NON-STATIONARY ALMOST SURE INVARIANCE PRINCIPLE FOR HYPERBOLIC SYSTEMS WITH SINGULARITIES

JIANYU CHEN, YUN YANG, AND HONG-KUN ZHANG

Dedicated to the memory of Nikolai Chernov.

Abstract. We investigate a wide class of two-dimensional hyperbolic systems with singularities, and prove the almost sure invariance principle (ASIP) for the random process generated by sequences of dynamically Hölder observables. The observables could be unbounded, and the process may be non-stationary and need not have linearly growing variances. Our results apply to Anosov diffeomorphisms, Sinai dispersing billiards and their perturbations. The random processes under consideration are related to the fluctuation of Lyapunov exponents, the shrinking target problem, etc.

Contents

1. Introduction

Since the pioneering work by Sinai [47] on dispersing billiards, the dynamical structures and stochastic properties have been extensively studied for chaotic billiards [31, 6, 7, 8, 10, 11, 41, 54, 18, 19, 12, 1, 17, 13, 20, 2, 53], and also for abstract

2010 *Mathematics Subject Classification.* 37D50, 37A25, 60F17.
Key words and phrases. ASIP, Non-stationarity, Hyperbolicity, Singularities.
H.-K. Zhang is partially supported by the NSF Career Award (DMS-1151762).

hyperbolic systems with or without singularities [48, 42, 37, 43, 46, 57, 58, 49, 24, 14, 21, 25, 26, 27]. Among all the physical measures, the SRB measures - named after Sinai [48], Ruelle [45] and Bowen [4, 5] - are shown to display several levels of ergodic properties, including the decay rate of correlations, the large deviation principles and the central limit theorem, etc.

In this paper, we shall focus on the almost sure invariance principle (ASIP) for a wide class of uniformly hyperbolic systems with singularities, which preserve a mixing SRB measure. The ASIP ensures that partial sum of a random process can be approximated by a Brownian motion with almost surely negligible error. More precisely, we say that a zero-mean random process $\mathbf{X} = \{X_n\}_{n\geq 0}$ with finite second moments satisfies an ASIP[1] for an error exponent $\lambda \in [0, \frac{1}{2})$, if there exists a Wiener process $W(\cdot)$ such that

$$\left|\sum_{k=0}^{n-1} X_k - W(\sigma_n^2)\right| = \mathcal{O}(\sigma_n^{2\lambda}), \quad \text{a.s.}, \tag{1.1}$$

where $\sigma_n^2 = \mathbb{E}\left(\sum_{k=0}^{n-1} X_k\right)^2$ is the variance of the n-th partial sum. In particular, if σ_n^2 grows linearly in n such that $\sigma_n^2 = n\sigma^2 + \mathcal{O}(1)$ for some $\sigma \in [0, \infty)$, it follows from (1.1) that

$$\left|\sum_{k=0}^{n-1} X_k - \sigma W(n)\right| = \mathcal{O}(n^{\lambda}), \quad \text{a.s.}.$$

The ASIP implies many other limit laws from statistics, such as the almost sure central limit theorem, the law of the iterated logarithm and the weak invariance principle (see [44] for the details).

There has been a great deal of work on the ASIP in the probability theroy, see for instance [44, 3, 29, 51, 50, 22, 55, 23]. In the context of the stationary process generated by bounded Hölder observables over smooth dynamical systems with singularities, the ASIP was first shown by Chernov [12] for Sinai dispersing billiards. Later, a scalar and a vector-valued ASIP were later proved by Melbourne and Nicol [39, 40] for the Young towers. By a purely spectral method, Gouëzel [32] extended the ASIP for stationary processes for a wide class of systems without assuming Young tower structure. Gouëzel [32] also provided a pure probabilistic condition, which was used by Stenlund [52] to show the ASIP for Sinai billiards with random scatterers.

The ultimate goal of our work is to establish ASIP for *non-stationary* process generated by *unbounded* observables, over a wide class of two-dimensional hyperbolic systems under the standard assumptions (**H1**)-(**H5**) in Section 2.1. Such class includes Anosov diffeomorphisms, Sinai dispersing billiards and their perturbations (See Section 5 for more details). Compared to existing results of ASIP for bounded stationary processes, our result is relatively new due to the following two features:

(1) Low regularity: the question on how large classes of observables satisfy the central limit theorem or other limit laws had been raised by several researchers, see, e.g., a survey by Denker [28]. Sometimes those classes are much larger than those of bounded Hölder continuous or bounded variation

[1] We may need to extend the partial sum process $\{\sum_{k=0}^{n-1} X_k\}_{n\geq 0}$ on a richer probability space without changing its distribution. In the rest of this paper, we always assume this technical operation when we mention ASIP.

functions. In this paper, we only assume dynamically Hölder continuity for observables, which could even be unbounded. A direct application is the fluctuation of Lyapunov exponents, for which the log unstable Jacobian blows up near singularity in billiard systems.

(2) Non-stationarity: time-dependent processes arise from the dynamical Borel-Cantelli Lemma and the shrinking target problem (see e.g. [35, 15, 30]). Recently, Haydn, Nicol, Török and Vaient [34] obtained ASIP for the shrinking target problem on a class of expanding maps. In analogy, under some mild conditions, we are able to apply our ASIP result to the shrinking target problem for two-dimensional hyperbolic systems with singularity.

The method we address here is rather transparent and efficient. We first construct a natural family of σ-algebras that are characterized by the singularities, and explore its exponentially α-mixing property. Extending the approach by Chernov [12] and applying a martingale version of ASIP by Shao [50], we are then able to prove the ASIP for the random process generated by a sequence of integrable dynamically Hölder observables. A crucial assumption is that the process satisfies the Marcinkiewicz-Zygmund type inequalities given by (2.1). We emphasize that those observables could be unbounded, and the growth of partial sum variances need not be linear. Furthermore, the error exponent λ in ASIP of the form (1.1) only depends on the constant κ_2 in (2.1).

This paper is organized as follows. In Section 2, we introduce the standard assumptions for the uniformly hyperbolic systems with singularities, and state our main results on the ASIP and other limit laws. We recall several useful theorems in probability theory in Section 3, and prove our main theorem on the ASIP in Section 4. In Section 5, we summarize the validity of the ASIP for a wide class of uniformly hyperbolic billiards, and discuss two practical process related to the fluctuation of ergodic average and the shrinking target problem.

2. Assumptions and Main Results

2.1. **Assumptions.** Let $T : M \to M$ be a piecewise C^2 diffeomorphism of a two-dimensional compact Riemannian manifold M with singularities S_1, that is, for each connected component $\Omega \subset M \setminus S_1$, the map $T : \Omega \to T(\Omega) \subset M$ is a C^2 diffeormophism which can be continuously extended to the closure of Ω. We denote by $S_{-1} := TS_1$ the singularity of the inverse map T^{-1}.

Let $d(\cdot, \cdot)$ denote the distance in M induced by the Riemannian metric. For any smooth curve $W \subset M$, we denote by $|W|$ its length, and by m_W the Lebesgue measure on W induced by the Riemannian metric restricted to W.

We now make several specific assumptions on the system $T : M \to M$. These assumptions are quite standard and have been made in many references [10, 14, 16, 21].

(H1) Uniform hyperbolicity of T. There exist two families of cones C^u_x (unstable) and C^s_x (stable) in the tangent spaces $\mathcal{T}_x M$, for all $x \in M$, and there exists a constant $\Lambda > 1$, with the following properties:

(1) $D_xT(C^u_x) \subset C^u_{Tx}$ and $D_xT(C^s_x) \supset C^s_{Tx}$, wherever D_xT exists.
(2) $\|D_xT(v)\| \geq \Lambda \|v\|$ for any $v \in C^u_x$, and $\|D_xT^{-1}(v)\| \geq \Lambda \|v\|$ for any $v \in C^s_x$.
(3) These families of cones are continuous on M and the angle between C^u_x and C^s_x is uniformly bounded away from zero.

We say that a smooth curve $W \subset M$ is an *unstable curve* for T if at every point $x \in W$ the tangent line $\mathcal{T}_x W$ belongs in the unstable cone C^u_x. Furthermore, a curve $W \subset M$ is an *unstable manifold* for T if $T^{-n}(W)$ is an unstable curve for all $n \geq 0$. We can define stable curves and stable manifolds in a similar fashion.

(**H2**) **Singularities.** The singularity set S_1 consists of a finite or countable union of smooth compact curves in M, including the boundary ∂M. We assume the following:

(1) ∂M is transversal to both stable and unstable cones.
(2) Every other smooth singularity curve in $S_1 \setminus \partial M$ is a stable curve, and every curve in S_1 terminates either inside another curve of S_1 or on ∂M.
(3) There exist $C > 0$ and $\beta_0 \in (0, 1)$ such that for any $x \in M \backslash S_1$,

$$\|D_x T\| \leq C d(x, S_1)^{-\beta_0}.$$

Similar assumptions are made for S_{-1}. We set $S_{\pm n} = \bigcup_{k=0}^{n-1} T^{\mp k} S_{\pm 1}$ for any $n \geq 1$, and it is clear that $S_{\pm n}$ be the singularity set of $T^{\pm n}$. Furthermore, we denote $S_{\pm\infty} = \bigcup_{n=0}^{\infty} S_{\pm n}$.

An unstable curve $W \subset M$ is said to be homogeneous if for any $n \geq 0$, $T^{-n}W$ is contained in a connected component of $M \setminus S_1$. In other words, $W \cap S_{-\infty} = \emptyset$. Similarly, we can define homogeneous stable curves.

Definition 1. For every $x, y \in M$, define $\mathbf{s}_+(x, y)$, the forward separation time for x and y, to be the smallest integer $n \geq 0$ such that x and y belong to distinct elements of $M \setminus S_n$. Similarly we define the backward separation time $\mathbf{s}_-(x, y)$.

(**H3**) **Regularity of smooth unstable curves**. We assume that there is a T-invariant family $\mathcal{W}^u_T$ of unstable curves such that

(1) **Bounded curvature** The curvature of any $W \in \mathcal{W}^u_T$ is uniformly bounded from above by a positive constant B.
(2) **Distortion bounds.** There exist $\gamma_0 \in (0, 1)$ and $C_T > 0$ such that for any $W \in \mathcal{W}^u_T$ and any $x, y \in W$,

$$|\ln \mathcal{J}_W(x) - \ln \mathcal{J}_W(y)| \leq C_T\, d(x, y)^{\gamma_0},$$

where $\mathcal{J}_W(x) = |D_x T|_{\mathcal{T}_x W}|$ is the Jacobian of T at x along the unstable curve W.
(3) **Absolute continuity.** Let $W_1, W_2 \in \mathcal{W}^u_T$ be two unstable curves close to each other. Denote

$$W_i' = \{x \in W_i \colon W^s(x) \cap W_{3-i} \neq \emptyset\}, \qquad i = 1, 2.$$

The map $\mathbf{h}\colon W_1' \to W_2'$ defined by sliding along stable manifolds is called the *stable holonomy* map. We assume that $\mathbf{h}_* m_{W_1'}$ is absolutely continuous with respect to $m_{W_2'}$. Furthermore, there exist $C_T > 0$ and $\vartheta_0 \in (0, 1)$ such that the Jacobian of $\mathbf{h}$ satisfies

$$|\ln \mathcal{J}\mathbf{h}(y) - \ln \mathcal{J}\mathbf{h}(x)| \leq C_T \vartheta_0^{\mathbf{s}_+(x,y)}, \quad \text{for any } \; x, y \in W_1'.$$

(**H4**) **SRB measure.** The map T preserves an SRB probability measure μ, that is, the conditional measure of μ on each unstable manifold W^u is absolutely continuous with respect to m_{W^u}. We further assume that μ is strongly mixing.

(H5) One-step expansion. Given an unstable curve $W \subset M$, we denote V_α as the connected component in TW with index α and $W_\alpha = T^{-1}V_\alpha$. There is $q_0 \in (0,1]$ such that

$$\liminf_{\delta\to 0} \sup_{W:\,|W|<\delta} \sum_\alpha \left(\frac{|W|}{|V_\alpha|}\right)^{q_0} \frac{|W_\alpha|}{|W|} < 1,$$

where the supremum is taken over all unstable curves W in M.

2.2. **Statement of the main results.** The main result in this paper is to prove the almost sure invariance principle for the system (M, T, μ), which satisfies Assumptions **(H1)-(H5)**, with respect to the process generated by a sequence of dynamically Hölder observables. We first recall the definition of such functions.

Definition 2. A measurable function $f : M \to \mathbb{R}$ is said to be forward dynamically Hölder continuous if there exists $\vartheta \in (0,1)$ such that

$$|f|^+_\vartheta := \sup\left\{\frac{|f(x)-f(y)|}{\vartheta^{\mathbf{s}_+(x,y)}} :\ x \neq y \text{ lie on a homogeneous unstable curve}\right\} < \infty,$$

where $\mathbf{s}_+(\cdot,\cdot)$ is the forward separation time given by Definition 1. The constant ϑ is called the dynamically Hölder exponent of f, and is usually denoted by ϑ_f. We denote the space of such functions by $\mathcal{H}^+_\vartheta$, and set $\mathcal{H}^+ := \cup_{\vartheta\in(0,1)}\mathcal{H}^+_\vartheta$.

In a similar fashion, we define the space $\mathcal{H}^-_\vartheta$ and $\mathcal{H}^-$ of backward dynamically Hölder continuous functions. Also, we denote $\mathcal{H}_\vartheta := \mathcal{H}^+_\vartheta \cap \mathcal{H}^-_\vartheta$, and $\mathcal{H} := \mathcal{H}^+ \cap \mathcal{H}^-$.

Remark 1. Note that any Hölder continuous function is automatically dynamical Hölder continuous. However, a dynamically Hölder function can be only piecewise continuous, and it may not be bounded.

We need to assume certain integrability for the observables. Given an L^s-integrable function f on M for some $s \geq 1$, we denote $\mathbb{E}(f) = \int f d\mu$ and $\|f\|_{L^s} = \mathbb{E}(|f|^s)^{1/s}$.

For convenience, we shall use the following notations: given two sequence a_n and b_n of non-negative numbers, we write $a_n = o(b_n)$ if $\lim_{n\to\infty} a_n/b_n = 0$; we write $a_n = \mathcal{O}(b_n)$ or $a_n \ll b_n$ if $a_n \leq Cb_n$ for some constant $C > 0$, which is independent of n; and we denote $a_n \asymp b_n$ if $a_n \ll b_n$ and $a_n \gg b_n$.

We are now ready to state our main result.

Theorem 1. *Let $\mathbf{X}_{\boldsymbol{f}} = \{X_n\}_{n\geq 0} := \{f_n \circ T^n\}_{n\geq 0}$ be a random process generated by a sequence $\boldsymbol{f} = \{f_n\}_{n\geq 0}$ of functions, which satisfies the following conditions:*

(1) *There are $\vartheta_{\boldsymbol{f}} \in (0,1)$ and $\beta_{\boldsymbol{f}} \in [0,\infty)$ such that $f_n \in \mathcal{H}_{\vartheta_{\boldsymbol{f}}}$ and*

$$|f_n|^+_{\vartheta_{\boldsymbol{f}}} + |f_n|^-_{\vartheta_{\boldsymbol{f}}} \ll n^{\beta_{\boldsymbol{f}}}.$$

(2) *There is $p > 4$ such that $f_n \in L^p$ with $\mathbb{E}(f_n) = 0$. Moreover, there are constants $\kappa_p \geq \kappa_2 > \frac{1}{4}$ such that*

$$\sigma_n := \left\|\sum_{k=0}^{n-1} X_k\right\|_{L^2} \gg n^{\kappa_2}, \quad \textit{and} \quad \sup_{m\geq 0}\left\|\sum_{k=m}^{m+n-1} X_k\right\|_{L^p} \ll n^{\kappa_p}. \tag{2.1}$$

Then the process $\mathbf{X}_{\boldsymbol{f}}$ *satisfies an ASIP for any error exponent* $\lambda \in \left(\max\{\frac{1}{4}, \frac{1}{8\kappa_2}\}, \frac{1}{2}\right)$, *that is, there exists a Wiener process* $W(\cdot)$ *such that*

$$\left|\sum_{k=0}^{n-1} X_k - W(\sigma_n^2)\right| = \mathcal{O}(\sigma_n^{2\lambda}), \quad a.s.. \tag{2.2}$$

Remark 2. It is well known that a zero-mean independent process $\mathbf{X} = \{X_n\}_{n\geq 0}$ with finite s-th moment (for some $s \geq 1$) satisfies the Marcinkiewicz-Zygmund inequalities, i.e., $\left\|\sum_{k=m}^{m+n-1} X_k\right\|_{L^s} \asymp \left\|\left(\sum_{k=m}^{m+n-1} X_k^2\right)^{\frac{1}{2}}\right\|_{L^s}$. Such type of inequalities were later generalized to martingale difference sequence, strongly mixing processes, etc (see e.g. [38, 56]). We note that the term $\left\|\left(\sum_{k=m}^{m+n-1} X_k^2\right)^{\frac{1}{2}}\right\|_{L^s}$ is of order $\sqrt{n}$ for stationary iid. processes. Due to the dependence and non-stationarity in our setting, there is no a priori information on $\left\|\left(\sum_{k=m}^{m+n-1} X_k^2\right)^{\frac{1}{2}}\right\|_{L^s}$. To this end, in terms of powers of n, we directly impose the 2nd moment lower bound and p-th moment upper bound in (2.1) for the partial sum $\sum_{k=m}^{m+n-1} X_k$.

Condition (1) in Theorem 1 implies that every function f_n is dynamically Hölder continuous with a common exponent $\vartheta_{\mathbf{f}}$, while the dynamically Hölder semi-norms of f_n are allowed to grow in a polynomial rate. Condition (2) implies that the growth rate of partial sum variances σ_n^2 is of order between $n^{2\kappa_2}$ and $n^{2\kappa_p}$. In particular, if $\kappa_2 = \kappa_p = \frac{1}{2}$, then the growth is asymptotically linear, i.e., $\sigma_n^2 \asymp n$.

We notice that the error exponent λ in (2.2) does not depend on the values of $\vartheta_{\mathbf{f}}$, $\beta_{\mathbf{f}}$ and κ_p, and it can be chosen arbitrarily close to $\frac{1}{4}$ if $\kappa_2 = \frac{1}{2}$. In the case when $p \in (2, 4]$ and $\kappa_2 > \frac{1}{p}$, our result still holds with $\lambda \in \left(\max\{\frac{1}{4}, \frac{1}{2p\kappa_2}\}, \frac{1}{2}\right)$, but requires advanced moment inequalities in the proof of a technical lemma - Lemma 14. For simplicity, we just prove the case when $p > 4$, which is sufficient for all of our applications.

Note that the ASIP is the strongest form - it implies many other limit laws, such as the weak invariance principle, the almost sure central limit theorem, and the law of iterated logarithm (see e.g. [44] for their proofs and more details).

Theorem 2. *Let* $\mathbf{X}_{\boldsymbol{f}} = \{X_n\}_{n\geq 0} := \{f_n \circ T^n\}_{n\geq 0}$ *be the random process satisfying the assumptions in Theorem 1. We have the following limit laws:*

(1) *Weak Invariance Principle: for any* $t \in [0, 1]$,

$$\frac{1}{\sigma_n} \sum_{k=0}^{\lfloor nt \rfloor - 1} f_k \circ T^k \xrightarrow{\text{in distribution}} W(t), \quad \text{as } n \to \infty,$$

where $W(\cdot)$ *is a Wiener process.*

(2) *Almost Sure Central Limit Theorem: we denote* $S_n = \sum_{k=0}^{n-1} f_k \circ T^k$, *and let* $\delta(\cdot)$ *be the Dirac measure on* $\mathbb{R}$, *then for* μ-*almost every* $x \in M$,

$$\frac{1}{\log \sigma_n^2} \sum_{k=1}^{n} \frac{1}{\sigma_k^2} \delta_{S_k(x)} \xrightarrow{\text{in distribution}} N(0, 1), \quad \text{as } n \to \infty,$$

where $N(0, 1)$ *is the standard normal distribution.*

(3) *Law of Iterated Logarithm: for μ-almost every $x \in M$,*

$$\limsup_{n\to\infty} \frac{\sum_{k=0}^{n-1} f_k \circ T^k(x)}{\sqrt{2\sigma_n^2 \log\log \sigma_n^2}} = 1.$$

3. Preliminaries from Probability Theory

In this section, we recall several useful theorems in the probability theory. Let (M, μ) be a standard probability space.

Lemma 3 (Borel-Cantelli lemma)**.** *If $\{E_n\}_{n\geq 1}$ is a sequence of events on (M,μ) such that $\sum_{n=1}^{\infty} \mu(E_n) < \infty$, then $\mu\left(\cap_{n=1}^{\infty} \cup_{k\geq n} E_k\right) = 0$.*

We introduce a special case of the results by Gal-Koksma (Theorem A1 in [44]).

Proposition 4. *Let $\{X_n\}_{n\geq 0}$ be a sequence of zero-mean random variables with finite second moments. Suppose there is $\kappa > 0$ such that for any $m \geq 0$, $n \geq 1$,*

$$\mathbb{E}\left(\sum_{k=m}^{m+n-1} X_k\right)^2 \ll (m+n)^\kappa - m^\kappa,$$

then for any $\delta > 0$,

$$\sum_{k=0}^{n-1} X_k = o\left(n^{\frac{\kappa}{2}+\delta}\right), \quad a.s..$$

Let $\mathfrak{F}$ and $\mathfrak{G}$ be two σ-algebras on the space (M, μ).

Definition 3. The α-mixing coefficient between $\mathfrak{F}$ and $\mathfrak{G}$ is given by

$$\alpha(\mathfrak{F}, \mathfrak{G}) := \sup_{A\in\mathfrak{F}} \sup_{B\in\mathfrak{G}} \left|\mu(A\cap B) - \mu(A)\mu(B)\right|. \tag{3.1}$$

Note that $\alpha(\mathfrak{F}, \mathfrak{G}) \leq 2$. We have the following covariance inequality.

Lemma 5 (Lemma 7.2.1 in [44])**.** *Let s_1, s_2 and s_3 be positive real numbers such that $1/s_1 + 1/s_2 + 1/s_3 = 1$. For any $X \in L^{s_1}(M, \mathfrak{F}, \mu)$ and any $Y \in L^{s_2}(M, \mathfrak{G}, \mu)$,*

$$|\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)| \leq 10\alpha(\mathfrak{F}, \mathfrak{G})^{\frac{1}{s_3}} \|X\|_{L^{s_1}} \|Y\|_{L^{s_2}}.$$

Definition 4. $\{(\xi_j, \mathfrak{G}_j)\}_{j\geq 1}$ is called a martingale difference sequence if

(1) $\mathfrak{G}_j$ is an increasing sequence of σ-algebras on (M, μ);
(2) Each ξ_j is L^1-integrable and $\mathfrak{G}_j$-measurable;
(3) $\mathbb{E}(\xi_j | \mathfrak{G}_{j-1}) = 0$ for any $j \geq 2$.

Here is a basic identity for martingale difference sequence $\{(\xi_j, \mathfrak{G}_j)\}_{j\geq 1}$:

$$\mathbb{E}(X\xi_j) = 0, \tag{3.2}$$

for any $\mathfrak{G}_{j-1}$-measurable random variable X, as long as $X\xi_j \in L^1$.

We shall need the following almost sure invariance principle by Shao [50] for the martingale difference sequences (in the L^4-integrable case).

Proposition 6 ([50])**.** *Let $\{(\xi_j, \mathfrak{G}_j)\}_{j\geq 1}$ be an L^4-integrable martingale difference sequence. Put $b_r^2 = \mathbb{E}\left(\sum_{j=1}^r \xi_j\right)^2$. Assume that there exists a sequence $\{a_r\}_{r\geq 1}$ of non-decreasing positive numbers with $\lim_{r\to\infty} a_r = \infty$ such that*

$$\sum_{j=1}^{r} \left[\mathbb{E}\left(\xi_j^2 \,|\mathfrak{G}_{j-1}\right) - \mathbb{E}\xi_j^2\right] = o(a_r), \text{ a.s.} \tag{3.3}$$

$$\sum_{j=1}^{\infty} a_j^{-2}\mathbb{E}|\xi_j|^4 < \infty. \tag{3.4}$$

Then $\{\xi_j\}_{j\geq 1}$ satisfies an ASIP of the following form: there exists a Wiener process $W(\cdot)$ such that

$$\left|\sum_{j=1}^{r} \xi_j - W(b_r^2)\right| = o\left(\left(a_r\left(\left|\log(b_r^2/a_r)\right| + \log\log a_r\right)\right)^{1/2}\right), \text{ a.s.}$$

4. Proof of Theorem 1

We shall prove our main theorem as follows. Firstly, we construct a natural family $\mathfrak{F}$ of σ-algebras on (M, μ), and show that such family is exponentially α-mixing. Secondly, we introduce blocks and approximate the sequence $\mathbf{f}$ by conditional expectation over a special sub-family of $\mathfrak{F}$ on each block. Furthermore, we divide the partial sum of $\mathbf{X_f}$ into a major part $\sum_{j=1}^{r(n)-1} Y_j$ and other negligible parts. Thirdly, we establish the martingale difference representation $\{\xi_j\}_{j\geq 1}$ for the process $\{Y_j\}_{j\geq 1}$, and obtain several preliminary norm estimates. Fourthly, we prove a technical lemma on Condition (3.3) and show an ASIP for the martingale difference sequence $\{\xi_j\}_{j\geq 1}$. Finally, we prove the ASIP for the original sequence $\mathbf{f}$.

4.1. **The strong mixing property.** We first recall the exponential decay of correlations for the system (M, T, μ) for bounded dynamically Hölder observables, which was proven in [21] by using the coupling lemma (see e.g. [16, 14]).

Proposition 7 ([21])**.** *There exist $C_0 > 0$ and $\vartheta_0 \in (0, 1)$ such that for any pair of functions $f \in \mathcal{H}^+_{\vartheta_f} \cap L^\infty(\mu)$ and $g \in \mathcal{H}^-_{\vartheta_g} \cap L^\infty(\mu)$ and $n \geq 0$,*

$$|\mathbb{E}(f \cdot g \circ T^n) - \mathbb{E}(f)\mathbb{E}(g)| \leq C_{f,g}\theta_{f,g}^n,$$

where $\theta_{f,g} = \max\{\vartheta_0, \vartheta_f^{1/4}, \vartheta_g^{1/4}\} < 1$, and

$$C_{f,g} = C_0\left(\|f\|_{L^\infty}\|g\|_{L^\infty} + \|f\|_{L^\infty}|g|^-_{\vartheta_g} + \|g\|_{L^\infty}|f|^+_{\vartheta_f}\right).$$

We then introduce the following natural family of σ-algebras for the system $T : M \to M$. Recall that $S_{\pm n}$ is the singularity set of $T^{\pm n}$ for $n \geq 1$. Let $\xi_0 := \{M\}$ be the trivial partition of M, and denote by $\xi_{\pm n}$ the partition of M into connected components of $M \backslash T^{\mp(n-1)} S_{\pm 1}$ for $n \geq 1$. Further, let

$$\xi_m^n := \xi_m \vee \cdots \vee \xi_n$$

for all $-\infty \leq m \leq n \leq \infty$. By Assumption (**H2**), ξ_0^∞ is the partition of M into maximal unstable manifolds, and $\xi_{-\infty}^0$ is that into maximal stable manifolds. Also, $\mu(\partial \xi_m^n) = 0$ by Assumption (**H4**), where $\partial \xi_m^n$ is the set of boundary curves for components in ξ_m^n.

Let $\mathfrak{F}_m^n$ be the Borel σ-algebra generated by the partition ξ_m^n. Notice that $\mathfrak{F}_{-\infty}^{\infty}$ coincides with the σ-algebra of all measurable subsets in M. We denote by $\mathfrak{F} := \{\mathfrak{F}_m^n\}_{-\infty\le m\le n\le\infty}$ the family of those σ-algebras.

Proposition 8. *The family $\mathfrak{F}$ is α-mixing with an exponential rate, i.e., there exist $C_0 > 0$ and $\vartheta_0 \in (0,1)$ (which are the same as in Proposition 7) such that*

$$\sup_{k\in\mathbb{Z}} \alpha(\mathfrak{F}_{-\infty}^k, \mathfrak{F}_{k+n}^{\infty}) \le C_0\vartheta_0^n,$$

where the definition of $\alpha(\cdot,\cdot)$ is given by (3.1).

Proof. By the fact that $T^{-k}\xi_m^n = \xi_{m+k}^{n+k}$ and the invariance of μ, it suffices to show that

$$\alpha(\mathfrak{F}_{-\infty}^0, \mathfrak{F}_n^{\infty}) = \sup_{A\in\mathfrak{F}_{-\infty}^0}\ \sup_{B\in\mathfrak{F}_n^{\infty}} |\mu(A\cap B) - \mu(A)\mu(B)| \le C_0\vartheta_0^n.$$

Since $A \in \mathfrak{F}_{-\infty}^0$ is a union of some maximal stable manifolds, we have that $\mathbf{1}_A \in \mathcal{H}^-$ such that $\|\mathbf{1}_A\|_{L^\infty} = 1$ and $|\mathbf{1}_A|_{\vartheta}^- = 0$ for any $\vartheta \in (0,1)$. Similarly, $B \in \mathfrak{F}_n^{\infty}$ implies that $T^{-n}(B) \in \mathfrak{F}_0^{\infty}$ is a union of some maximal unstable manifolds, and thus $\mathbf{1}_{T^{-n}B} \in \mathcal{H}^+$ such that $\|\mathbf{1}_{T^{-n}B}\|_{L^\infty} = 1$ and $|\mathbf{1}_{T^{-n}B}|_{\vartheta}^+ = 0$ for any $\vartheta \in (0,1)$. Therefore, by Proposition 7, for any $A \in \mathfrak{F}_{-\infty}^0$ and $B \in \mathfrak{F}_n^{\infty}$,

$$|\mu(A\cap B) - \mu(A)\mu(B)| = |\mathbb{E}(\mathbf{1}_{T^{-n}B}\cdot \mathbf{1}_A\circ T^n) - \mathbb{E}(\mathbf{1}_{T^{-n}B})\mathbb{E}(\mathbf{1}_A)| \le C_0\vartheta_0^n.$$

This completes the proof of Proposition 8. □

4.2. **Blocks and approximations.** Let $\mathbf{f} = \{f_n\}_{n\ge0}$ be a sequence of functions satisfying the assumptions of Theorem 1.

From now on, we fix a error exponent $\lambda \in \left(\max\{\frac14, \frac{1}{8\kappa_2}\}, \frac12\right)$, and choose a sufficiently small constant $\epsilon > 0$ such that

$$2\epsilon\kappa_p + \frac{1}{4-\epsilon} < 2\kappa_2\lambda, \quad\text{and}\quad \frac{\epsilon\kappa_p}{\kappa_2} < 4\lambda - 1. \tag{4.1}$$

We partition the time interval $[0,\infty)$ into a sequence of consecutive blocks $\Delta_j = [\tau_j, \tau_{j+1})$ for $j \ge 1$, where $\tau_j = \sum_{i=0}^{j-1}\lceil i^\epsilon\rceil$. Note that the block Δ_j is of length $\lceil j^\epsilon\rceil$, and $\tau_j \asymp j^{1+\epsilon}$. For convenience, we set $\tau_0 = -1$.

For any $k \in \Delta_j$, we define the approximated function of f_k by

$$g_k = \mathbb{E}\left(f_k \middle| \mathfrak{F}_{-\lceil 0.2j^\epsilon\rceil}^{\lceil 0.2j^\epsilon\rceil}\right). \tag{4.2}$$

It is clear that $\mathbb{E}(g_k) = 0$. Since the separation times are adapted to the natural family $\mathfrak{F}$ of σ-algebras, we have the following uniform L^∞-bounds on the difference sequence $\{(f_k - g_k)\}_{k\ge0}$.

Lemma 9. $\sup_{k\in\Delta_j}\|f_k - g_k\|_{L^\infty} \ll \vartheta_{\boldsymbol{f}}^{0.1j^\epsilon}$.

Proof. Note that $k \le \tau_{j+1} \asymp j^{1+\epsilon}$ for any $k \in \Delta_j$. For any measurable set $A \in \xi_{-\lceil 0.2j^\epsilon\rceil}^{\lceil 0.2j^\epsilon\rceil}$, and any two points $x, y \in A$, there is a point $z \in A$ such that x and z belong to one unstable curve, and y and z belong to one stable curve. It follows that $\mathbf{s}_+(x,z) > \lceil 0.2j^\epsilon\rceil$ and $\mathbf{s}_-(y,z) > \lceil 0.2j^\epsilon\rceil$, and thus by Condition (1) of Theorem 1,

$$\begin{aligned}|f_k(x) - f_k(y)| &\le |f_k(x) - f_k(z)| + |f_k(y) - f_k(z)| \\ &\le \left(|f_k|_{\vartheta_{\mathbf{f}}}^+ + |f_k|_{\vartheta_{\mathbf{f}}}^-\right)\vartheta_{\mathbf{f}}^{\lceil 0.2j^\epsilon\rceil}\end{aligned}$$

$$\ll \quad k^{\beta_{\mathbf{f}}}\vartheta_{\mathbf{f}}^{\lceil 0.2j^\epsilon\rceil} \le j^{(1+\epsilon)\beta_{\mathbf{f}}}\vartheta_{\mathbf{f}}^{\lceil 0.2j^\epsilon\rceil} \ll \vartheta_{\mathbf{f}}^{0.1j^\epsilon}.$$

Hence for any $k \in \Delta_j$, we have

$$\begin{aligned} |f_k(x) - g_k(x)| &= \left| f_k(x) - \frac{1}{\mu(A)}\int_A f_k(y)d\mu(y)\right| \\ &\le \frac{1}{\mu(A)}\int_A |f_k(x) - f_k(y)|\, d\mu(y) \ll \vartheta_{\mathbf{f}}^{0.1j^\epsilon}. \end{aligned}$$

The proof of Lemma 9 is complete. □

For any $n \ge 0$, there is a unique $r(n) \ge 1$ such that $n \in \Delta_{r(n)}$. Note that $r(n) \asymp n^{\frac{1}{1+\epsilon}}$. We now decompose the partial sum of the process $\mathbf{X_f} = \{X_n\}_{n\ge 0} = \{f_n \circ T^n\}_{n\ge 0}$ as follows:

$$\begin{aligned} \sum_{k=0}^{n-1} X_k &= \sum_{j=1}^{r(n)-1}\left(\sum_{k\in\Delta_j} g_k\circ T^k\right) + \sum_{k=0}^{\tau_{r(n)}-1}(f_k - g_k)\circ T^k + \sum_{k=\tau_{r(n)}}^{n-1} X_k \\ &=: \sum_{j=1}^{r(n)-1} Y_j + U_n + V_n. \end{aligned} \tag{4.3}$$

It turns out that the major contribution for ASIP is given by $\sum_{j=1}^{r(n)-1} Y_j$, while the rest terms are negligible.

Lemma 10. *Let U_n and V_n be given by* (4.3). *Then*

(i) $\|U_n\|_{L^p} = \mathcal{O}(1)$, *and* $|U_n| = \mathcal{O}(1)$, *a.s.*.
(ii) $\|V_n\|_{L^p} = \mathcal{O}(n^{\epsilon\kappa_p})$, *and* $|V_n| = \mathcal{O}\left(n^{2\kappa_2\lambda}\right)$, *a.s.*.

Proof. (i) Note that $U_n = \sum_{j=1}^{r(n)-1}\sum_{k\in\Delta_j}(f_k - g_k)\circ T^k$. By Lemma 9 and Minkowski's inequality, we have

$$\begin{aligned} \|U_n\|_{L^p} \le \sum_{j=1}^{\infty}\sum_{k\in\Delta_j}\|(f_k-g_k)\circ T^k\|_{L^p} &\le \sum_{j=1}^{\infty}\sum_{k\in\Delta_j}\|f_k - g_k\|_{L^p} \\ &\ll \sum_{j=1}^{\infty}\lceil j^\epsilon\rceil\, \vartheta_{\mathbf{f}}^{0.1j^\epsilon} < \infty. \end{aligned}$$

Moreover, $\sum_{j=1}^\infty\sum_{k\in\Delta_j}\|(f_k - g_k)\circ T^k\|_{L^p} < \infty$ implies that $\sum_{j=1}^\infty\sum_{k\in\Delta_j}|(f_k - g_k)\circ T^k| < \infty$ a.s., and thus $|U_n| = \mathcal{O}(1)$ a.s..

(ii) By (2.1), we obtain

$$\|V_n\|_{L^p} = \left\|\sum_{k=\tau_{r(n)}}^{n-1} X_k\right\|_{L^p} \ll \left(n - \tau_{r(n)}\right)^{\kappa_p} \ll (r(n)^\epsilon)^{\kappa_p} \ll n^{\epsilon\kappa_p}.$$

Moreover, by Markov's inequality and (4.1),

$$\mu\{|V_n| \ge n^{2\kappa_2\lambda}\} \le n^{-2p\kappa_2\lambda}\ \mathbb{E}|V_n|^p \ll n^{p(-2\kappa_2\lambda+\epsilon\kappa_p)} \ll n^{-\frac{p}{4-\epsilon}},$$

and hence $\sum_{n=1}^\infty \mu\{|V_n| \ge n^{2\kappa_2\lambda}\} < \infty$. By the Borel-Cantelli lemma (Lemma 3), we get $\mu\left(\bigcap_{k=1}^\infty\bigcup_{n\ge k}\left\{|V_n| \ge n^{2\kappa_2\lambda}\right\}\right) = 0$. In other words, $|V_n| \ll n^{2\kappa_2\lambda}$, a.s.. □

4.3. **Martingale representation for $\{Y_j\}_{j\ge 1}$.** In this subsection, we introduce a martingale representation for the random process $\{Y_j\}_{j\ge 1}$ as defined in (4.3). Such representation is given by Lemma 7.4.1 in [44], but has better norm estimates in our context.

We first establish the following preliminary estimates for Y_j.

Lemma 11. *For any $j \ge 1$, the random variable Y_j is $\mathfrak{F}_{\tau_{j-1}}^{\tau_{j+2}}$-measurable such that $\mathbb{E}Y_j = 0$ and $\|Y_j\|_{L^p} \ll j^{\epsilon\kappa_p}$. Furthermore, $\|\sum_{j=1}^r Y_j\|_{L^2} \gg r^{\kappa_2}$.*

Proof. By (4.2) and (4.3), we have for any $j \ge 1$,

$$Y_j = \sum_{k\in\Delta_j} g_k \circ T^k = \sum_{k\in\Delta_j} \mathbb{E}\left(f_k \left| \mathfrak{F}_{-\lceil 0.2j^\epsilon\rceil}^{\lceil 0.2j^\epsilon\rceil}\right.\right) \circ T^k = \sum_{k\in\Delta_j} \mathbb{E}\left(f_k \circ T^k \left| \mathfrak{F}_{k-\lceil 0.2j^\epsilon\rceil}^{k+\lceil 0.2j^\epsilon\rceil}\right.\right)$$

is $\mathfrak{F}_{\tau_j-\lceil 0.2j^\epsilon\rceil}^{\tau_{j+1}+\lceil 0.2j^\epsilon\rceil}$-measurable, and thus $\mathfrak{F}_{\tau_{j-1}}^{\tau_{j+2}}$-measurable. It is clear that $\mathbb{E}Y_j = 0$ since each f_k is of zero mean. Moreover, by Lemma 9,

$$\left\| Y_j - \sum_{k\in\Delta_j} X_k \right\|_{L^\infty} \le \sum_{k\in\Delta_j} \|f_k - g_k\|_{L^\infty} \ll \lceil j^\epsilon \rceil \vartheta_{\mathbf{f}}^{0.1j^\epsilon} \le \sup_{j\ge 1} \lceil j^\epsilon \rceil \vartheta_{\mathbf{f}}^{0.1j^\epsilon} < \infty.$$

Therefore, by (2.1),

$$\|Y_j\|_{L^p} \le \left\| \sum_{k\in\Delta_j} X_k \right\|_{L^p} + \left\| Y_j - \sum_{k\in\Delta_j} X_k \right\|_{L^\infty} \ll \lceil j^\epsilon \rceil^{\kappa_p} + \mathcal{O}(1) \ll j^{\epsilon\kappa_p}.$$

Furthermore,

$$\begin{aligned} \left\| \sum_{j=1}^r Y_j \right\|_{L^2} &\ge \left\| \sum_{k=0}^{\tau_{r+1}-1} X_k \right\|_{L^2} - \sum_{j=1}^r \left\| Y_j - \sum_{k\in\Delta_j} X_k \right\|_{L^\infty} \\ &\gg \tau_{r+1}^{\kappa_2} - \sum_{j=1}^\infty \lceil j^\epsilon \rceil \vartheta_{\mathbf{f}}^{0.1j^\epsilon} \\ &\gg (r+1)^{\kappa_2(1+\epsilon)} - \mathcal{O}(1) \gg r^{\kappa_2}. \end{aligned}$$

□

Now we denote $\mathfrak{G}_j$ the σ-algebra generated by $Y_1, Y_2, \dots, Y_j$, and it is immediate from Lemma 11 that $\mathfrak{G}_j \subset \mathfrak{F}_{-1}^{\tau_{j+2}}$. We also set $\mathfrak{G}_0 := \{\emptyset, M\}$ to be the trivial σ-algebra.

Lemma 12. *For any $j \ge 1$, we set $\xi_j := Y_j - u_j + u_{j+1}$, where $u_j \in L^4$ is given by*

$$u_j := \sum_{k=0}^\infty \mathbb{E}\left(Y_{j+k} \left| \mathfrak{G}_{j-1}\right.\right). \tag{4.4}$$

Then $\{(\xi_j, \mathfrak{G}_j)\}_{j\ge 1}$ is a martingale difference sequence. Moreover, $\mathbb{E}u_j = \mathbb{E}\xi_j = 0$ and

$$\|u_j\|_{L^4} \ll j^{\epsilon\kappa_p}, \quad \textit{and} \quad \|\xi_j\|_{L^4} \ll j^{\epsilon\kappa_p}.$$

Proof. We first show that each u_j, given by (4.4), is well-defined as an L^4 function. Denote for short $v_{jk} := \mathbb{E}\left(Y_{j+k} \left| \mathfrak{G}_{j-1}\right.\right)$, which is $\mathfrak{G}_{j-1}$-measurable. Then

$$\mathbb{E}|v_{jk}|^4 = \mathbb{E}\left(v_{jk} \cdot v_{jk}^3\right) = \mathbb{E}\left(\mathbb{E}\left(Y_{j+k} \left| \mathfrak{G}_{j-1}\right.\right) \cdot v_{jk}^3\right) = \mathbb{E}\left(Y_{j+k} \cdot v_{jk}^3\right). \tag{4.5}$$

By Lemma 11, Y_{j+k} is $\mathfrak{F}_{\tau_{j+k-1}}^{\tau_{j+k+2}}$-measurable, and also $\mathfrak{G}_{j-1} \subset \mathfrak{F}_{-1}^{\tau_{j+1}}$. We choose $s_1 = p$, $s_2 = \frac{4}{3}$ and $s_3 = \frac{4p}{p-4}$, and apply Lemma 5 to the last term of (4.5),

$$\begin{aligned} \mathbb{E}|v_{jk}|^4 &\leq 10\alpha(\mathfrak{F}_{\tau_{j+k-1}}^{\tau_{j+k+2}}, \mathfrak{F}_{-1}^{\tau_{j+1}})^{\frac{1}{s_3}} \|Y_{j+k}\|_{L^{s_1}} \left\|v_{jk}^3\right\|_{L^{s_2}} \\ &= 10\alpha(\mathfrak{F}_{\tau_{j+k-1}}^{\tau_{j+k+2}}, \mathfrak{F}_{-1}^{\tau_{j+1}})^{\frac{1}{s_3}} \|Y_{j+k}\|_{L^p} \left[\mathbb{E}|v_{jk}|^4\right]^{\frac{3}{4}}. \end{aligned}$$

Dividing $\left[\mathbb{E}|v_{jk}|^4\right]^{\frac{3}{4}}$ on both sides, and then using Proposition 8 and Lemma 11, we have that for any $j \geq 1$,

$$\begin{aligned} \|v_{jk}\|_{L^4} &\leq 10\alpha(\mathfrak{F}_{\tau_{j+k-1}}^{\infty}, \mathfrak{F}_{-\infty}^{\tau_{j+1}})^{\frac{1}{s_3}} \|Y_{j+k}\|_{L^p} \\ &\ll \begin{cases} (j+k)^{\epsilon\kappa_p}, & 0 \leq k < 3, \\ (j+k)^{\epsilon\kappa_p}\, \vartheta_0^{\frac{\lceil (j+k-2)^\epsilon \rceil}{s_3}}, & k \geq 3, \end{cases} \\ &\ll \begin{cases} j^{\epsilon\kappa_p}, & 0 \leq k < 3, \\ \vartheta_0^{\frac{(k-2)^\epsilon}{2s_3}}, & k \geq 3. \end{cases} \end{aligned}$$

Therefore, for any $j \geq 1$,

$$\sum_{k=0}^{\infty} \|v_{jk}\|_{L^4} = \sum_{k=0}^{2} \|v_{jk}\|_{L^4} + \sum_{k=3}^{\infty} \|v_{jk}\|_{L^4} \ll 3j^{\epsilon\kappa_p} + \sum_{k=3}^{\infty} \vartheta_0^{\frac{(k-2)^\epsilon}{2s_3}} \ll j^{\epsilon\kappa_p},$$

which implies that $u_j = \sum_{k=0}^{\infty} v_{jk}$ is well-defined in L^4, and $\|u_j\|_{L^4} \ll j^{\epsilon\kappa_p}$.

By the formula $\xi_j := Y_j - u_j + u_{j+1}$, it is easy to check that $\{(\xi_j, \mathfrak{G}_j)\}_{j\geq 1}$ is a martingale difference sequence (see Definition 4). Moreover,

$$\|\xi_j\|_{L^4} \leq \|Y_j\|_{L^p} + \|u_j\|_{L^4} + \|u_{j+1}\|_{L^4} \ll j^{\epsilon\kappa_p}.$$

The proof of Lemma 12 is complete. □

The following lemma shows that $\sum_{j=1}^r Y_j$ is well approximated by $\sum_{j=1}^r \xi_j$.

Lemma 13. *We have the following estimates:*

$$\left\|\sum_{j=1}^{r} (Y_j - \xi_j)\right\|_{L^4} = \mathcal{O}\left(r^{\epsilon\kappa_p}\right), \quad \text{and} \quad \left|\sum_{j=1}^{r} (Y_j - \xi_j)\right| = \mathcal{O}\left(r^{2\kappa_2\lambda}\right), \ \text{a.s.}.$$

Proof. By Lemma 12, we have $\sum_{j=1}^r (Y_j - \xi_j) = u_1 - u_{r+1}$, and thus

$$\left\|\sum_{j=1}^{r} (Y_j - \xi_j)\right\|_{L^4} = \|u_1 - u_{r+1}\|_{L^4} \ll 1 + (r+1)^{\epsilon\kappa_p} \ll r^{\epsilon\kappa_p}.$$

Moreover, by Markov's inequality and (4.1),

$$\begin{aligned} \mu\left\{\left|\sum_{j=1}^{r} (Y_j - \xi_j)\right| \geq r^{2\kappa_2\lambda}\right\} \leq r^{-8\kappa_2\lambda}\, \mathbb{E}\left|\sum_{j=1}^{r} (Y_j - \xi_j)\right|^4 &\ll r^{4(-2\kappa_2\lambda + \epsilon\kappa_p)} \\ &\ll r^{-\frac{4}{4-\epsilon}}, \end{aligned}$$

and hence $\sum_{r=1}^{\infty} \mu\left\{\left|\sum_{j=1}^r (Y_j - \xi_j)\right| \geq r^{2\kappa_2\lambda}\right\} < \infty$. By the Borel-Cantelli lemma (Lemma 3), we get $\left|\sum_{j=1}^r (Y_j - \xi_j)\right| \ll r^{2\kappa_2\lambda}$, a.s.. □

According to Lemma 10 and Lemma 13, we shall focus on proving ASIP for the process $\{\xi_j\}_{j\geq 1}$.

4.4. **ASIP for $\{\xi_j\}_{j\geq 1}$.** In this subsection, we shall use Proposition 6 to prove a version of ASIP for the martingale difference sequence $\{(\xi_j, \mathfrak{G}_j)\}_{j\geq 1}$. We first need a technical lemma with the following almost sure estimate.

Lemma 14. $\sum_{j=1}^{r}\left[\mathbb{E}\left(\xi_j^2\,|\mathfrak{G}_{j-1}\right) - \mathbb{E}\xi_j^2\right] = o(r^{4\kappa_2\lambda})$, *a.s.*

Proof. We denote $R_j := \mathbb{E}\left(\xi_j^2\,|\mathfrak{G}_{j-1}\right) - \mathbb{E}\xi_j^2 = \mathbb{E}\left(\xi_j^2 - \mathbb{E}\xi_j^2\,|\mathfrak{G}_{j-1}\right)$, and note that R_j is $\mathfrak{G}_{j-1}$-measurable and $\mathbb{E}R_j = 0$. Moreover, by Lemma 12 and Jensen's inequality,

$$\|R_j\|_{L^2} \leq \sqrt{\mathbb{E}\left(\xi_j^2 - \mathbb{E}\xi_j^2\right)^2} \leq \|\xi_j\|_{L^4}^2 \ll j^{2\epsilon\kappa_p}.$$

If for any $m \geq 1$ and any $r \geq 1$,

$$\mathbb{E}\left(\sum_{j=m}^{m+r-1} R_j\right)^2 \ll (m+r)^{1+8\epsilon\kappa_p} - m^{1+8\epsilon\kappa_p}, \tag{4.6}$$

then Lemma 14 immediately follows from (4.1) and Gal-Koksma (Proposition 4). In the rest of the proof, we shall prove (4.6). Using that $\mathbb{E}R_j = 0$ and $\mathbb{E}R_j^2 \geq 0$, we first notice that

$$\begin{aligned}
\mathbb{E}\left(\sum_{j=m}^{m+r-1} R_j\right)^2 &\leq 2\sum_{j=m}^{m+r-1}\sum_{k=0}^{m+r-1-j}\mathbb{E}(R_jR_{j+k})\\
&= 2\sum_{j=m}^{m+r-1}\sum_{k=0}^{m+r-1-j}\mathbb{E}\left(R_j\mathbb{E}\left(\xi_{j+k}^2 - \mathbb{E}\xi_{j+k}^2\,|\mathfrak{G}_{j+k-1}\right)\right)\\
&= 2\sum_{j=m}^{m+r-1}\sum_{k=0}^{m+r-1-j}\mathbb{E}\left(R_j\left(\xi_{j+k}^2 - \mathbb{E}\xi_{j+k}^2\right)\right)\\
&= 2\sum_{j=m}^{m+r-1}\mathbb{E}\left(R_j\sum_{k=0}^{m+r-1-j}\xi_{j+k}^2\right)\\
&= 2\sum_{j=m}^{m+r-1}\mathbb{E}\left(R_j\left(\sum_{k=0}^{m+r-1-j}\xi_{j+k}\right)^2\right)\\
&\quad -4\sum_{j=m}^{m+r-1}\sum\sum_{0\leq k<\ell\leq m+r-1-j}\mathbb{E}\left(R_j\xi_{j+k}\xi_{j+\ell}\right)\\
&= 2\sum_{j=m}^{m+r-1}\mathbb{E}\left(R_j\left(\sum_{k=0}^{m+r-1-j}\xi_{j+k}\right)^2\right).
\end{aligned}$$

In the last step, we use (3.2) to conclude that $\mathbb{E}\left(R_j\xi_{j+k}\xi_{j+\ell}\right) = 0$ if $k < \ell$. By Lemma 12, we further obtain

$$\mathbb{E}\left(\sum_{j=m}^{m+r-1} R_j\right)^2$$

$$
\begin{aligned}
&\leq\ 2\sum_{j=m}^{m+r-1}\mathbb{E}\left(R_j\ \left[\sum_{k=0}^{m+r-1-j}Y_{j+k}+(u_{m+r-1}-u_j)\right]^2\right)\\
&\leq\ 2\sum_{j=m}^{m+r-1}\left\{\sum_{k=0}^{m+r-1-j}\mathbb{E}\left(R_jY_{j+k}^2\right)+2\sum_{k=0}^{m+r-1-j}\sum_{\ell=1}^{m+r-1-j-k}\mathbb{E}\left(R_jY_{j+k}Y_{j+k+\ell}\right)\right.\\
&\qquad\qquad \left.+2\sum_{k=0}^{m+r-1-j}\mathbb{E}\left(R_jY_{j+k}u_{m+r-1}\right)-2\sum_{k=0}^{m+r-1-j}\mathbb{E}\left(R_jY_{j+k}u_j\right)\right\}\\
&\quad +2\sum_{j=m}^{m+r-1}\mathbb{E}\left(R_j\left(u_{m+r-1}-u_j\right)^2\right)\\
&=:\ 2\sum_{j=m}^{m+r-1}(I_1+I_2+I_3+I_4)+2I_5.
\end{aligned}
$$

To prove (4.6), it suffices to show that

$$|I_i|\ll j^{8\epsilon\kappa_p},\ \text{for } i=1,2,3,4,\quad\text{and}\quad |I_5|\ll(m+r)^{1+8\epsilon\kappa_p}-m^{1+8\epsilon\kappa_p}.$$

For I_1: Recall that $\|R_j\|_{L^2}\ll j^{2\epsilon\kappa_p}$, and R_j is $\mathfrak{G}_{j-1}$- and thus $\mathfrak{F}_{-1}^{\tau_{j+1}}$-measurable. By Lemma 11, $\|Y_{j+k}^2\|_{L^{p/2}}\leq\|Y_{j+k}\|_{L^p}^2\ll(j+k)^{2\epsilon\kappa_p}$, and Y_{j+k}^2 is $\mathfrak{F}_{\tau_{j+k-1}}^{\tau_{j+k+2}}$- and thus $\mathfrak{F}_{\tau_{j+k-1}}^{\infty}$-measurable. Applying Lemma 5 and Proposition 8, we take $s=\frac{2p}{p-4}$ and get

$$
\begin{aligned}
|I_1|\ &\leq\ \sum_{k=0}^{m+r-1-j}\left|\mathbb{E}\left(R_jY_{j+k}^2\right)\right|\\
&\leq\ \sum_{k=0}^{m+r-1-j}10\alpha(\mathfrak{F}_{-1}^{\tau_{j+1}},\mathfrak{F}_{\tau_{j+k-1}}^{\infty})^{\frac{1}{s}}\,\|R_j\|_{L^2}\left\|Y_{j+k}^2\right\|_{L^{p/2}}\\
&\ll\ \sum_{k=0}^{2}j^{2\epsilon\kappa_p}(j+k)^{2\epsilon\kappa_p}+\sum_{k=3}^{m+r-1-j}\vartheta_0^{\frac{\lceil(j+k-2)^\epsilon\rceil}{s}}j^{2\epsilon\kappa_p}(j+k)^{2\epsilon\kappa_p}\\
&\ll\ j^{4\epsilon\kappa_p}\left[\mathcal{O}(1)+\sum_{k=3}^{m+r-1-j}\vartheta_0^{\frac{\lceil(k-2)^\epsilon\rceil}{s}}(1+k)^{2\epsilon\kappa_p}\right]\ll j^{8\epsilon\kappa_p}.
\end{aligned}
$$

For I_2: we split the double sum into the cases $k\geq\ell$ and $k<\ell$, that is,

$$|I_2|\leq 2\sum_{k=0}^{m+r-1-j}\sum_{1\leq\ell\leq k}\left|\mathbb{E}\left(R_j(Y_{j+k}Y_{j+k+\ell})\right)\right|+2\sum_{\ell=1}^{m+r-1-j}\sum_{0\leq k<\ell}\left|\mathbb{E}\left((R_jY_{j+k})Y_{j+k+\ell}\right)\right|.$$

In the first summation, we note that $\ell\leq k$ and

$$\|Y_{j+k}Y_{j+k+\ell}\|_{L^{p/2}}\leq\|Y_{j+k}\|_{L^p}\|Y_{j+k+\ell}\|_{L^p}\ll(j+k)^{\epsilon\kappa_p}(j+k+\ell)^{\epsilon\kappa_p}\ll(j+2k)^{2\epsilon\kappa_p}.$$

Applying Lemma 5 and Proposition 8, we take $s=\frac{2p}{p-4}$ and get

$$\sum_{k=0}^{m+r-1-j}\sum_{1\leq\ell\leq k}\left|\mathbb{E}\left(R_j(Y_{j+k}Y_{j+k+\ell})\right)\right|$$

$$\begin{aligned}
&\leq \sum_{k=0}^{m+r-1-j} \sum_{1\leq \ell\leq k} 10\alpha(\mathfrak{F}_{-1}^{\tau_{j+1}}, \mathfrak{F}_{\tau_{j+k-1}}^{\infty})^{\frac{1}{s}} \|R_j\|_{L^2} \|Y_{j+k}Y_{j+k+\ell}\|_{L^{p/2}} \\
&\ll \sum_{k=0}^{2} kj^{2\epsilon\kappa_p}(j+2k)^{2\epsilon\kappa_p} + \sum_{k=3}^{m+r-1-j} k\vartheta_0^{\frac{\lceil (j+k-2)^\epsilon\rceil}{s}} j^{2\epsilon\kappa_p}(j+2k)^{2\epsilon\kappa_p} \\
&\ll j^{4\epsilon\kappa_p}\left[\mathcal{O}(1) + \sum_{k=3}^{m+r-1-j} \vartheta_0^{\frac{\lceil (k-2)^\epsilon\rceil}{s}}(1+2k)^{1+2\epsilon\kappa_p}\right] \ll j^{8\epsilon\kappa_p}.
\end{aligned}$$

In the second summation, we note that $k < \ell$ and hence

$$\|R_jY_{j+k}\|_{L^{4/3}} \leq \|R_j\|_{L^2}^{4/3}\|Y_{j+k}\|_{L^4}^{4/3} \ll \left[j^{2\epsilon\kappa_p}(j+k)^{\epsilon\kappa_p}\right]^{4/3} \ll (j+k)^{4\epsilon\kappa_p} \leq (j+\ell)^{4\epsilon\kappa_p}.$$

Also, $\|Y_{j+k+\ell}\|_{L^p} \ll (j+k+\ell)^{\epsilon\kappa_p} \leq (j+2\ell)^{\epsilon\kappa_p}$. Applying Lemma 5 and Proposition 8, we take $s' = \frac{4p}{p-4}$ and get

$$\begin{aligned}
&\sum_{\ell=1}^{m+r-1-j} \sum_{0\leq k<\ell} |\mathbb{E}\left((R_jY_{j+k})Y_{j+k+\ell}\right)| \\
\leq &\sum_{\ell=1}^{m+r-1-j} \sum_{0\leq k<\ell} 10\alpha(\mathfrak{F}_{-1}^{\tau_{j+k+1}}, \mathfrak{F}_{\tau_{j+k+\ell-1}}^{\infty})^{\frac{1}{s'}} \|R_jY_{j+k}\|_{L^{4/3}} \|Y_{j+k+\ell}\|_{L^p} \\
\ll &\sum_{\ell=1}^{2} \ell(j+\ell)^{4\epsilon\kappa_p}(j+2\ell)^{\epsilon\kappa_p} + \sum_{\ell=3}^{m+r-1-j} \ell\vartheta_0^{\frac{\lceil (j+k+\ell-2)^\epsilon\rceil}{s'}}(j+\ell)^{4\epsilon\kappa_p}(j+2\ell)^{\epsilon\kappa_p} \\
\ll &j^{5\epsilon\kappa_p}\left[\mathcal{O}(1) + \sum_{\ell=3}^{m+r-1-j} \vartheta_0^{\frac{\lceil (\ell-2)^\epsilon\rceil}{s}}(1+2\ell)^{1+5\epsilon\kappa_p}\right] \ll j^{8\epsilon\kappa_p}.
\end{aligned}$$

Therefore, $|I_2| \ll j^{8\epsilon\kappa_p}$.

For I_3: by the definition of u_j in (4.4), we rewrite

$$\begin{aligned}
\sum_{k=0}^{m+r-1-j} \mathbb{E}\left(R_jY_{j+k}u_{m+r-1}\right) &= \sum_{k=0}^{m+r-1-j} \sum_{\ell=0}^{\infty} \mathbb{E}\left(R_jY_{j+k}\mathbb{E}\left(Y_{m+r-1+\ell}|\mathfrak{G}_{m+r-2}\right)\right) \\
&= \sum_{k=0}^{m+r-1-j} \sum_{\ell=0}^{\infty} \mathbb{E}\left(R_jY_{j+k}Y_{m+r-1+\ell}\right) \\
&= \sum_{k=0}^{m+r-1-j} \sum_{\ell=m+r-1-j-k}^{\infty} \mathbb{E}\left(R_jY_{j+k}Y_{j+k+\ell}\right)
\end{aligned}$$

We can split I_3 into the cases $k \geq \ell$ and $k < \ell$, and obtain $|I_3| \ll j^{8\epsilon\kappa_p}$ by applying similar estimates for I_2.

For I_4: Note that $\|R_ju_j\|_{L^{4/3}} \leq \|R_j\|_{L^2}^{4/3}\|u_j\|_{L^4}^{4/3} \ll \left[j^{2\epsilon\kappa_p}j^{\epsilon\kappa_p}\right]^{4/3} = j^{4\epsilon\kappa_p}$. Applying Lemma 5 and Proposition 8, we take $s' = \frac{4p}{p-4}$ and get

$$\begin{aligned}
|I_4| &\leq 2\sum_{k=0}^{m+r-1-j} |\mathbb{E}\left(R_jY_{j+k}u_j\right)| \\
&\leq 2\sum_{k=0}^{m+r-1-j} 10\alpha(\mathfrak{F}_{-1}^{\tau_{j+1}}, \mathfrak{F}_{\tau_{j+k-1}}^{\infty})^{\frac{1}{s'}} \|R_ju_j\|_{L^{4/3}} \|Y_{j+k}\|_{L^p}
\end{aligned}$$

$$
\begin{aligned}
&\ll \sum_{k=0}^{2} j^{4\epsilon\kappa_p}(j+k)^{\epsilon\kappa_p} + \sum_{k=3}^{m+r-1-j} \vartheta_0^{\frac{\lceil (j+k-2)^{\epsilon}\rceil}{s'}} j^{4\epsilon\kappa_p}(j+k)^{\epsilon\kappa_p} \\
&\ll j^{5\epsilon\kappa_p}\left[\mathcal{O}(1) + \sum_{k=3}^{m+r-1-j} \vartheta_0^{\frac{\lceil (k-2)^{\epsilon}\rceil}{s'}} (1+k)^{5\epsilon\kappa_p}\right] \ll j^{8\epsilon\kappa_p}.
\end{aligned}
$$

For I_5: by Cauchy-Schwartz inequality,

$$
\begin{aligned}
|I_5| \le \sum_{j=m}^{m+r-1} \left|\mathbb{E}\left(R_j\left(u_{m+r-1}-u_j\right)^2\right)\right| &\le \sum_{j=m}^{m+r-1} \|R_j\|_{L^2}\|u_{m+r-1}-u_j\|_{L^4}^2 \\
&\le \sum_{j=m}^{m+r-1} \|R_j\|_{L^2}\left(\|u_{m+r-1}\|_{L^4}+\|u_j\|_{L^4}\right)^2 \\
&\ll \sum_{j=m}^{m+r-1} j^{2\epsilon\kappa_p}\left[(m+r-1)^{\epsilon\kappa_p}+j^{\epsilon\kappa_p}\right]^2 \\
&\ll (m+r)^{2\epsilon\kappa_p}\sum_{j=m}^{m+r-1} j^{2\epsilon\kappa_p} \\
&\le (m+r)^{2\epsilon\kappa_p}\sum_{j=m}^{m+r-1} j^{6\epsilon\kappa_p} \\
&\le (m+r)^{1+8\epsilon\kappa_p} - m^{1+8\epsilon\kappa_p}.
\end{aligned}
$$

The proof of Lemma 14 is complete. □

We are now ready to show an ASIP for the sequence $\{\xi_j\}_{j\ge 1}$.

Lemma 15. *$\{\xi_j\}_{j\ge1}$ satisfies an ASIP as follows: put $b_r^2 = \mathbb{E}\left(\sum_{j=1}^r \xi_j\right)^2$. There exists a Wiener process $W(\cdot)$ such that*

$$
\left|\sum_{j=1}^r \xi_j - W(b_r^2)\right| = o\left(r^{2\kappa_2\lambda(1+\epsilon)}\right), \ a.s..
$$

Proof. We directly apply Proposition 6 to the L^4-integrable martingale difference sequence $\{(\xi_j, \mathfrak{G}_j)\}_{j\ge1}$. Set $a_r = r^{4\kappa_2\lambda}$, then Condition (3.3) holds by Lemma 14. Condition (3.4) also holds, since by Lemma 12 and (4.1), we have

$$
\sum_{j=1}^{\infty} a_j^{-2}\mathbb{E}|\xi_j|^4 \ll \sum_{j=1}^{\infty} j^{-8\kappa_2\lambda} j^{4\epsilon\kappa_p} \le \sum_{j=1}^{\infty} j^{-\frac{4}{4-\epsilon}} < \infty.
$$

On the one hand, by (4.1), $\epsilon\kappa_p < \kappa_2(4\lambda-1) < \kappa_2$, and thus

$$
b_r = \left\|\sum_{j=1}^r \xi_j\right\|_{L^2} \le \sum_{j=1}^r \|\xi_j\|_{L^2} \ll \sum_{j=1}^r j^{\epsilon\kappa_p} \ll r^{1+\epsilon\kappa_p} \ll r^{1+\kappa_2}.
$$

On the other hand, by Lemma 11 and Lemma 13,

$$
b_r = \left\|\sum_{j=1}^r \xi_j\right\|_{L^2} \ge \left\|\sum_{j=1}^r Y_j\right\|_{L^2} - \left\|\sum_{j=1}^r (Y_j-\xi_j)\right\|_{L^4} \gg r^{\kappa_2} - \mathcal{O}(r^{\epsilon\kappa_p}) \gg r^{\kappa_2}.
$$

Therefore, $r^{2\kappa_2(1-2\lambda)} \ll b_r^2/a_r \ll r^{2\kappa_2(1-2\lambda)+2}$, and hence $\left|\log(b_r^2/a_r)\right| \ll r^{4\kappa_2\lambda\epsilon}$. It is obvious that $\log\log a_r \ll r^{4\kappa_2\lambda\epsilon}$ as well. By Proposition 6, we have

$$\begin{aligned}\left|\sum_{j=1}^{r}\xi_j - W(b_r^2)\right| &= o\left(\left(a_r\left(\left|\log(b_r^2/a_r)\right| + \log\log a_r\right)\right)^{1/2}\right),\ a.s. \\ &\leq o\left(r^{2\kappa_2\lambda(1+\epsilon)}\right),\ a.s..\end{aligned}$$

□

4.5. **ASIP for $X_{\mathbf{f}}$.** Finally, we prove Theorem 1 - the ASIP for the random process $X_{\mathbf{f}} = \{X_n\}_{n\geq 0} = \{f_n \circ T^n\}_{n\geq 0}$. By the previous subsections, we can now write

$$\sum_{k=0}^{n-1} X_k = \sum_{j=1}^{r(n)-1}\xi_j + \sum_{j=1}^{r(n)-1}(Y_j - \xi_j) + U_n + V_n. \tag{4.7}$$

We first compare the variances $\sigma_n^2 = \mathbb{E}\left(\sum_{k=0}^{n-1} X_k\right)^2$ and $b_{r(n)-1}^2 = \mathbb{E}\left(\sum_{j=1}^{r(n)-1}\xi_j\right)^2$.

Lemma 16. $\left|\sigma_n - b_{r(n)-1}\right| \ll n^{\epsilon\kappa_p}$. *As a result, for any Wiener process* $W(\cdot)$,

$$\left|W(\sigma_n^2) - W(b_{r(n)-1}^2)\right| = \mathcal{O}\left(\sigma_n^{2\lambda}\right),\ a.s.$$

Proof. By (4.7), Lemma 10 and Lemma 13,

$$\begin{aligned}\left|\sigma_n - b_{r(n)-1}\right| &\leq \left\|\sum_{j=1}^{r(n)-1}(Y_j - \xi_j) + U_n + V_n\right\|_{L^2} \\ &\leq \left\|\sum_{j=1}^{r(n)-1}(Y_j - \xi_j)\right\|_{L^4} + \|U_n\|_{L^p} + \|V_n\|_{L^p} \\ &= \mathcal{O}\left((r(n)-1)^{\epsilon\kappa_p}\right) + \mathcal{O}(1) + \mathcal{O}\left(n^{\epsilon\kappa_p}\right) \ll n^{\epsilon\kappa_p}.\end{aligned}$$

In the last step, we use the fact that $r(n) \asymp n^{\frac{1}{1+\epsilon}} \ll n$. By (2.1) and (4.1),

$$\begin{aligned}\left|\sigma_n^2 - b_{r(n)-1}^2\right| &\leq \left|\sigma_n - b_{r(n)-1}\right|\left(2\sigma_n + \left|\sigma_n - b_{r(n)-1}\right|\right) \\ &\ll n^{\epsilon\kappa_p}\left(2\sigma_n + n^{\epsilon\kappa_p}\right) \ll \sigma_n^{\epsilon\kappa_p/\kappa_2+1}.\end{aligned}$$

For any Wiener process $W(\cdot)$, the random variables $Z_n := W(\sigma_n^2) - W(b_{r(n)-1}^2)$ follows the normal distribution $N\left(0, \left|\sigma_n^2 - b_{r(n)-1}^2\right|\right)$. By (4.1), we can choose a sufficiently large $s > \frac{4\max\{1,1/\kappa_2\}}{4\lambda-1-\epsilon\kappa_p/\kappa_2} > 4$. Then by Markov's inequality and Jensen's inequality, we have

$$\begin{aligned}\mu\{|Z_n| \geq \sigma_n^{2\lambda}\} \leq \sigma_n^{-2\lambda s}\mathbb{E}|Z_n|^s \leq \sigma_n^{-2\lambda s}\left(\mathbb{E}|Z_n|^2\right)^{s/2} &\ll \sigma_n^{\frac{s}{2}[\epsilon\kappa_p/\kappa_2+1-4\lambda]} \\ &\ll \sigma_n^{-2/\kappa_2} \ll n^{-2},\end{aligned}$$

which implies that $\sum_{n=1}^{\infty}\mu\{|Z_n| \geq \sigma_n^{2\lambda}\} < \infty$. By the Borel-Cantelli lemma (Lemma 3), we get $|Z_n| \ll \sigma_n^{2\lambda}$, a.s.. □

We are now ready to prove our main theorem.

Proof of Theorem 1. First, by (4.7), Lemma 10 and Lemma 13, we have

$$\begin{aligned}\left|\sum_{k=0}^{n-1} X_k - \sum_{j=1}^{r(n)-1} \xi_j\right| &\leq \left|\sum_{j=1}^{r(n)-1} (Y_j - \xi_j)\right| + |U_n| + |V_n| \\ &= \mathcal{O}\left((r(n)-1)^{2\kappa_2\lambda}\right) + \mathcal{O}(1) + \mathcal{O}\left(n^{2\kappa_2\lambda}\right), \ a.s. \\ &= \mathcal{O}\left(n^{2\kappa_2\lambda}\right) = \mathcal{O}\left(\sigma_n^{2\lambda}\right), \ a.s.\end{aligned}$$

By Lemma 15 and Lemma 16, there exists a Wiener process $W(\cdot)$ such that

$$\begin{aligned}\left|\sum_{k=0}^{n-1} X_k - W(\sigma_n^2)\right| &\leq \left|\sum_{k=0}^{n-1} X_k - \sum_{j=1}^{r(n)-1} \xi_j\right| + \left|\sum_{j=1}^{r(n)-1} \xi_j - W(b_{r(n)-1}^2)\right| \\ &\quad + \left|W(\sigma_n^2) - W(b_{r(n)-1}^2)\right| \\ &= \mathcal{O}\left(\sigma_n^{2\lambda}\right) + o\left((r(n)-1)^{2\kappa_2\lambda(1+\epsilon)}\right) + \mathcal{O}\left(\sigma_n^{2\lambda}\right) = \mathcal{O}\left(\sigma_n^{2\lambda}\right), \ a.s..\end{aligned}$$

Here we use the fact that $r(n) \asymp n^{\frac{1}{1+\epsilon}}$ and hence $(r(n)-1)^{2\kappa_2\lambda(1+\epsilon)} \asymp n^{2\kappa_2\lambda} \ll \sigma_n^{2\lambda}$. This completes the proof of Theorem 1. □

5. Applications to Random Processes for Concrete Systems

5.1. **Concrete hyperbolic systems.** Our main result applies to a large class of two-dimensional uniformly hyperbolic systems, including Anosov diffeomorphisms[2] and chaotic billiards. We shall focus on the Sinai dispersing billiards and their conservative perturbations. Since such models were studied in [21, 25, 26], we only remind some basic facts here.

We first recall standard definitions, see [7, 8, 10]. A two-dimensional billiard is a dynamical system where a point moves freely at the unit speed in a domain $Q \subset \mathbb{R}^2$ and bounces off its boundary ∂Q by the laws of elastic reflection. A billiard is dispersing if ∂Q is a finite union of mutually disjoint C^3-smooth curves with strictly positive curvature. Four broad classes of perturbations of the dispersing billiards were considered in [25, 26]:

(a) Tables with shifted, rotated or deformed scatterers;
(b) Billiards under small external forces which bend trajectories during flight;
(c) Billiards with kicks or twists at reflections, including slips along the disk;
(d) Random perturbations comprised of maps with uniform properties (including any of the above classes, or a combination of them).

We treat all the above systems in a universal coordinate system. More precisely, let $M = \partial Q \times [-\pi/2, \pi/2]$ be the collision space, which is a standard cross-section of the billiard flow. The canonical coordinate in M is denoted by (r, φ), where r is the arc length parameter on ∂Q and $\varphi \in [-\pi/2, \pi/2]$ is the angle of reflection. The collision map $T : M \to M$ takes an outward unit vector (r, φ) at ∂Q to the outward unit vector after the next collision, and the singularity of T is caused by the tangential collisions, that is, $S_1 = \partial M \cup T^{-1}(\partial M)$.

[2] By adding the boundaries of the finite Markov partition, a topological mixing C^2 Anosov diffeomorphism satisfies our Assumptions **(H1)-(H5)**.

It was proven in [21, 25, 26] that all the above collision map $T : M \to M$ preserves a mixing SRB measure μ, and the systems (M, T, μ) satisfy the assumptions **(H1)-(H5)** in Section 2.1. Therefore, under conditions in Theorem 1, the ASIP holds for the non-stationary process generated by unbounded observables over those systems.

5.2. **Practical random process.** In this subsection, we discuss some practical processes over the concrete systems in Section 5.1.

5.2.1. *Fluctuation of Lyapunov exponents.* By Birkhoff's ergodic theorem, Pesin entropy formula and the mixing property of the system (M, T, μ), we have

$$\lim_{n\to\infty} \frac{1}{n} \log |D_x^u T^n| = \int \log |D_x^u T| d\mu = h_\mu(T),$$

where $|D_x^u T^n|$ is the Jacobian of T^n at x along the unstable direction, and $h_\mu(T)$ is the metric entropy of the SRB measure μ. We would like to study the fluctuation of the convergence for the ergodic sum given by

$$\log |D_x^u T^n| - nh_\mu(T) = \sum_{k=0}^{n-1} [\log |D_x^u T| - h_\mu(T)] \circ T^k.$$

Unlike in Anosov systems, the log unstable Jacobian function $x \mapsto \log |D_x^u T|$ is unbounded in billiard systems. Recall that M is the phase space of a billiard system with coordinates $x = (r, \varphi)$, then $\log |D_x^u T| \asymp -\log \cos \varphi$ blows up near the singularities $\{\varphi = \pm\frac{\pi}{2}\}$. Nevertheless, $\log |D_x^u T|$ is dynamically Hölder continuous by Assumption (**H3**), and it belongs to L^p for any $p \geq 1$, as

$$\int |\log |D_x^u T||^p \, d\mu \asymp \iint |\log \cos \varphi|^p \cos \varphi dr d\varphi < \infty.$$

More generally, it follows from Theorem 1 that an ASIP holds for the ergodic sum of a dynamically Hölder observable. Here we only assume higher order integrability rather than boundedness for the observable.

Theorem 17. *Suppose that $f \in \mathcal{H} \cap L^p$ for some $p > 4$, such that $\mathbb{E}f = 0$ and the first moment of its auto-correlations is finite, i.e.,*

$$\sum_{n=0}^{\infty} n \, |\mathbb{E}(f \cdot f \circ T^n)| < \infty. \tag{5.1}$$

Then the stationary process $\mathbf{X}_f := \{f \circ T^n\}_{n\geq 0}$ satisfies an ASIP for any error exponent $\lambda \in \left(\frac{1}{4}, \frac{1}{2}\right)$, that is, there is a Wiener process $W(\cdot)$ such that

$$\left|\sum_{k=0}^{n-1} f \circ T^k - \sigma_f W(n)\right| = \mathcal{O}(n^\lambda), \quad a.s.. \tag{5.2}$$

where σ_f^2 is given by the Green-Kubo formula, i.e.,

$$\sigma_f^2 := \sum_{n=-\infty}^{\infty} \mathbb{E}\left(f \cdot f \circ T^n\right) \in [0, \infty). \tag{5.3}$$

Proof. First, we note that the series in (5.3) converges absolutely by Condition (5.1). By direct computation, we have

$$\sigma_n^2 = \mathbb{E}\left(\sum_{k=0}^{n-1} f \circ T^k\right)^2 \quad = \quad n\sigma_f^2 - \sum_{|k|>n} n \; \mathbb{E}\left(f \cdot f \circ T^k\right) - 2\sum_{k=1}^{n-1} k \; \mathbb{E}\left(f \cdot f \circ T^k\right)$$

$$= \quad n\sigma_f^2 + \mathcal{O}(1).$$

Therefore, $\sigma_f^2 = \lim_{n\to\infty} \sigma_n^2/n \in [0,\infty)$. If $\sigma_f = 0$, then σ_n^2 is uniformly bounded. In such case, it is well known that f is a coboundary, i.e., there exists an L^2 function $g : M \to \mathbb{R}$ such that $f = g - g\circ T$ (see e.g. Theorem 18.2.2 in [36]), and thus (5.2) is automatic for any error exponent $\lambda > 0$.

We now focus on the case when $\sigma_f > 0$. Condition (1) in Theorem 1 automatically holds since $f \in \mathcal{H}$. For Condition (2), we have $\sigma_n \asymp \sqrt{n}$ since $0 < \sigma_f < \infty$, that is, $\kappa_2 = \frac{1}{2}$. Also, by stationarity and Minkowski's inequality,

$$\sup_{m\geq 0} \left\| \sum_{k=m}^{m+n-1} f\circ T^k \right\|_{L^p} = \left\| \sum_{k=0}^{n-1} f\circ T^k \right\|_{L^p} \leq n\|f\|_{L^p} \ll n.$$

In other words, $\kappa_p = 1$. By Theorem 1, we obtain the ASIP for any $\lambda \in (\frac{1}{4}, \frac{1}{2})$, that is, there exists a Wiener process $W(\cdot)$ such that

$$\left| \sum_{k=0}^{n-1} f\circ T^k - W(\sigma_n^2) \right| = \mathcal{O}(\sigma_n^{2\lambda}) = \mathcal{O}\left(n^\lambda\right), \quad \text{a.s..}$$

Note that $Z_n := W(\sigma_n^2) - \sigma_f W(n)$ follows the normal distribution $N\left(0, \left|\sigma_n^2 - n\sigma_f^2\right|\right)$, and recall that $\left|\sigma_n^2 - n\sigma_f^2\right| = \mathcal{O}(1)$. Then by Jensen's inequality,

$$\mu\{|Z_n| \geq n^{\frac{1}{4}}\} \leq n^{-2}\mathbb{E}|Z_n|^8 \leq n^{-2}\left(\mathbb{E}|Z_n|^2\right)^4 \ll n^{-2},$$

which implies that $\sum_{n=1}^\infty \mu\{|Z_n| \geq n^{\frac{1}{4}}\} < \infty$. By the Borel-Cantelli lemma (Lemma 3), we get $|Z_n| \ll n^{\frac{1}{4}}$, a.s.. Therefore,

$$\left| \sum_{k=0}^{n-1} f\circ T^k - \sigma_f W(n) \right| \leq \left| \sum_{k=0}^{n-1} f\circ T^k - W(\sigma_n^2) \right| + |Z_n| = \mathcal{O}(n^\lambda), \ \textit{a.s..}$$

□

Remark 3. The stationary ASIP in the special case when $p = \infty$ had been shown by Chernov [12] and many other authors. In this case, Condition (5.1) holds due to the exponential decay of correlations for bounded dynamically Hölder observables.

Moreover, we can relax Condition (5.1) to sub-linear first moment of autocorrelations, i.e., $\sum_{n=0}^\infty n\,|\mathbb{E}(f\cdot f\circ T^n)| \ll n^\eta$ for some $\eta < 1$, then by a slight modification in the proof, we can show that the ASIP in (5.2) holds for any error exponent $\lambda \in \left(\max\{\frac{1}{4}, \frac{\eta}{2}\}, \frac{1}{2}\right)$.

Now we can directly apply Theorem 17 to study the fluctuations of Lyapunov exponents in generic billiard systems for which Markov sieves exist (See Corollary 1.8 and Theorem 7.2 in [9] for more details). For such generic billiards, Condition (5.1) holds for $f = \log|D_x^u T| - h_\mu(T)$ and a broader class of observables. Therefore, by Theorem 17, for any $\lambda \in (\frac{1}{4}, \frac{1}{2})$, there is a Wiener process $W(\cdot)$ such that

$$|\log|D_x^u T^n| - nh_\mu(T) - \sigma_f W(n)| = \mathcal{O}(n^\lambda), \quad \text{a.s..}$$

5.2.2. *Shrinking target problem.* Let $\{A_n\}_{n\geq 0}$ be a sequence of nested Borel subsets of M, i.e., $A_n \supset A_{n+1}$ for any $n \geq 0$. Given $x \in M$, we can study the absolute frequency that the trajectory of x hits the shrinking targets A_n. More precisely, for any $n \geq 1$, we denote

$$N_n(x) = \#\{k \in [0,n) :\ T^k x \in A_k\} = \sum_{k=0}^{n-1} \mathbf{1}_{A_k} \circ T^k(x). \tag{5.4}$$

Note that $\mathbb{E}N_n = \sum_{k=0}^{n-1} \mu(A_k)$ by the invariance of μ under T. We say that the sequence $\{A_n\}_{n\geq 0}$ is dynamically Borel-Cantelli if $\lim_{n\to\infty} N_n = \infty$, a.s..

Similar to a recent result by Haydn, Nicol, Török and Vaienti in [34] (see Theorem 5.1 therein), we obtain the following ASIP for the frequency process N_n of the shrinking target problem.

Theorem 18. *Let $\{A_n\}_{n\geq 0}$ be a sequence of nested Borel subsets such that*

(i) *There is $\beta \in [0,\infty)$ such that $\mathbf{1}_{A_n} \in \mathcal{H}_{0.5}$ and $|\mathbf{1}_{A_n}|^+_{0.5} + |\mathbf{1}_{A_n}|^-_{0.5} \ll n^\beta$;*
(ii) *There is $\gamma \in (0, \frac{3}{4})$ such that $\mu(A_n) \gg n^{-\gamma}$. Moreover, $\mu(A_n) = o(\frac{1}{\log n})$.*

Then the process N_n (as defined in (5.4)*) satisfies the ASIP for any error exponent $\lambda \in \left(\max\{\frac{1}{4}, \frac{1}{8(1-\gamma)}\}, \frac{1}{2}\right)$, that is, there exists a Wiener process $W(\cdot)$ such that*

$$\left|N_n - \mathbb{E}N_n - W(\sigma_n^2)\right| = \mathcal{O}(\sigma_n^{2\lambda}), \quad a.s., \tag{5.5}$$

where $\sigma_n^2 = \mathbb{E}N_n^2 - (\mathbb{E}N_n)^2$.

Remark 4. Here is a particular choice of the sequence $\{A_n\}_{n\geq 0}$ such that Condition (i) in Theorem 18 holds: let A_n be an open subset with boundaries in the singular set $S_{-w(n)} \cup S_{w(n)}$, where $w(n)$ is an sequence of positive integers such that $w(n) \ll \log_2 n$. Then each $\mathbf{1}_{A_n} \in \mathcal{H}_{0.5}$ and $|\mathbf{1}_{A_n}|^+_{0.5} + |\mathbf{1}_{A_n}|^-_{0.5} \leq 2^{w(n)} \ll n^\beta$ for some $\beta > 0$.

Proof of Theorem 18. Without loss of generality, we may assume that $\mu(A_0) \leq \frac{1}{2}$. We take $f_n = \mathbf{1}_{A_n} - \mu(A_n)$, then $N_n - \mathbb{E}N_n = \sum_{k=0}^{n-1} f_k \circ T^k$. It follows from Condition (i) that $\{f_n\}_{n\geq 0}$ satisfies Condition (1) in Theorem 1. For Condition (2), the second moment inequality in (2.1) is automatic since for any $p \geq 1$, any $m \geq 0$ and any $n \geq 1$,

$$\left\| \sum_{k=m}^{m+n-1} f_k \circ T^k \right\|_{L^p} \leq \sum_{k=m}^{m+n-1} \|f_k\|_{L^\infty} \leq 2n.$$

That is, $\kappa_p = 1$.

It remains to show the first moment inequality in (2.1). We follow the arguments of Lemma 2.4 in [33]. First, we claim the following long term iterations: there exists $c > 0$ such that for any $k \geq 0$,

$$\sum_{\ell > k + c\log(k+1)} \left|\mathbb{E}(f_k \circ T^k \cdot f_\ell \circ T^\ell)\right| \ll (k+1)^{-2}. \tag{5.6}$$

Indeed, by Proposition 7, we take $\theta := \max\{\vartheta_0, 2^{-1/4}\} < 1$ and $c > \frac{3+\beta}{-\log\theta}$. Then together by Condition (i), for any $0 \leq k < \ell$ and such that $\ell - k > c\log(k+1)$,

$$\begin{aligned} \left|\mathbb{E}(f_k \circ T^k \cdot f_\ell \circ T^\ell)\right| &\ll C_0\left(4 + 2k^\beta + 2\ell^\beta\right)\theta^{\ell-k} \\ &\leq C_0\left(4 + 2k^\beta + 2^{1+\beta}\left[k^\beta + (\ell-k)^\beta\right]\right)\theta^{\ell-k} \\ &\ll \mathcal{O}(1) + k^\beta\theta^{\ell-k} + (\ell-k)^\beta\theta^{\ell-k} \end{aligned}$$

$$\ll \quad \mathcal{O}(1) + k^{\beta}(k+1)^{-3-\beta} + \mathcal{O}(1) \ll (k+1)^{-3},$$

which immediately implies (5.6). Now we have

$$\begin{aligned}
\sigma_n^2 = \mathbb{E}\left(\sum_{k=0}^{n-1} f_k \circ T^k\right)^2 \quad &= \quad \sum_{k=0}^{n-1} \mathbb{E}(f_k^2) + 2\sum_{k=0}^{n-1} \sum_{\substack{k<\ell<n,\\ \ell\le k+c\log(k+1)}} \mathbb{E}(f_k \circ T^k \cdot f_\ell \circ T^\ell) \\
&\quad +2\sum_{k=0}^{n-1} \sum_{\substack{k<\ell<n,\\ \ell> k+c\log(k+1)}} \mathbb{E}(f_k \circ T^k \cdot f_\ell \circ T^\ell) \\
&= \quad \sum_{k=0}^{n-1} \left(\mu(A_k) - \mu(A_k)^2\right) \\
&\quad +2\sum_{k=0}^{n-1} \sum_{\substack{k<\ell<n,\\ \ell\le k+c\log(k+1)}} \Big(\mu(A_k \cap T^{-(\ell-k)}A_\ell) - \mu(A_k)\mu(A_\ell)\Big) \\
&\quad +2\sum_{k=0}^{n-1} \mathcal{O}\left((k+1)^{-2}\right) \\
&\ge \quad \sum_{k=0}^{n-1} \left(\mu(A_k) - \mu(A_k)^2\right) - 2\sum_{k=0}^{n-1} \sum_{\substack{k<\ell<n,\\ \ell\le k+c\log(k+1)}} \mu(A_k)\mu(A_\ell) + \mathcal{O}(1) \\
&\ge \quad \frac{1}{2}\sum_{k=0}^{n-1} \mu(A_k) - 2c\sum_{k=0}^{n-1} \log(k+1)\mu(A_k)^2 + \mathcal{O}(1) \\
&= \quad \frac{1}{2}\sum_{k=0}^{n-1} \mu(A_k)\left[1 - 4c\log(k+1)\mu(A_k)\right] + \mathcal{O}(1),
\end{aligned}$$

where the last inequality uses the fact that $\mu(A_\ell) \le \mu(A_k) \le \mu(A_0) \le \frac{1}{2}$. By Condition (ii), we further get

$$\sigma_n^2 \gg \sum_{k=0}^{n-1} \mu(A_k)(1 - o(1)) \gg n^{1-\gamma}.$$

In other words, $\kappa_2 = 1 - \gamma$. Applying our main theorem - Theorem 1, we obtain the ASIP for N_n given by (5.5). □

Acknowledgements

The authors would like to thank Nicolai Haydn and Huyi Hu for helpful discussions and suggestions.

References

[1] P. Balint and S. Gouëzel. *Limit theorems in the stadium billiard*, Comm. Math. Phys. **263** (2006), no. 2, 461–512.

[2] P. Balint, N. Chernov, and D. Dolgopyat. *Limit theorems for dispersing billiards with cusps*, Comm. Math. Phys. **308** (2011), 479–510.

[3] I. Berkes and W. Philipp. *Approximation theorems for independent and weakly dependent random vectors,* Ann. Probab. **7** (1979), no. 1, 29–54.

[4] R. Bowen. *Markov partitions for Axiom A diffeomorphisms*, Amer. J. Math. **92** (1970), 725–747.

[5] R. Bowen. *Equilibrium States and the Ergodic Theory of Axiom A Diffeomorphisms*, Lect. Notes in Math. **470**, Springer-Verlag, Berlin-New York, 1975.

[6] L. A. Bunimovich and Ya. G. Sinai *Markov partitions for dispersing billiards*, Commun. Math. Phys. **73** (1980), 247–280.

[7] L. A. Bunimovich, Ya. G. Sinai, and N. I. Chernov. *Markov partitions for two-dimensional hyperbolic billiards*, Russian Math. Surveys **45** (1990), 105–152.

[8] L. A. Bunimovich, Ya. G. Sinai, and N. I. Chernov. *Statistical properties of two-dimensional hyperbolic billiards*, Russian Math. Surveys. **46** (1991), 47–106.

[9] N. I. Chernov. *Limit theorems and Markov approximations for chaotic dynamical systems*, Probab. Theory Related Fields **101** (1995), no. 3, 321–362.

[10] N. I. Chernov. *Decay of correlations in dispersing billiards*, J. Statist. Phys. **94** (1999), 513–556.

[11] N. I. Chernov. *Sinai billiards under small external forces.* Ann. Henri Poincare, **2** (2001), 197–236.

[12] N. I. Chernov. *Advanced statistical properties of dispersing billiards*, J. Statist. Phys. **122** (2006), 1061–1094.

[13] N. I. Chernov. *Sinai billiards under small external forces II.* Ann. Henri Poincare, **9** (2008), 91–107.

[14] N. I. Chernov and D. Dolgopyat. *Brownian Brownian Motion-I*, Memoirs of AMS. **198** (2009).

[15] N. I. Chernov, and D. Kleinbock, *Dynamical Borel-Cantelli lemmas for Gibbs measures*, Israel J. Math. **122**, 2001, 1–27.

[16] N. I. Chernov, and R. Markarian. *Chaotic Billiards*, Math. Surveys Monographs **127**, AMS, Providence, 2006.

[17] N. Chernov, and R. Markarian. *Dispersing billiards with cusps: slow decay of correlations*, Comm. Math. Phys. **270**, 2007, 727–758.

[18] N. Chernov and H.-K. Zhang. *Billiards with polynomial mixing rates*, Nonlineartity **4** (2005), 1527–1553.

[19] N. Chernov and H.-K. Zhang. *A family of chaotic billiards with variable mixing rates*, Stochast. Dynam. **5** (2005), 535–553.

[20] N. Chernov and H.-K. Zhang. *Improved estimate for correlations in billiards*, Commun. Math. Phys. **277** (2008), 305–321.

[21] N. Chernov and H.-K. Zhang. *On statistical properties of hyperbolic systems with singularities*, J. Statist. Phys. **136** (2009), 615–642.

[22] C. Cuny. *Pointwise ergodic theorems with rate with applications to limit theorems for stationary processes*, Stoch. Dyn. **11** (2011), no. 1, 135–155.

[23] C. Cuny, and F. Merlevède. *Strong invariance principles with rate for "reverse" martingale differences and applications*, J. Theoret. Probab. **28** (2015), no. 1, 137–183.

[24] M. Demers and C. Liverani. Stability of statistical properties in two-dimensional piecewise hyperbolic maps, *Trans. Amer. Math. Soc.* **360** (2008), 4777–4814.

[25] M. Demers and H.-K. Zhang. *Spectral analysis of the transfer operator for the Lorentz gas*, J. Modern Dynamics **5** (2011), 665–709.

[26] M. Demers and H.-K. Zhang. *A functional analytic approach to perturbations of the Lorentz gas*, Commun. Math. Phys. **324** (2013), 767–830.

[27] M. Demers and H.-K. Zhang. *Spectral analysis of hyperbolic systems with singularities*, Nonlinearity **27** (2014), 379–433.

[28] M. Denker. *The central limit theorem for dynamical systems,* Dynamical systems and ergodic theory (Warsaw, 1986), 33–62, Banach Center Publ. **23**, PWN, Warsaw, 1989.

[29] E. Eberlein. *On strong invariance principles under dependence assumptions,* Ann. Probab. **14** (1986), no. 1, 260–270.

[30] B. Fayad. *Mixing in the absence of the shrinking target property,* Bull. London Math. Soc. **38** (2006), no. 5, 829–838.

[31] G. Gallavotti and D. Ornstein. *Billiards and Bernoulli schemes*, Commun. Math. Phys. **38** (1974), 83–101.

[32] S. Gouëzel. *Almost sure invariance principle for dynamical systems by spectral methods*, Ann. Probab. **38** (2010), no. 4, 1639–1671.

[33] N. Haydn, M. Nicol, S. Vaienti and L. Zhang. *Central limit theorems for the shrinking target problem*, J. Stat. Phys. **153** (2013), no. 5, 864–887.

[34] N. Haydn, M. Nicol, A. Török and S. Vaienti. *Almost sure invariance principle for sequential and non-stationary dynamical systems*, Trans. Amer. Math. Soc. **369** (2017), no. 8, 5293–5316.

[35] R. Hill and S. Velani. *Ergodic theory of shrinking targets,* Invent. Math. **119**, (1995), no. 1, 175–198.

[36] I.A. Ibragimov and Y.V. Linnik. *Independent and stationary sequences of random variables*, Wolters-Noordhoff, Gröningen, 1971.

[37] A. Katok and J. M. Strelcyn. *Invariant Manifolds, Entropy and Billiards; Smooth Maps with Singularities*, Lect. Notes Math. **1222**, Springer, New York 1986.

[38] J. Marcinkiewicz and A. Zygmund. *Sur les fonctions indépendantes*, Fund. Math. **29** (1937), 60–90.

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

[40] I. Melbourne and M. Nicol. *A vector-valued almost sure invariance principle for hyperbolic dynamical systems*, Ann. Probab. **37** (2009), no. 2, 478–505.

[41] R. Markarian. *Billiards with polynomial decay of correlations*, Ergod. Th. Dynam. Syst. **24** (2004), 177–197.

[42] Ya. Pesin. *Characteristic Lyapunov exponents, and smooth ergodic theory*, Russ. Math. Surv. **32** (1977), no. 4, 55–114.

[43] Ya. Pesin. *Dynamical systems with generalized hyperbolic attractors: hyperbolic, ergodic and topological properties*, Ergod. Th. Dynam. Syst. **12** (1992), 123–152.

[44] W. Philipp and W. Stout. *Almost sure invariance principles for partial sums of weakly dependent random variables*, Memoir. Amer. Math. Soc. **161**: (1975).

[45] D. Ruelle. *Thermodynamic formalism*, Encyclopedia of Mathematics and its Applications, vol 5, Addison-Wesley, 1978.

[46] E. Sataev. *Invariant measures for hyperbolic maps with singularities*, Russ. Math. Surv. **47** (1992), 191–251.

[47] Ya. G. Sinai. *Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards*, Russ. Math. Surv. **25** (1970), 137–189.

[48] Ya. G. Sinai. *Gibbs measures in ergodic theory*, Russ. Math. Surv. **27** (1972), 21–69.

[49] O. Sarig. *Subexponential decay of correlations*, Invent. Math. **150** (2002), 629–653.

[50] Q.M. Shao. *Almost sure invariance principles for mixing sequences of random variables,* Stochastic Process. Appl. **48** (1993), no. 2, 319–334.

[51] Q.M. Shao and C.R. Lu, *Strong approximations for partial sums of weakly dependent random variables,* Sci. Sinica Ser. A **30** (1987), no. 6, 575–587.

[52] M. Stenlund. *A vector-valued almost sure invariance principle for Sinai billiards with random scatterers*, Comm. Math. Phys. **325** (2014), no. 3, 879–916.

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

[54] D. Szász and T. Varjú. *Local limit theorem for Lorentz process and its recurrence in the plane,* Ergodic Theory Dynam. Systems **24** (2004), 257–278.

[55] W.B. Wu. *Strong invariance principles for dependent random variables*, Ann. Probab. **35** (2007), no. 6, 2294–2320.

[56] R. Yokoyama. *Moment bounds for stationary mixing sequences*, Z. Wahrsch. Verw. Gebiete **52** (1980), no.1, 45–57.

[57] L. S. Young. *Statistical properties of systems with some hyperbolicity including certain billiards*, Ann. Math. **147** (1998), 585–650.

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

Department of Mathematics & Statistics, University of Massachusetts Amherst
E-mail address: `jchen@math.umass.edu.`

Mathematics Department, The Graduate Center, City University of New York
E-mail address: `yyang@gc.cuny.edu.`

Department of Mathematics & Statistics, University of Massachusetts Amherst
E-mail address: `hongkun@math.umass.edu.`