+1}n^{-\beta}.$$

Therefore, $\int_Y 1_{\{h_n=k\}}R\,\mathrm{d}\mu\leqslant 2C_1k^{\beta+2}n^{-\beta}$, and

$$\int_X\operatorname{diam}\left(\phi(T^nW^s)\right)\,\mathrm{d}\mu\leqslant 2CC_1n^{-\beta}\sum_{k=1}^{\infty}\gamma^kk^{\beta+2}=O(n^{-\beta}),$$

as required. □

3.2. Verifying condition (a) in theorem 3.1

For completeness, we show that theorem 3.1 includes examples that fit within the Chernov–Markarian–Zhang setup [13, 14, 33] (in the summable decay of correlations regime, so $\beta > 1$) for Hölder mean zero observables $\phi : X \to \mathbb{R}$. In particular, we recover limit theorems that have been obtained previously for such invertible examples [27, 37, 41, 42]. Since there are no new results here, we only sketch the construction from [13, 33].

Remark 3.3. When treating examples falling within the Chernov–Markarian–Zhang setup, a significant (over) simplification is to suppose that there is exponential (or rapid) contraction of stable leaves under the underlying dynamics. For billiards with subexponential decay of correlations, such a condition fails since on average stable directions contract as slowly as unstable directions expand. In general, one should assume that there is an inducing set (called Y below) such that expansion *and* contraction occurs only on visits to Y. This general point of view is the one adopted here, as codified by the random variable h_n in lemma 3.2.

It is part of the setup that X is a metric space and $T : X \to X$ is the canonical billiard map corresponding to the first collision with the boundary of the billiard table. It is assumed (and for many classes of billiards explicitly constructed) that there is a set $Y \subset X$ and a first return map $F = T^R : Y \to Y$ such that F is uniformly hyperbolic and the return time has tail bounds satisfying $\mu(R > n) = O(n^{-(\beta_0+1)})$, where we assume that $\beta_0 > 1$ (see [13, section 4]). Moreover, Y is modelled by a Young tower with exponential tails [50]. A standard argument (see for example [13, theorem 4]) shows that $T : X \to X$ is modelled by a Young tower $f : \Delta \to \Delta$ with polynomial tails [51], with tail rate $O(n^{-(\beta+1)})$ for all $\beta < \beta_0$. In particular, there is a measure-preserving semiconjugacy $\pi : \Delta \to X$, so we can work with $f : \Delta \to \Delta$ instead of $T : X \to X$ and observables $\hat{\phi} = \phi \circ \pi : \Delta \to \mathbb{R}$ where $\phi : X \to \mathbb{R}$ is Hölder.

The final part of the set up that we require is that Δ is covered by stable leaves $\mathcal{W}^s$ satisfying $T(W(x)) \subseteq W(Tx)$, for all $x \in \Delta$, where $W(x)$ is the element of $\mathcal{W}^s$ containing x. Due to the uniform hyperbolicity of $F = T^R$, the contraction condition in lemma 3.2 holds [13, section 4.2]. Hence $f : \Delta \to \Delta$ satisfies condition (b) of theorem 3.1 and it remains to verify condition (a).

Let $\bar{f} : \bar{\Delta} \to \bar{\Delta}$ denote the quotient (one-sided) Young tower obtained by quotienting along stable leaves. Consider observables $\bar{\phi} : \bar{\Delta} \to \mathbb{R}$ that are Lipschitz with respect to a symbolic metric on $\bar{\Delta}$, with Lipschitz norm $\|\bar{\phi}\|$. By [51, theorem 3], there is a constant $C > 0$ such that

$$\left| \int_{\bar{\Delta}} \bar{\phi}\, \bar{\psi} \circ \bar{f}^n \, \mathrm{d}\bar{\mu}_\Delta - \int_{\bar{\Delta}} \bar{\phi} \, \mathrm{d}\bar{\mu}_\Delta \int_{\bar{\Delta}} \bar{\psi} \, \mathrm{d}\bar{\mu}_\Delta \right| \leqslant C\|\bar{\phi}\||\bar{\psi}|_\infty n^{-\beta}, \tag{3.1}$$

for all $\bar{\phi} : \bar{\Delta} \to \mathbb{R}$ Lipschitz, $\bar{\psi} \in L^\infty(\bar{\Delta})$, $n \geqslant 1$. (The dependence on $\|\bar{\phi}\|$ and $|\bar{\psi}|_\infty$ is not stated explicitly in [51, theorem 3] but follows by a standard argument using the uniform boundedness principle. Alternatively, see [30] for a direct argument).

Returning to the two-sided tower $f : \Delta \to \Delta$ and the lifted observable $\hat{\phi} = \phi \circ \pi : \Delta \to \mathbb{R}$, it follows for instance from [30, proposition 5.3] that there exists a choice (depending only on the Hölder exponent of ϕ) of symbolic metric on $\bar{\Delta}$ and a sequence of observables $\tilde{\phi}_\ell \in L^\infty(\Delta)$, $\ell \geqslant 1$, such that

(i) $\tilde{\phi}_\ell$ is $\mathcal{F}_0$-measurable and hence projects down to an observable $\bar{\phi}_\ell : \bar{\Delta} \to \mathbb{R}$.
(ii) $\sup_{\ell \geqslant 1} \|L^\ell \bar{\phi}_\ell\| < \infty$.
(iii) $\lim_{\ell \to \infty} |\phi \circ f^\ell - \tilde{\phi}_\ell|_1 = 0$.

Here, $\mathcal{F}_0$ is the σ-algebra generated by $\mathcal{W}^s$ and L is the transfer operator corresponding to $\bar{f}:\bar{\Delta}\to\bar{\Delta}$.

Let $\psi\in L^\infty(\mathcal{F}_0)$ with projection $\bar{\psi}\in L^\infty(\bar{\Delta})$. Following [30, proof of corollary 5.4],

$$\int_\Delta \phi\,\psi\circ f^n\,\mathrm{d}\mu_\Delta=\int_\Delta \phi\circ f^\ell\,\psi\circ f^{\ell+n}\,\mathrm{d}\mu_\Delta=I_1+I_2+I_3,$$

where

$$I_1=\int_\Delta(\phi\circ f^\ell-\tilde{\phi}_\ell)\,\psi\circ f^{\ell+n}\,\mathrm{d}\mu_\Delta,\qquad I_2=\int_\Delta\tilde{\phi}_\ell\,\mathrm{d}\mu_\Delta\int_\Delta\psi\,d\mu_\Delta,$$
$$I_3=\int_\Delta\tilde{\phi}_\ell\,\psi\circ f^{\ell+n}\,\mathrm{d}\mu_\Delta-\int_\Delta\tilde{\phi}_\ell\,\mathrm{d}\mu_\Delta\int_\Delta\psi\,\mathrm{d}\mu_\Delta.$$

Now $|I_1|\leqslant|\phi\circ f^\ell-\tilde{\phi}_\ell|_1|\psi|_\infty$. Also, $I_2=\int_\Delta(\tilde{\phi}_\ell-\phi\circ f^\ell)\,\mathrm{d}\mu_\Delta\int_\Delta\psi\,\mathrm{d}\mu_\Delta$, so $|I_2|\leqslant|\phi\circ f^\ell-\tilde{\phi}_\ell|_1|\psi|_1$. By (iii), $\lim_{\ell\to\infty}I_j=0$ for $j=1,2$. By (i),

$$\begin{aligned}I_3&=\int_{\bar{\Delta}}\bar{\phi}_\ell\,\bar{\psi}\circ\bar{f}^{\ell+n}\,\mathrm{d}\bar{\mu}_\Delta-\int_{\bar{\Delta}}\bar{\phi}_\ell\,\mathrm{d}\bar{\mu}_\Delta\int_{\bar{\Delta}}\bar{\psi}\,\mathrm{d}\bar{\mu}_\Delta\\&=\int_{\bar{\Delta}}L^\ell\bar{\phi}_\ell\,\bar{\psi}\circ\bar{f}^n\,\mathrm{d}\bar{\mu}_\Delta-\int_{\bar{\Delta}}L^\ell\bar{\phi}_\ell\,\mathrm{d}\bar{\mu}_\Delta\int_{\bar{\Delta}}\bar{\psi}\,\mathrm{d}\bar{\mu}_\Delta,\end{aligned}$$

so by (3.1) and (ii), $|I_3|\leqslant C\|L^\ell\bar{\phi}_\ell\||\bar{\psi}|_\infty\,n^{-\beta}\ll|\psi|_\infty\,n^{-\beta}$. Together, these estimates establish condition (a) in theorem 3.1.

4. Application to Lorentz gases

In this section, we use the results of [3] to show that the hypotheses of theorem 3.1 (with $\beta>1$ arbitrarily large) are satisfied for the time-one map corresponding to a finite horizon planar periodic Lorentz gas for all Hölder observables ϕ. Hence the results of section 2 hold for all $p\geqslant 1$, establishing the results listed in the introduction.

4.1. Setting and main result for Lorentz gases

Let $\mathbb{T}^2=\mathbb{R}^2/\mathbb{Z}^2$ denote the two-torus, and let $B_i\subset\mathbb{T}^2$, $i=1,\ldots,d$, denote open convex sets such that their closures are pairwise disjoint and their boundaries are C^3 curves with strictly positive curvature. We refer to the sets B_i as scatterers. The billiard flow Φ_t is defined by the motion of a point particle in $Q=\mathbb{T}^2\backslash\bigcup_{i=1}^d B_i$ undergoing elastic collisions at the boundaries of the scatterers and moving at constant velocity with unit speed between collisions. Hence Φ_t is defined on the three dimensional phase space

$$X=Q\times\mathbb{S}^1,\qquad\mathbb{S}^1=[0,2\pi]/\sim,$$

where $\sim$ indicates that 0 and 2π are identified.

Between collisions, $\Phi_t(x_1,x_2,\theta)=(x_1+t\cos\theta,x_2+t\sin\theta,\theta)$, while at collisions the point (x,θ^-) becomes (x,θ^+) where θ^- and θ^+ are the pre- and post-collisions angles, respectively. Defining $X_0=X/\sim$, where we identify $(x,\theta^-)\sim(x,\theta^+)$ at collisions, we obtain a continuous flow $\Phi_t:X_0\to X_0$.

Let $M=\bigcup_{i=1}^d\partial B_i\times[-\pi/2,\pi/2]$. The billiard map $F:M\to M$ is the discrete-time map which maps one collision to the next. Parametrizing each ∂B_i by an arclength coordinate r (oriented clockwise) and letting φ denote the angle that the post-collision velocity vector makes

with the normal to the scatterer (directed inwards in Q), we obtain the standard coordinates (r, φ) on M.

For $x \in X$, define the collision time $\tau(x)$ to be the first time $t > 0$ that $\Phi_t(x) \in M$. Since the closures of the scatterers are disjoint, there exists $\tau_{\min} > 0$ such that $\tau(x) \geqslant \tau_{\min}$ for all $x \in M$. In addition, we assume that the billiard has finite horizon so that there exists $\tau_{\max} < \infty$ such that $\tau(x) \leqslant \tau_{\max}$ for all $x \in X$.

It is well known (see [12, section 3.3]) that the flow preserves the contact form

$$\omega = \cos\theta \; \mathrm{d}x_1 + \sin\theta \; \mathrm{d}x_2,$$

so that the contact volume is $\omega \wedge \mathrm{d}\omega = \mathrm{d}x_1 \wedge \mathrm{d}\theta \wedge \mathrm{d}x_2$. We denote by μ the normalized Lebesgue measure on X, which by the preceding calculation is preserved by the flow.

The main result of this section is the following.

Theorem 4.1. *Let T be the time-one map corresponding to a finite horizon Lorentz gas as described above, and let $\phi : X \to \mathbb{R}$ be a mean zero Hölder observable. Then conditions (a) and (b) of theorem 3.1 hold with $n^{-\beta}$ replaced by e^{-cn} for some $c > 0$.*

As a consequence, conditions (2.1) and (2.3) hold for all $p \geqslant 1$, and all the results described in section 2 apply in this setting.

We remark that the observable ϕ is assumed to be Hölder continuous only on X, not X_0. Thus ϕ is allowed to be discontinuous at the boundary of X, i.e. at collisions. In particular, theorem 4.1 applies to the velocity.

4.2. Proof of theorem 4.1

The remainder of this section is devoted to the proof of theorem 4.1, which consists of verifying the conditions of theorem 3.1. First we recall some of the essential properties and main constructions used in [3].

4.2.1. Hyperbolicity and singularities. The singularities for both the collision map and the flow are created by tangential collisions with the scatterers. Let $\mathcal{S}_0 = \left\{(r, \varphi) \in M : \varphi = \pm\frac{\pi}{2}\right\}$. Away from the set $\mathcal{S}_1 = \mathcal{S}_0 \cup F^{-1}\mathcal{S}_0$ (resp. $\mathcal{S}_{-1} = \mathcal{S}_0 \cup F\mathcal{S}_0$) the map F (resp. F^{-1}) is uniformly hyperbolic: letting

$$\Lambda = 1 + 2\tau_{\min}\mathcal{K}_{\min}, \tag{4.1}$$

where $\mathcal{K}_{\min}$ denotes the minimum curvature of the scatterers, there exist stable $\bar{C}^{\mathrm{s}}$ and unstable $\bar{C}^{\mathrm{u}}$ cones in the tangent space of M such that stable and unstable vectors in these cones undergo uniform expansion and contraction at an exponential rate given by Λ. Flowing $\bar{C}^{\mathrm{s}}$ backward and $\bar{C}^{\mathrm{u}}$ forward between collisions allows us to define two families of stable C^{s} and unstable C^{u} cones for the flow that lie in the kernel of the contact form. (Hence they are 'flat' two-dimensional cones in the tangent space of the flow; see [3, section 2.1] for an explicit definition of these cones.)

Let $P^{\pm}$ denote the projections from X onto M under the forward and backward flow. Then C^{u} is continuous on X away from the surface $\mathcal{S}_{-1}^{-} = \{x \in X : P^{+}(x) \in \mathcal{S}_{-1}\}$, and C^{s} is continuous on X away from the surface $\mathcal{S}_{1}^{+} = \{x \in X : P^{-}(x) \in \mathcal{S}_{1}\}$. To maintain control of distortion, we define the standard homogeneity strips

$$\mathbb{H}_k = \left\{(r, \varphi) \in M : \frac{\pi}{2} - \frac{1}{k^2} \leqslant \varphi \leqslant \frac{\pi}{2} - \frac{1}{(k+1)^2}\right\}, \quad k \geqslant k_0,$$

for some $k_0 \geqslant 1$ which is determined to ensure a one-step expansion condition. A similar set of homogeneity strips $\mathbb{H}_{-k}$, $k \geqslant k_0$, is defined for φ near $-\frac{\pi}{2}$.

Following [3], we define a set of admissible stable curves $\mathcal{A}^s$ for the flow. A C^2 curve W belongs to $\mathcal{A}^s$ if the tangent vector at each point of W belongs to C^s, and W has curvature bounded by B_0 and length $|W|$ bounded by δ_0. Here, $\delta_0 > 0$ is chosen to satisfy a complexity bound (see [3, lemma 3.8]) and B_0 is chosen large enough that the family $\mathcal{A}^s$ is invariant under Φ_{-t}, $t \geqslant 0$ (once long pieces are subdivided according to the length δ_0). We call $W \in \mathcal{A}^s$ *homogeneous* if $P^+(W)$ lies in a single homogeneity strip.

We define $\mathcal{W}^s$ to be the family of maximal C^2 connected homogeneous stable manifolds for the flow. Note that $\mathcal{W}^s$ forms a partition of X (mod μ-measure 0). Moreover, each element of $\mathcal{W}^s$ (up to subdivision due to the length δ_0) belongs to $\mathcal{A}^s$. When we define a homogeneous stable manifold $W \in \mathcal{W}^s$, we take into account cuts introduced at the boundary of the extended singularity set, which includes the boundaries of the homogeneity strips. Thus $P^+(\Phi_t W)$ lies in a single homogeneity strip for all $t \geqslant 0$.[8] Let $\mathcal{F}_0$ denote the sigma algebra generated by elements of $\mathcal{W}^s$. Since $\mathcal{W}^s$ forms a partition of X, it follows that $\mathcal{F}_0$ comprises countable unions of elements of $\mathcal{W}^s$.

4.2.2. Norms and Banach spaces. With the class of admissible stable curves defined, we can now describe the Banach spaces used to prove decay of correlations in [3].

Let $\alpha \in (0, \frac{1}{3}]$. For $W \in \mathcal{A}^s$, let $C^\alpha(W)$ denote the closure of C^1 functions in the Holder norm defined by

$$|\psi|_{C^\alpha(W)} = \sup_{x \in W} |\psi(x)| + \sup_{\substack{x,x' \in W \\ x \neq x'}} |\psi(x) - \psi(x')|\, d_W(x,x')^{-\alpha},$$

where d_W is arclength distance along W. Define the weak norm of $\phi \in C^0(X)$ by

$$|\phi|_w = \sup_{W \in \mathcal{A}^s} \sup_{\substack{\psi \in C^\alpha(W) \\ |\psi|_{C^\alpha(W)} \leqslant 1}} \int_W \phi\, \psi \, \mathrm{d}m_W,$$

where m_W denotes arclength measure on W. The weak space $\mathcal{B}_w$ is defined as the completion of the set $\{\phi \in C^0(X_0) : |\phi|_{\mathrm{w}} < \infty\}$.

The strong norm $\|\phi\|_{\mathcal{B}}$ is defined as in [3, section 2.3]. The space $\mathcal{B}$ is similarly defined as the completion of a class of smooth functions on X_0 in the $\|\cdot\|_{\mathcal{B}}$ norm. Since we do not need the precise definition of $\|\cdot\|_{\mathcal{B}}$ here, we omit its definition; however, the following lemma summarizes some of the important properties of these spaces.

Lemma 4.2 ([3, *lemmas 3.9 and 3.10*]). *We have the inclusions*

$$C^1(X) \cap C^0(X_0) \subset \mathcal{B} \subset \mathcal{B}_w \subset (C^\alpha(X))^*,$$

where the first two inclusions are injective. Moreover, $|\cdot|_w \leqslant \|\cdot\|_{\mathcal{B}} \leqslant C|\cdot|_{C^1(X)}$ *and the unit ball of* $\mathcal{B}$ *is compactly embedded in* $\mathcal{B}_w$.

When we refer to functions $\phi \in C^0(X)$ as elements of $\mathcal{B}$ or $\mathcal{B}_w$, we identify ϕ with the measure $\phi \mathrm{d}\mu$. With this identification, the two definitions of $\mathcal{L}_t \phi$ given in the next section are reconciled.

[8] Due to our definition of C^s, if $W \in \mathcal{A}^s$, then $P^+(W)$ is a stable curve for the map; and if $W \in \mathcal{W}^s$, then $P^+(W)$ is a local homogeneous stable manifold for the map.

The following lemma is central to our verification of condition (a) in theorem 3.1, and is a strengthening of [3, lemma 2.11]. Let $C^\alpha(\mathcal{W}^{\rm s})$ denote those functions which are in $C^\alpha(W)$ for all $W \in \mathcal{W}^{\rm s}$ with $|\psi|_{C^\alpha(\mathcal{W}^{\rm s})} = \sup_{W\in\mathcal{W}^{\rm s}} |\psi|_{C^\alpha(W)}$ finite.

Lemma 4.3. *There exists $C > 0$ such that for $\phi \in \mathcal{B}_{\rm w}$ and $\psi \in C^\alpha(\mathcal{W}^{\rm s})$,*

$$|\phi(\psi)| \leqslant C|\phi|_{\rm w}|\psi|_{C^\alpha(\mathcal{W}^{\rm s})}.$$

Again, due to our identification, when $\phi \in C^0(X)$, we intend $\phi(\psi) = \int_X \phi\psi \mathrm{d}\mu$. Lemma 4.3 is proved at the end of this section.

4.2.3. Transfer operator. We define the transfer operator $\mathcal{L}_t$, for $t \geqslant 0$, by $\mathcal{L}_t\phi = \phi \circ \Phi_{-t}$, for $\phi \in C^0(X_0)$. This can be extended to any element of $\mathcal{B}_w$, and more generally a distribution of order α by

$$\mathcal{L}_t\phi(\psi) = \phi(\psi \circ \Phi_t), \quad \text{for all } \psi \in C^\alpha(\mathcal{A}^{\rm s}),\ \phi \in (C^\alpha(\mathcal{A}^{\rm s}))^*.$$

By [3, lemma 4.9], the map $(t, \phi) \mapsto \mathcal{L}_t\phi$ from $[0, \infty) \times \mathcal{B}$ to $\mathcal{B}$ is jointly continuous, so $\{\mathcal{L}_t\}_{t\geqslant 0}$ is a semi-group of bounded operators on $\mathcal{B}$.

Define the generator of the semi-group by $Z\phi = \lim_{t\downarrow 0} \frac{\mathcal{L}_t\phi - \phi}{t}$ for $\phi \in C^1(X)$. While Z is not a bounded operator on $\mathcal{B}$, the strong continuity of $\mathcal{L}_t$ implies that Z is closed with domain dense in $\mathcal{B}$. Indeed, by [3, lemma 7.5] the domain of Z contains all $\phi \in C^2(X) \cap C^0(X_0)$ such that $(\nabla\phi) \cdot \widehat{\eta} \in C^0(X_0)$, where $\widehat{\eta}$ denotes the flow direction, and there is a constant $C > 0$ such that

$$\|Z\phi\|_{\mathcal{B}} \leqslant C|\phi|_{C^2(X)} \quad \text{for all such } \phi. \tag{4.2}$$

4.2.4. Condition (b) of theorem 3.1. Recall that T and F denote the time-one map for the flow and the collision map, respectively. By the finite horizon condition, any $W \in \mathcal{W}^{\rm s}$ must undergo $k \geqslant \lfloor n/\tau_{\max} \rfloor$ collisions after n iterates by T. By [3, lemmas 3.3 and 3.4],

$$\operatorname{diam}(T^n W) = |\Phi_n W| \leqslant C|F^{k-1}(P^+(W))| \leqslant C\Lambda^{-(k-1)}|P^+(W)| \leqslant C'\Lambda^{-n/\tau_{\max}}|W|,$$

where $\Lambda > 1$ is the hyperbolicity constant defined in (4.1). We have used here that the lengths of $P^+(W)$ and W are bounded multiples of one another (indeed the Jacobian of this map is $C^{\frac{1}{2}}$, see [3, lemma 3.4]).

Let $\phi : X \to \mathbb{R}$ be C^η. Then $\operatorname{diam}(\phi(T^n W)) \leqslant |\phi|_\eta \operatorname{diam}(T^n W)^\eta \leqslant C\Lambda^{-n\eta/\tau_{\max}}$. Hence condition (b) holds with $n^{-\beta}$ replaced by $\Lambda^{-n\eta/\tau_{\max}}$.

4.2.5. Condition (a) of theorem 3.1. By [3, theorem 1.4], Z has a spectral gap on $\mathcal{B}$ and, using results of [10], $\mathcal{L}_t$ admits the following decomposition: there exists $\nu > 0$, a finite rank projector $\Pi : \mathcal{B} \to \mathcal{B}$ and a family of bounded operators $\mathcal{P}_t$ on $\mathcal{B}$ satisfying $\Pi\mathcal{P}_t = \mathcal{P}_t\Pi = 0$, and a matrix $\widehat{Z} : \Pi(\mathcal{B}) \to \Pi(\mathcal{B})$ with eigenvalues $0, z_1, \ldots, z_N \in \mathbb{C}$ satisfying $\mathrm{Re} z_j < -\nu$ for $j = 1, \ldots, N$, such that

$$\mathcal{L}_t = \mathcal{P}_t + \mathrm{e}^{t\widehat{Z}}\Pi \quad \text{for all } t \geqslant 0. \tag{4.3}$$

Moreover, there exists $C_\nu > 0$ such that for all ϕ in $\mathrm{Dom}(Z) \subset \mathcal{B}$,

$$|\mathcal{P}_t\phi|_{\rm w} \leqslant C_\nu \mathrm{e}^{-\nu t}\|Z\phi\|_{\mathcal{B}} \quad \text{for all } t \geqslant 0. \tag{4.4}$$

Now suppose $\phi \in C^2(X) \cap C^0(X_0)$ is of mean zero with $(\nabla\phi)\cdot\widehat{\eta} \in C^0(X_0)$, and $\psi \in C^\alpha(\mathcal{W}^{\rm s})$. By (4.3),

$$\int_X \phi\,\psi\circ\Phi_t\,\mathrm{d}\mu = \int_X \mathcal{L}_t\phi\,\psi\,\mathrm{d}\mu = \int_X \mathcal{P}_t\phi\,\psi\,\mathrm{d}\mu + \int_X \mathrm{e}^{t\widehat{Z}}\Pi\phi\,\psi\,\mathrm{d}\mu.$$

Hence by lemma 4.3,

$$\left|\int_X \phi\,\psi\circ\Phi_t\,\mathrm{d}\mu\right| \leqslant C\left\{|\mathcal{P}_t\phi|_{\rm w} + |\mathrm{e}^{t\widehat{Z}}\Pi\phi|_w\right\}|\psi|_{C^\alpha(\mathcal{W}^{\rm s})}. \tag{4.5}$$

Letting Π_0 denote the projector corresponding to the simple eigenvalue 0, we see that $\Pi_0\phi = \int\phi\mathrm{d}\mu = 0$ since μ is the conformal probability measure with respect to $\mathcal{L}_t$. Hence by lemma 4.2,

$$|\mathrm{e}^{t\widehat{Z}}\Pi\phi|_{\rm w} = |\mathrm{e}^{t\widehat{Z}}(\Pi-\Pi_0)\phi|_{\rm w} \leqslant C\mathrm{e}^{-\nu t}|\phi|_{\rm w} \leqslant C'\mathrm{e}^{-\nu t}|\phi|_{C^1(X)}.$$

By (4.2) and (4.4),

$$|\mathcal{P}_t\phi|_{\rm w} \leqslant C_\nu\mathrm{e}^{-\nu t}\|Z\phi\|_{\mathcal{B}} \leqslant C'\mathrm{e}^{-\nu t}|\phi|_{C^2(X)}.$$

Substituting these estimates in (4.5),

$$\left|\int_X \phi\,\psi\circ T^n\,\mathrm{d}\mu\right| = \left|\int_X \phi\,\psi\circ\Phi_n\,\mathrm{d}\mu\right| \leqslant C\mathrm{e}^{-\nu n}|\phi|_{C^2(X)}|\psi|_{C^\alpha(\mathcal{W}^{\rm s})} \quad \text{for all } n \geqslant 0.$$

The result extends to $\phi \in C^\eta(X)$ as in [3] by a standard mollification argument. (Exponential contraction persists with a rate dependent on η.) In particular, there are constants $c, C > 0$ such that

$$\left|\int_X \phi\,\psi\circ T^n\,\mathrm{d}\mu\right| \leqslant C\mathrm{e}^{-cn}|\phi|_{C^\eta(X)}|\psi|_{C^\alpha(\mathcal{W}^{\rm s})} \quad \text{for all } n \geqslant 0,$$

for all $\phi \in C^\eta(X)$, $\psi \in C^\alpha(\mathcal{W}^{\rm s})$.

Let $K(\mathcal{W}^{\rm s})$ denote the set of bounded functions on X that are constant on elements of $\mathcal{W}^{\rm s}$, and let $|\cdot|_{C^0(\mathcal{W}^{\rm s})} = \sup_{W\in\mathcal{W}^{\rm s}}|\cdot|_{C^0(W)}$. Note that these functions are $\mathcal{F}_0$-measurable. Moreover, $K(\mathcal{W}^{\rm s}) \subset C^\alpha(\mathcal{W}^{\rm s})$ and $|\psi|_{C^0(\mathcal{W}^{\rm s})} = |\psi|_{C^\alpha(\mathcal{W}^{\rm s})}$ for $\psi \in K(\mathcal{W}^{\rm s})$. Hence

$$\left|\int_X \phi\,\psi\circ T^n\,\mathrm{d}\mu\right| \leqslant C\mathrm{e}^{-cn}|\phi|_{C^\eta(X)}|\psi|_{C^0(\mathcal{W}^{\rm s})} \quad \text{for all } n \geqslant 0,$$

for all $\phi \in C^\eta(X)$, $\psi \in K(\mathcal{W}^{\rm s})$.

Finally, let $\phi \in C^\eta(X)$, $\psi \in L^\infty(\mathcal{F}_0)$. Recall that $L^\infty(\mathcal{F}_0)$ is the set of functions in $L^\infty(\mu)$ which are $\mathcal{F}_0$-measurable, so there exists a pointwise representative ψ' in the equivalence class of ψ in $L^\infty(\mu)$ that is constant on local stable manifolds and such that $\sup|\psi'| = |\psi|_\infty$. In particular, $\psi' \in K(\mathcal{W}^{\rm s})$ with $|\psi'|_{C^0(\mathcal{W}^{\rm s})} = |\psi|_\infty$ and $\int_X|\psi-\psi'|\mathrm{d}\mu = 0$. Hence

$$\left|\int_X \phi\,\psi\circ T^n\,\mathrm{d}\mu\right| = \left|\int_X \phi\,\psi'\circ T^n\,\mathrm{d}\mu\right| \leqslant C\mathrm{e}^{-cn}|\phi|_{C^\eta(X)}|\psi|_\infty.$$

Hence condition (a) holds with $n^{-\beta}$ replaced by e^{-cn}.

As promised, we end this section by proving lemma 4.3.

Proof of lemma 4.3. By density of $C^0(X_0)$ in $\mathcal{B}_w$, it suffices to prove the lemma for $\phi \in C^0(X_0)$ and $\psi \in C^\alpha(\mathcal{W}^{\rm s})$.

The normalized Lebesgue measure μ on X projects to the measure $\bar{\mu} = (2|\partial Q|)^{-1} \cos\varphi \, \mathrm{d}r\mathrm{d}\varphi$ on M; this is the unique smooth invariant probability measure for the billiard map F. Let $\overline{\mathcal{W}}^{\mathrm{s}}$ denote the set of maximal connected homogeneous stable manifolds for F. Note that $P^+(\mathcal{W}^{\mathrm{s}}) = \overline{\mathcal{W}}^{\mathrm{s}}$. Indexing elements of $\overline{\mathcal{W}}^{\mathrm{s}}$, we write $\overline{\mathcal{W}}^{\mathrm{s}} = \{V_\xi\}_{\xi\in\Xi}$, which defines a (mod 0) partition of M. We disintegrate $\bar{\mu}$ into conditional measures $\bar{\mu}_\xi$ on V_ξ, $\xi \in \Xi$, and a factor measure λ on Ξ. Indeed, the conditional measures are smooth on each V_ξ, and we can write

$$\mathrm{d}\bar{\mu}_\xi = \bar{\rho}_\xi \, \mathrm{d}\bar{m}_\xi \, \mathrm{d}\lambda(\xi),$$

where $\bar{m}_\xi$ is arclength measure along V_ξ (in M), and

$$|\log \bar{\rho}_\xi|_{C^{\frac{1}{3}}(V_\xi)} \leqslant C, \qquad |\bar{\rho}_\xi|_{C^0(V_\xi)} \leqslant C|V_\xi|^{-1}, \tag{4.6}$$

for some $C > 0$ depending only on the table Q (see [12, corollary 5.30]). The exponent $\frac{1}{3}$ comes from the definition of the homogeneity strips. This is the standard decomposition of $\bar{\mu}$ into a proper standard family[9] (see [12, example 7.21]). We further subdivide $\Xi = \bigcup_{i=1}^d \Xi_i$, where Ξ_i is the index set corresponding to each component $M_i = \partial B_i \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ of M.

Write $X = \bigcup_{i=1}^d X_i$ where $X_i = \{x \in X : P^+(X) \in M_i\}$. On each X_i, we represent Lebesgue measure as $\mathrm{d}\mu = c\cos\varphi \mathrm{d}r\mathrm{d}\varphi\mathrm{d}s$, where c is a normalizing constant, (r, φ) range over M_i, and s ranges from 0 to the maximum free flight time under the backwards flow, which we denote by $t_i \leqslant \tau_{\max}$.

Next, for each $\xi \in \Xi_i$, the flow surface $V_\xi^- = \{x \in X_i : P^+(x) \in V_\xi\}$ is smoothly foliated by elements of $\mathcal{W}^{\mathrm{s}}$, which are simply flow translates of one another. For each s and V_ξ, let $W_{\xi,s} = \Phi_{-t(s)}V_\xi$, where $t(s, Z)$ is defined for $z \in V_\xi$ so that $W_{\xi,s}$ lies in the kernel of ω, i.e. it is an element of $\mathcal{W}^{\mathrm{s}}$. Note that for $s < \delta_0$, some points in V_ξ may not have lifted off of M. For such small times, $W_{\xi,s}$ denotes only those points that have lifted off of M. Similarly, for $s > \tau_{\min}$, some part of $\Phi_{-t(s)}V_\xi$ may have collided with a scatterer. For such times, $W_{\xi,s}$ only denotes those points which have not yet undergone a collision. Thus $\bigcup_{s\in[0,t_i]} W_{\xi,s} = V_\xi^-$.

Using this decomposition, we may represent Lebesgue measure on each X_i by

$$\mathrm{d}\mu(x) = \rho_\xi(x) \, \mathrm{d}m_{W_{\xi,s}}(x) \, \mathrm{d}\lambda(\xi) \, \mathrm{d}s,$$

where ρ_ξ is smooth along each $W_{\xi,s}$, satisfying analogous bounds to (4.6), since the contact form is C^∞ on X_i and the projection P^+ is sufficiently smooth (see [3, lemma 3.4]), so that the arclength of $W_{\xi,s}$ varies smoothly with that of V_ξ.

Using the fact that each $W_{\xi,s} \in \mathcal{W}^{\mathrm{s}}$ can be subdivided into at most $C\delta_0^{-1}$ elements of $\mathcal{A}^{\mathrm{s}}$, we are ready to estimate

$$\begin{aligned}
\left|\int_X \phi\,\psi \,\mathrm{d}\mu\right| &\leqslant \sum_{i=1}^d \left|\int_{X_i} \phi\,\psi\,\mathrm{d}\mu\right| \leqslant \sum_i \left|\int_0^{t_i}\int_{\Xi_i}\int_{W_{\xi,s}} \phi\,\psi\,\rho_\xi \,\mathrm{d}m_{W_{\xi,s}}\,\mathrm{d}\lambda(\xi)\,\mathrm{d}s\right| \\
&\leqslant \sum_i \int_0^{t_i}\int_{\Xi_i} C\delta_0^{-1}|\phi|_w|\psi|_{C^\alpha(W_{\xi,s})}|\rho_\xi|_{C^\alpha(W_{\xi,s})}\,\mathrm{d}\lambda(\xi)\,\mathrm{d}s \\
&\leqslant C\delta_0^{-1}\tau_{\max}|\phi|_w|\psi|_{C^\alpha(\mathcal{W}^{\mathrm{s}})}\int_\Xi |V_\xi|^{-1}\mathrm{d}\lambda(\xi).
\end{aligned}$$

[9] Standard families in [12] are standard pairs defined on local unstable manifolds, while here we use local stable manifolds. The decompositions of μ have equivalent properties due to the symmetry of the map F under time reversal.

This last integral is finite by [12, exercise 7.15] since our decomposition of $\bar{\mu}$ constitutes a proper standard family, yielding the desired estimate for $\phi(\psi)$.

For completeness, we finish by proving [12, exercise 7.15]. For $x \in V_\xi$, let $r^s(X)$ denote the distance measured along V_ξ from x to the nearest endpoint of V_ξ. By [12, theorem 5.17], there exists $C_0 > 0$ such that

$$\sup_{\varepsilon>0} \frac{\bar{\mu}(x \in M : r^s(x) < \varepsilon)}{\varepsilon} \leqslant C_0.$$

We claim this quantity provides an upper bound on the relevant integral. To see this, we use the decomposition (4.6) to write,

$$\begin{aligned} C_0 &= \sup_{\varepsilon>0} \frac{\bar{\mu}(x \in M : r^s(x) < \varepsilon)}{\varepsilon} = \sup_{\varepsilon>0} \int_\Xi \frac{\bar{\mu}_\xi(r^s(x) < \varepsilon)}{\varepsilon} \, \mathrm{d}\lambda(\xi) \\ &\geqslant \sup_{\varepsilon>0} \int_\Xi \frac{C}{|V_\xi|} \frac{|V_\xi \cap \{r^s < \varepsilon\}|}{\varepsilon} \, \mathrm{d}\lambda(\xi) \\ &\geqslant \sup_{\varepsilon>0} 2C \int_{\{\xi : |V_\xi| > 2\varepsilon\}} \frac{1}{|V_\xi|} \, \mathrm{d}\lambda(\xi) = 2C \int_\Xi \frac{1}{|V_\xi|} \, \mathrm{d}\lambda(\xi), \end{aligned}$$

where we have used the fact that $|V_\xi| > 0$ for λ-a.e. ξ, and the bound $|V_\xi \cap \{r^s(X) < \varepsilon\}| = 2\varepsilon$ if $|V_\xi| > 2\varepsilon$. (One can also prove a reverse inequality, but we do not need this here.) □

Acknowledgments

We are very grateful to the referees for several helpful suggestions, especially with regard to the issue mentioned in remark 2.8. This research resulted from a Research in Pairs on *Techniques des martingales et espaces de Banach anisotropes* (*Martingale techniques and anisotropic Banach spaces*) at CIRM, Luminy, during August 2017. M D was supported in part by NSF Grant DMS 1800321. I M was supported in part by European Advanced Grant *StochExtHomog* (ERC AdG 320977). M N was supported in part by NSF Grant DMS 1600780.

References

[1] Araújo V, Melbourne I and Varandas P 2015 Rapid mixing for the Lorenz attractor and statistical limit laws for their time-1 maps *Commun. Math. Phys.* **340** 901–38
[2] Baladi V 2017 The quest for the ultimate anisotropic Banach space *J. Stat. Phys.* **166** 525–57
[3] Baladi V, Demers M F and Liverani C 2018 Exponential decay of correlations for finite horizon Sinai billiard flows *Invent. Math.* **211** 39–177
[4] Bálint P and Gouëzel S 2006 Limit theorems in the stadium billiard *Commun. Math. Phys.* **263** 461–512
[5] Bálint P and Melbourne I 2018 Statistical properties for flows with unbounded roof function, including the Lorenz attractor *J. Stat. Phys.* **172** 1101–126
[6] Blank M, Keller G and Liverani C 2002 Ruelle–Perron–Frobenius spectrum for Anosov maps *Nonlinearity* **15** 1905–973
[7] Bowen R 1975 *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms* (Lecture Notes in Math.) vol 470 (Berlin: Springer)
[8] Bunimovich L A, Sinaĭ Y G and Chernov N I 1991 Statistical properties of two-dimensional hyperbolic billiards *Uspekhi Mat. Nauk* **46** 43–92
[9] Burkholder D L 1973 Distribution function inequalities for martingales *Ann. Probability* **1** 19–42

[10] Butterley O 2016 A note on operator semigroups associated to chaotic flows *Ergodic Theory Dynam. Systems* **36** 1396–408
[11] Chernov N 2007 A stretched exponential bound on time correlations for billiard flows *J. Stat. Phys.* **127** 21–50
[12] Chernov N and Markarian R 2006 *Chaotic Billiards* (Mathematical Surveys and Monographs) vol 127 (Providence, RI: American Mathematical Society)
[13] Chernov N and Zhang H-K 2005 Billiards with polynomial mixing rates *Nonlinearity* **18** 1527–553
[14] Chernov N and Zhang H-K 2005 A family of chaotic billiards with variable mixing rates *Stochastics and Dynamics* **5** 535–53
[15] Chevyrev I, Friz P K, Korepanov A, Melbourne I and Zhang H 2019 Multiscale systems, homogenization, and rough paths. *Probability and Analysis in Interacting Physical Systems: In Honor of S.R.S. Varadhan, Berlin, August, 2016* (Springer Proceedings in Mathematics & Statistics) ed P Friz *et al* vol 283 pp 17–48
[16] Chevyrev I, Friz P K, Korepanov A, Melbourne I and Zhang H 2019 Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 2 (arxiv:1903.10418)
[17] Dedecker J, Merlevède F and Pène F 2013 Empirical central limit theorems for ergodic automorphisms of the torus *ALEA Lat. Am. J. Probab. Math. Stat.* **10** 731–66
[18] Dedecker J and Rio E 2000 On the functional central limit theorem for stationary processes *Ann. Inst. H. Poincaré Probab. Statist.* **36** 1–34
[19] Demers M F 2018 A gentle introduction to anisotropic Banach spaces *Chaos Solitons Fractals* **116** 29–42
[20] Dolgopyat D 1998 Prevalence of rapid mixing in hyperbolic flows *Ergodic Theory Dynam. Systems* **18** 1097–114
[21] Friz P K and Hairer M 2014 *A Course on Rough Paths* (Universitext) (Berlin: Springer)
[22] Gordin M I 1969 The central limit theorem for stationary processes *Soviet Math. Dokl.* **10** 1174–176
[23] Gouëzel S 2010 Almost sure invariance principle for dynamical systems by spectral methods *Ann. Probab.* **38** 1639–671
[24] Gouëzel S 2015 Limit theorems in dynamical systems using the spectral method hyperbolic dynamics, fluctuations and large deviations *Proc. Sympos. Pure Math.* (Providence, RI: Amer. Math. Soc.) vol 89 pp 161–93
[25] Hennion H and Hervé L 2001 *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness* (Lecture Notes in Math) vol 1766 (Berlin: Springer)
[26] Heyde C C 1975 On the central limit theorem and iterated logarithm law for stationary processes *Bull. Austral. Math. Soc.* **12** 1–8
[27] Kelly D and Melbourne I 2016 Smooth approximation of stochastic differential equations *Ann. Probab.* **44** 479–520
[28] Kelly D and Melbourne I 2017 Homogenization for deterministic fast-slow systems with multidimensional multiplicative noise *J. Funct. Anal.* **272** 4063–102
[29] Kipnis C and Varadhan S R S 1986 Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions *Commun. Math. Phys.* **104** 1–19
[30] Korepanov A, Kosloff Z and Melbourne I 2019 Explicit coupling argument for nonuniformly hyperbolic transformations *Proc. R. Soc. Edinburgh* A **149** 101–30
[31] Liverani C 1996 Central limit theorem for deterministic systems *International Conference on Dynamical Systems* ed F Ledrappier, J Lewowicz and S Newhouse (Pitman Research Notes in Math) vol 362 (Harlow: Longman Group Ltd) pp 56–75
[32] Lyons T J 1998 Differential equations driven by rough signals *Rev. Mat. Iberoamericana* **14** 215–310
[33] Markarian R 2004 Billiards with polynomial decay of correlations *Ergodic Theory Dynam. Systems* **24** 177–97
[34] Maxwell M and Woodroofe M 2000 Central limit theorems for additive functionals of Markov chains *Ann. Probab.* **28** 713–24
[35] Melbourne I 2007 Rapid decay of correlations for nonuniformly hyperbolic flows *Trans. Amer. Math. Soc.* **359** 2421–441
[36] Melbourne I 2018 Superpolynomial and polynomial mixing for semiflows and flows *Nonlinearity* **31** R268–316
[37] Melbourne I and Nicol M 2005 Almost sure invariance principle for nonuniformly hyperbolic systems *Commun. Math. Phys.* **260** 131–46
[38] Melbourne I and Nicol M 2008 Large deviations for nonuniformly hyperbolic systems *Trans. Amer. Math. Soc.* **360** 6661–676

[39] Melbourne I and Török A 2002 Central limit theorems and invariance principles for time-one maps of hyperbolic flows *Commun. Math. Phys.* **229** 57–71
[40] Melbourne I and Török A 2004 Statistical limit theorems for suspension flows *Isr. J. Math.* **144** 191–209
[41] Melbourne I and Török A 2012 Convergence of moments for Axiom A and nonuniformly hyperbolic flows *Ergodic Theory Dynam. Systems* **32** 1091–100
[42] Melbourne I and Varandas P 2016 A note on statistical properties for nonuniformly hyperbolic systems with slow contraction and expansion *Stoch. Dyn.* **16** 1660012
[43] Peligrad M and Utev S 2005 A new maximal inequality and invariance principle for stationary sequences *Ann. Probab.* **33** 798–815
[44] Ruelle D 1978 *Thermodynamic Formalism: The Mathematical Structures of Classical Equilibrium Statistical Mechanics* (Encyclopedia of Math. and its Applications) vol 5 (Reading, MA: Addison Wesley)
[45] Sinaĭ Y G 1972 Gibbs measures in ergodic theory *Russ. Math. Surv.* **27** 21–70
[46] Tyran-Kamińska M 2005 An invariance principle for maps with polynomial decay of correlations *Commun. Math. Phys.* **260** 1–15
[47] Viana M 1997 *Stochastic Dynamics of Deterministic Systems* (Rio de Janeiro: IMPA)
[48] Volný D 1993 Approximating martingales and the central limit theorem for strictly stationary processes *Stochastic Process. Appl.* **44** 41–74
[49] Volný D 2007 A nonadapted version of the invariance principle of Peligrad and Utev *C. R. Math. Acad. Sci. Paris* **345** 167–69
[50] Young L-S 1998 Statistical properties of dynamical systems with some hyperbolicity *Ann. Math.* **147** 585–650
[51] Young L-S 1999 Recurrence times and rates of mixing *Isr. J. Math.* **110** 153–88