Ergod. Th. & Dynam. Sys. (2020), **40**, 2317–2348 © *Cambridge University Press, 2019*
doi:10.1017/etds.2019.2

Rates in almost sure invariance principle for slowly mixing dynamical systems

C. CUNY†, J. DEDECKER‡, A. KOREPANOV§ and F. MERLEVÈDE¶

† *Université de Brest, LMBA, UMR CNRS 6205, France*
(e-mail: christophe.cuny@univ-brest.fr)

‡ *Université Paris Descartes, Sorbonne Paris Cité, Laboratoire MAP5 (UMR 8145), France*
(e-mail: jerome.dedecker@parisdescartes.fr)

§ *Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK*
(e-mail: a.korepanov@warwick.ac.uk)

¶ *Université Paris-Est, LAMA (UMR 8050), UPEM, CNRS, UPEC, France*
(e-mail: florence.merlevede@u-pem.fr)

(*Received* 20 *February* 2018 *and accepted in revised form* 3 *January* 2019)

Abstract. We prove the one-dimensional almost sure invariance principle with essentially optimal rates for slowly (polynomially) mixing deterministic dynamical systems, such as Pomeau–Manneville intermittent maps, with Hölder continuous observables. Our rates have form $o(n^{\gamma}L(n))$, where $L(n)$ is a slowly varying function and γ is determined by the speed of mixing. We strongly improve previous results where the best available rates did not exceed $O(n^{1/4})$. To break the $O(n^{1/4})$ barrier, we represent the dynamics as a Young-tower-like Markov chain and adapt the methods of Berkes–Liu–Wu and Cuny–Dedecker–Merlevède on the Komlós–Major–Tusnády approximation for dependent processes.

Key words: strong invariance principle, KMT approximation, non-uniformly expanding dynamical systems, Markov chain

2010 Mathematics Subject Classification: 37E05 (Primary); 60F17 (Secondary)

1. *Introduction and statement of results*

In their study of turbulent bursts, Pomeau and Manneville [**25**] introduced simple dynamical systems, exhibiting intermittent transitions between 'laminar' and 'turbulent' behaviour. Over the last few decades, such maps have been very popular in dynamical systems. We consider a version of Liverani, Saussol and Vaienti [**19**], where for a fixed

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

$\gamma \in (0, 1)$, the map $f\colon [0, 1] \to [0, 1]$ is given by

$$f(x) = \begin{cases} x(1 + 2^{\gamma} x^{\gamma}), & x \le 1/2, \\ 2x - 1, & x > 1/2. \end{cases} \tag{1.1}$$

There exists a unique absolutely continuous f-invariant probability measure μ on $[0, 1]$, which is equivalent to the Lebesgue measure.

The intermittent behaviour comes from the fact that 0 is a fixed point with $f'(0) = 1$. Hence if a point x is *close* to 0, then its orbit $(f^n(x))_{n\ge 0}$ stays around 0 for a *long* time. The degree of intermittency is given by the parameter γ and is quantified by choosing an interval away from 0 such as $Y =]1/2, 1]$ and considering the first return time $\tau\colon Y \to \mathbb{N}$,

$$\tau(x) = \min\{n \ge 1\colon f^n(x) \in Y\}.$$

It is straightforward to verify [**8**, **32**] that for some $C > 0$ all $n \ge 1$,

$$C^{-1} n^{-1/\gamma} \le \mathrm{Leb}\,(\tau \ge n) \le C n^{-1/\gamma}, \tag{1.2}$$

where Leb denotes the Lebesgue measure on Y.

Suppose that $\varphi\colon [0, 1] \to \mathbb{R}$ is a Hölder continuous observable with $\int \varphi\, d\mu = 0$ and let

$$S_n(\varphi) = \sum_{k=0}^{n-1} \varphi \circ f^k.$$

We consider $S_n(\varphi)$ as a discrete time random process on the probability space $([0, 1], \mu)$. Since μ is f-invariant, the increments $(\varphi \circ f^n)_{n\ge 0}$ are stationary. Using the bound (1.2), Young [**32**] proved that the correlations decay polynomially:

$$\left| \int \varphi\, \varphi \circ f^n \, d\mu \right| = O(n^{-(1-\gamma)/\gamma}). \tag{1.3}$$

If $\gamma < 1/2$, then $S_n(\varphi)$ satisfies the central limit theorem (CLT), that is $n^{-1/2} S_n(\varphi)$ converges in distribution to a normal random variable with variance

$$c^2 = \int \varphi^2 \, d\mu + 2 \sum_{n=1}^{\infty} \int \varphi\, \varphi \circ f^n \, d\mu. \tag{1.4}$$

By (1.3), the series above converges absolutely. The asymptotics in (1.3) is sharp [**7**, **8**, **12**, **30**, **32**], and for each $\gamma \ge 1/2$ there are observables φ for which the series for c^2 diverges, and the CLT does not hold. We are interested in the case when the CLT holds, so from here on we restrict to $\gamma < 1/2$.

In parallel with (1.1), we consider a very similar map

$$f(x) = \begin{cases} x(1 + x^{\gamma} \rho(x)), & x \le 1/2, \\ 2x - 1, & x > 1/2, \end{cases} \tag{1.5}$$

where, following Holland [**11**] and Gouëzel [**8**], $\rho \in \mathcal{C}^2((0, 1/2], (0, \infty))$ is slowly varying at 0 and satisfies:

- $x\rho'(x) = o(\rho(x))$ and $x^2 \rho''(x) = o(\rho(x))$;
- $f(1/2) = 1$ and $f'(x) > 1$ for all $x \ne 0$;

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

- $\displaystyle\int_0^{1/2} \frac{1}{x(\rho(x))^{1/\gamma}}\,dx < \infty.$

For example, take $\rho(x) = C|\log x|^{(1+\varepsilon)\gamma}$ with $\varepsilon > 0$ and $C = 2^{\gamma}(\log 2)^{-(1+\varepsilon)\gamma}$.

Then in place of the bound Leb $(\tau \geq n) \leq Cn^{-1/\gamma}$ in (1.2) we have a slightly stronger bound [**8**, Theorem 1.4.10, Proposition 1.4.12, Lemma 1.4.14]:

$$\int_Y \tau^{1/\gamma}\, d\,\text{Leb} \ < \infty. \tag{1.6}$$

Remark 1.1. The analysis above for the map (1.1) applies to the map (1.5) with minor differences: the correlations decay slightly faster and the CLT holds also for $\gamma = 1/2$ (see [**8**]).

Further we use f to denote either of the maps (1.1) and (1.5), specifying which one we refer to where it makes a difference.

A strong generalization of the CLT and the aim of our work is the following property.

Definition 1.2. We say that a real-valued random process $(S_n)_{n\geq 1}$ satisfies the *almost sure invariance principle* (ASIP) (also known as a *strong invariance principle*) with rate $o(n^{\beta})$, $\beta \in (0,\, 1/2)$, and variance c^2 if one can redefine $(S_n)_{n\geq 1}$ without changing its distribution on a (richer) probability space on which there exists a Brownian motion $(W_t)_{t\geq 0}$ with variance c^2 such that

$$S_n = W_n + o(n^{\beta}) \quad \text{almost surely (a.s.).}$$

We define similarly the ASIP with rates $o(r_n)$ or $O(r_n)$ for deterministic sequences $(r_n)_{n\geq 1}$.

For the map (1.1) with Hölder continuous observables φ, the ASIP for $S_n(\varphi)$ has been first proved by Melbourne and Nicol [**20**], albeit without explicit rates. In [**21**, Theorem 1.6 and Remark 1.7], the same authors obtained the ASIP with rates

$$S_n(\varphi) - W_n = \begin{cases} o(n^{\gamma/2+1/4+\varepsilon}), & \gamma \in\,]1/4,\, 1/2[\\ o(n^{3/8+\varepsilon}), & \gamma \in\,]0,\, 1/4], \end{cases}$$

for all $\varepsilon > 0$. Their proof is based on Philipp and Stout [**26**, Theorem 7.1]. This result has been subsequently improved. Using the approach for the reverse martingales of Cuny and Merlevède [**5**], Korepanov, Kosloff and Melbourne [**17**] proved the ASIP with rates

$$S_n(\varphi) - W_n = \begin{cases} o(n^{\gamma+\varepsilon}), & \gamma \in [1/4,\, 1/2[\\ O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}), & \gamma \in\,]0,\, 1/4[\end{cases}$$

for all $\varepsilon > 0$. (Section 5.2 provides some more details.)

When φ is not Hölder continuous, the situation is more delicate. For instance, functions with discontinuities are not easily amenable to the method of Young towers used in [**17**, **20**, **21**]. For φ of bounded variation, using the conditional quantile method, Merlevède and Rio [**22**] proved the ASIP with rates

$$S_n(\varphi) - W_n = O(n^{\gamma'}(\log n)^{1/2}(\log\log n)^{(1+\varepsilon)\gamma'})$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

for all $\varepsilon > 0$, where $\gamma' = \max\{\gamma, 1/3\}$. Besides considering observables of bounded variation, the results of [**22**] also cover a large class of unbounded observables.

In all the papers above, the rates are not better than $O(n^{1/4})$, which could be perceived as largely suboptimal when $0 < \gamma < 1/4$ due to the intuition coming from the processes with independent and identically distributed (iid) increments [**13**, **14**] and recent related works [**3**, **4**, **23**], where better rates are reached for certain special models. All such models have an underlying independent structure (either functions of a sequence of iid random variables in [**3**, **4**] or independent blocks constructed via the regeneration technique in [**23**]). This allowed using precise results on coupling with Brownian motion for independent random variables such as those in Sakhanenko [**29**] and Zaitsev [**33**]. Recently Korepanov [**16**] applied the result of [**3**] to non-uniformly expanding dynamical systems with exponential tails of the return times using a suitable representation of the system as a Bernoulli shift. The method of [**16**] does not cover systems with slower than exponential decay of return times such as the intermittent maps.

In this paper, representing the dynamics as a Young-tower-like Markov chain similar to [**16**] and adapting the proof of [**4**], we obtain the following result.

THEOREM 1.3. *Let $\gamma \in (0, 1/2)$ and $\varphi\colon [0, 1] \to \mathbb{R}$ be a Hölder continuous observable with $\int \varphi \, d\mu = 0$. For the map (1.1), the random process $S_n(\varphi)$ satisfies the ASIP with variance c^2 given by (1.4) and rate $o(n^\gamma (\log n)^{\gamma+\varepsilon})$ for all $\varepsilon > 0$. For the map (1.5), the random process $S_n(\varphi)$ satisfies the ASIP with variance c^2 given by (1.4) and rate $o(n^\gamma)$.*

The rates in Theorem 1.3 are optimal in the following sense.

PROPOSITION 1.4. *Let f be the map (1.1). There exists a Hölder continuous observable φ with $\int \varphi \, d\mu = 0$ such that*

$$\limsup_{n\to\infty} (n \log n)^{-\gamma} |S_n(\varphi) - W_n| > 0$$

for all Brownian motions $(W_t)_{t\geq 0}$ defined on the same (possibly enlarged) probability space as $(S_n(\varphi))_{n\geq 0}$. Hence, one cannot take $\varepsilon = 0$ in Theorem 1.3.

Remark 1.5. If $c^2 = 0$, the rate in the ASIP can be improved to $O(1)$. Indeed, then it is well known that φ is a *coboundary* in the sense that $\varphi = u - u \circ f$ with some $u\colon [0, 1] \to \mathbb{R}$. By [**8**, Proposition 1.4.2], u is bounded, thus $S_n(\varphi)$ is bounded uniformly in n.

Remark 1.6. It is possible to relax the assumption that φ is Hölder continuous. As a simple example, Theorem 1.3 holds if φ is Hölder on $(0, 1/2)$ and on $(1/2, 1)$, with a discontinuity at $1/2$. See §4.3 for further extensions.

Remark 1.7. Intermittent maps are prototypical examples of *non-uniformly expanding dynamical systems,* to which our results apply in a general setup, and so does the discussion of rates preceding Theorem 1.3. We focus on the maps (1.1) and (1.5) for simplicity only, and discuss the generalization in §5.

The paper is organized as follows. In §2, following Korepanov [**15**], we represent the dynamical systems (1.1) and (1.5) as a function of the trajectories of a particular Markov chain; further, in §3, we introduce a *meeting time* related to the Markov chain and estimate its moments. While §§2 and 3 have an independent interest, they present crucial tools

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

for the proof of Theorem 1.3 as we explain in its outline in §4.1. Next, in §4 we prove Theorem 1.3 for our new process (which is a function of the whole future trajectories of the Markov chain) by adapting the ideas of Berkes, Liu and Wu [**3**] and Cuny, Dedecker and Merlevède [**4**]. In §5 we generalize our results to the class of non-uniformly expanding dynamical systems and show the optimality of the rates.

Throughout, we use the notation $a_n \ll b_n$ and $a_n = O(b_n)$ interchangeably, meaning that there exists a positive constant C not depending on n such that $a_n \le Cb_n$ for all sufficiently large n. As usual, $a_n = o(b_n)$ means that $\lim_{n\to\infty} a_n/b_n = 0$. Recall that $v\colon X \to \mathbb{R}$ is a Hölder observable (with a Hölder exponent $0 < \eta \le 1$) on a bounded metric space (X, d) if $\|v\|_\eta = |v|_\infty + |v|_\eta < \infty$ where $|v|_\infty = \sup_{x\in X} |v(x)|$ and $|v|_\eta = \sup_{x\neq y} |v(x) - v(y)|/d(x, y)^\eta$. All along the paper, we use the notation $\mathbb{N} = \{0, 1, 2, \ldots\}$.

2. *Reduction to a Markov chain*

2.1. *Outline.* In this section we construct a stationary Markov chain $g_0, g_1, \ldots$ on a countable state space S, the space of all possible future trajectories Ω and an observable $\psi\colon \Omega \to \mathbb{R}$ such that the random process $(X_n)_{n\ge0}$ where $X_n = \psi(g_n, g_{n+1}, \ldots)$ has the same distribution as $(\varphi \circ f^n)_{n\ge0}$, the increments of $(S_n(\varphi))_{n\ge1}$.

Our Markov chain is in the spirit of the classical Young towers [**32**]. Just as the Young towers for the maps (1.1) and (1.5), our construction enjoys recurrence properties related to the choice of γ, and we supply Ω with a metric, with respect to which ψ is Lipschitz.

We follow the ideas of [**16**], though in the setup of the maps (1.1) and (1.5) we are able to make the proofs simpler and hopefully easier to read.

2.2. *Basic properties of intermittent maps.* A standard way to work with maps (1.1), (1.5) is an inducing scheme. As in §1, set $Y =]1/2, 1]$ and let $\tau\colon Y \to \mathbb{N}$ be the *inducing time,* $\tau(x) = \min\{k \ge 1\colon f^k(x) \in Y\}$. Let $F\colon Y \to Y$ be the *induced map*, $F(x) = f^{\tau(x)}(x)$. Let α be the partition of Y into the intervals where τ is constant. Let $\beta = 1/\gamma$.

We remark that $\gcd\{\tau(a)\colon a \in \alpha\} = 1$.

Let m denote the Lebesgue measure on Y, normalized so that it is a probability measure. Recall that we have the bounds:

- $m(\tau \ge n) \le Cn^{-\beta}$ for all $n \ge 1$ for the map (1.1);
- $\int \tau^\beta \, dm < \infty$ for the map (1.5).

The induced map F satisfies the following properties:

- (full image) $F\colon a \to Y$ is a bijection for each $a \in \alpha$;
- (expansion) there is $\lambda > 1$ such that $|F'| \ge \lambda$;
- (bounded distortion) there is a constant $C_d \ge 0$ such that

$$|\log |F'(x)| - \log |F'(y)|| \le C_d |F(x) - F(y)|$$

for all $x, y \in a, a \in \alpha$.

2.3. *Disintegration of the Lebesgue measure.* The properties in §2.2 allow a disintegration of the measure m, as described in this subsection.

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

Let $\mathcal{A}$ denote the set of all finite words in the alphabet α, not including the empty word. For $w = a_0 \cdots a_{n-1} \in \mathcal{A}$, let $|w| = n$ and let Y_w denote the cylinder of points in Y, which follow the itinerary of letters of w under the iteration of F:

$$Y_w = \{y \in Y \colon F^k(y) \in a_k \text{ for } 0 \le k \le n-1\}.$$

Let also $h \colon \mathcal{A} \to \mathbb{N}$, $h(w) = \tau(a_0) + \cdots + \tau(a_{n-1})$ for $w = a_0 \cdots a_{n-1}$.

For $w_0, \ldots, w_n \in \mathcal{A}$, let $w_0 \cdots w_n \in \mathcal{A}$ denote the concatenation.

PROPOSITION 2.1. *For each infinite sequence $a_0, a_1, \ldots \in \alpha$, there exists a unique $y \in Y$ such that $F^n(y) \in a_n$ for all $n \ge 0$.*

In particular, for each sequence $w_0, w_1, \ldots \in \mathcal{A}$ there exists a unique $y \in Y$ such that $y \in Y_{w_0}$, $F^{|w_0|}(y) \in Y_{w_1}$, $F^{|w_0|+|w_1|}(y) \in Y_{w_2}$, and so on.

Proof. Uniqueness of y follows from expansion of F, so it is enough to show existence.

Let $w_n = a_0 \cdots a_{n-1}$. Note that Y_{w_n}, $n \ge 0$, is a nested sequence of intervals with shrinking to 0 length, closed on the right and open on the left. Let y be the only point in the intersection of their closures, $\{y\} = \bigcap_n \bar{Y}_{w_n}$.

Suppose that $y \notin \bigcap_n Y_{w_n}$. Then $y \in \bar{Y}_{w_n} \setminus Y_{w_n}$ for some n, thus y is a left end-point of Y_{w_n}. Observe that $\bar{Y}_{w_{n+1}}$ is contained in $\bar{Y}_{w_n}$ but cannot contain its left end-point, i.e. $y \notin \bar{Y}_{w_{n+1}}$. This is a contradiction, proving that $y \in \bigcap_n Y_{w_n}$. Hence $F^n(y) \in a_n$ for all n, as required. □

PROPOSITION 2.2. *There exist a probability measure $\mathbb{P}_{\mathcal{A}}$ on $\mathcal{A}$ and a disintegration*

$$m = \sum_{w \in \mathcal{A}} \mathbb{P}_{\mathcal{A}}(w) m_w,$$

where:

- *each m_w is a probability measure supported on Y_w;*
- *$(F^{|w|})_* m_w = m$;*
- *$\mathbb{P}_{\mathcal{A}}(w) > 0$ for each w;*
- *for the map (1.1), $\mathbb{P}_{\mathcal{A}}(h \ge k) \le C_\beta k^{-\beta}$ for all $k \ge 1$, where $C_\beta > 0$ is a constant;*
- *for the map (1.5), $\int h^\beta \, d\mathbb{P}_{\mathcal{A}} < \infty$.*

The disintegration in Proposition 2.2 was introduced in [**35**] and called *regenerative partition of unity*. The bounds on the tail of h are proved in [**15**]. This disintegration is the basis of the Markov chain construction.

2.4. *Construction of the Markov chain.* Let $g_0, g_1, \ldots$ be a Markov chain with state space

$$S = \{(w, \ell) \in \mathcal{A} \times \mathbb{Z} \colon 0 \le \ell < h(w)\}$$

and transition probabilities

$$\mathbb{P}(g_{n+1} = (w, \ell) \mid g_n = (w', \ell')) = \begin{cases} 1, & \ell = \ell' + 1 \text{ and } \ell' + 1 < h(w) \text{ and } w = w', \\ \mathbb{P}_{\mathcal{A}}(w), & \ell = 0 \text{ and } \ell' + 1 = h(w'), \\ 0 & \text{else.} \end{cases} \tag{2.1}$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

The Markov chain $g_0, g_1, \ldots$ has a unique (hence ergodic) invariant probability measure ν on S, given by

$$\nu(w, \ell) = \frac{\mathbb{P}_{\mathcal{A}}(w)\mathbf{1}_{\{0\le\ell<h(w)\}}}{\sum_{(w,\ell)\in S}\mathbb{P}_{\mathcal{A}}(w)} = \frac{\mathbb{P}_{\mathcal{A}}(w)\mathbf{1}_{\{0\le\ell<h(w)\}}}{\mathbb{E}_{\mathcal{A}}(h)}. \tag{2.2}$$

The Markov chain $g_0, g_1, \ldots$ starting from ν defines a probability measure $\mathbb{P}_\Omega$ on the space $\Omega \subset S^{\mathbb{N}}$ of sequences, which correspond to non-zero probability transitions. Let $\sigma : \Omega \to \Omega$ be the left shift action,

$$\sigma(g_0, g_1, \ldots) = (g_1, g_2, \ldots).$$

Remark 2.3. There exists $w \in \mathcal{A}$ with $\mathbb{P}_{\mathcal{A}}(w) > 0$ and $h(w) = 1$. Therefore, the Markov chain $g_0, g_1, \ldots$ is aperiodic. Aperiodicity is used in the proof of the ASIP (namely, in the proof of Lemma 3.1 to apply Lindvall's result [**18**]). However, in the general case, as far as the ASIP is concerned, aperiodicity is not necessary (see §5).

We supply the space Ω with a *separation time* $s : \Omega \times \Omega \to \mathbb{N} \cup \{\infty\}$, measured in terms of the number of visits to $S_0 = \{(w, \ell) \in S : \ell = 0\}$ as follows. For $a, b \in \Omega$,

$$\begin{aligned} a &= (g_0, \ldots, g_N, g_{N+1}, \ldots), \\ b &= (g_0, \ldots, g_N, g'_{N+1}, \ldots) \end{aligned} \tag{2.3}$$

with $g_{N+1} \neq g'_{N+1}$, we set

$$s(a, b) = \#\{0 \le n \le N : g_n \in S_0\}.$$

We define a *separation metric* d on Ω by

$$d(a, b) = \lambda^{-s(a,b)}. \tag{2.4}$$

For $g = (w, \ell) \in S$, define $X_g \subset [0, 1]$, $X_g = f^\ell(Y_w)$. Then, similar to Proposition 2.1, to each $(g_0, g_1, \ldots) \in \Omega$ there corresponds a unique $x \in [0, 1]$ such that $f^n(x) \in X_{g_n}$ for all $n \ge 0$ (but for a given x, there may be many such $(g_0, g_1, \ldots) \in \Omega$).

Thus we introduce a projection $\pi : \Omega \to [0, 1]$, with $\pi(g_0, g_1, \ldots) = x$ where $f^n(x) \in X_{g_n}$ for all $n \ge 0$ as above.

The key properties of the projection π are the following.

LEMMA 2.4.

- *π is Lipschitz: $|\pi(a) - \pi(b)| \le d(a, b)$ for all $a, b \in \Omega$;*
- *π is a measure preserving map between the probability spaces $(\Omega, \mathbb{P}_\Omega)$ and $([0, 1], \mu)$;*
- *π is a semiconjugacy between $\sigma : \Omega \to \Omega$ and $f : [0, 1] \to [0, 1]$, i.e. the following diagram commutes:*

$$\begin{array}{ccc} \Omega & \xrightarrow{\sigma} & \Omega \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ [0, 1] & \xrightarrow{f} & [0, 1] \end{array}$$

COROLLARY 2.5. *Suppose that $\varphi : [0, 1] \to \mathbb{R}$ is Hölder continuous. Let $\psi = \varphi \circ \pi$ and $X_k = \psi(g_k, g_{k+1}, \ldots)$ for $k \ge 0$. Then:*

(a) *ψ is Hölder continuous;*

(b) *The process $(X_k)_{k\ge0}$ on the probability space $(\Omega, \mathbb{P}_\Omega)$ is equal in law to $(\varphi \circ f^k)_{k\ge0}$ on $([0, 1], \mu)$.*

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

2.5. *Proof of Lemma 2.4.* The last item, namely, the property that $\pi \circ \sigma = f \circ \pi$ follows directly from the construction of σ and π.

We prove now the first item. Suppose that $a, b \in \Omega$ are as in (2.3) and write

$$\begin{aligned} &g_0, \dots, g_N \\ &\quad = (w_0, \ell_0), \dots, (w_0, h(w_0) - 1), (w_1, 0), \dots, (w_1, h(w_1) - 1), \dots, \\ &\qquad (w_k, 0), \dots, (w_k, \ell_k), \end{aligned}$$

where $0 \le \ell_0 < h(w_0)$, $0 \le \ell_k < h(w_k)$ and $h(w_0) - \ell_0 + \sum_{i=1}^{k-1} h(w_i) + \ell_k = N$. Then both $\pi(a)$ and $\pi(b)$ belong to $f^{\ell_0}(Y_{w_0 \cdots w_k})$.

Suppose that $\ell_0 \neq 0$. Then $s(a, b) = k$. Since $|f'| \ge 1$ and $|F'| \ge \lambda$,

$$\operatorname{diam} f^{\ell_0}(Y_{w_0 \cdots w_k}) \le \operatorname{diam} Y_{w_1 \cdots w_k} \le \lambda^{-k}.$$

Then $|\pi(a) - \pi(b)| \le \lambda^{-s(a,b)} = d(a, b)$. If $\ell_0 = 0$, then $s(a, b) = k + 1$ and we have $\operatorname{diam} f^{\ell_0}(Y_{w_0 \cdots w_k}) \le \lambda^{-(k+1)}$. Again, $|\pi(a) - \pi(b)| \le d(a, b)$, as required.

It remains to prove the second item, namely: $\pi_* \mathbb{P}_\Omega = \mu$. Let $\Omega_0 = \{(g_0, g_1, \dots) \in \Omega \colon g_0 \in S_0\}$. Then $\mathbb{P}_\Omega(\Omega_0) > 0$. Let

$$\mathbb{P}_{\Omega_0}(\cdot) = \frac{\mathbb{P}_\Omega(\cdot \cap \Omega_0)}{\mathbb{P}_\Omega(\Omega_0)}$$

be the corresponding conditional probability measure. We shall use the following intermediate result whose proof is given later.

PROPOSITION 2.6. $\pi_* \mathbb{P}_{\Omega_0} = m$.

Let us complete the proof of the second item with the help of this proposition. Note that $\sigma \colon \Omega \to \Omega$ preserves the ergodic probability measure $\mathbb{P}_\Omega$. Since $f \circ \pi = \pi \circ \sigma$, the measure $\upsilon := \pi_* \mathbb{P}_\Omega$ on $[0, 1]$ is f-invariant and ergodic, as is μ.

Suppose that υ and μ are different measures. Since they are both f-invariant and ergodic, they are singular with respect to each other: there exists $A \subset [0, 1]$ such that $\mu(A) = 1$ and $\upsilon(A) = 0$.

Let $\upsilon|_Y$ and $\mu|_Y$ denote the restrictions on Y. By Proposition 2.6, $m \ll \upsilon|_Y$. Since in turn $\mu|_Y \ll m$, it follows that $\mu|_Y \ll \upsilon|_Y$. Hence $\mu(A \cap Y) = \upsilon(A \cap Y) = 0$. Also, $\mu(Y \setminus A) = 0$, so $\mu(Y) = 0$, which contradicts the fact that μ is equivalent to the Lebesgue measure on $[0, 1]$. Thus $\mu = \upsilon$.

To end the proof of the second item, it remains to show Proposition 2.6.

Proof of Proposition 2.6. Our strategy is to show that for each $w \in \mathcal{A}$,

$$\mathbb{P}_{\Omega_0}(\pi^{-1}(Y_w)) = m(Y_w).$$

Then the result follows from Carathéodory's extension theorem.

Let $m = \sum_{w \in \mathcal{A}} \mathbb{P}_{\mathcal{A}}(w) m_w$ be the decomposition from Proposition 2.2. Recall that each m_w is supported on Y_w and $(F^{|w|})_* m_w = m$. Since $F^{|w|} \colon Y_w \to Y$ is a diffeomorphism between two intervals, the measures m_w are uniquely determined by these properties. It is straightforward to write $m_w = \sum_{w' \in \mathcal{A}} \mathbb{P}_{\mathcal{A}}(w') m_{ww'}$ for each w. (Here ww' is the

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

concatenation of w, w' and the measures $m_{ww'}$ are from the same decomposition.) Thus we obtain a decomposition

$$m = \sum_{w,w' \in \mathcal{A}} \mathbb{P}_{\mathcal{A}}(w)\mathbb{P}_{\mathcal{A}}(w')m_{ww'}.$$

Further, for $n \geq 0$, we write

$$m = \sum_{w_0,\dots,w_n \in \mathcal{A}} \mathbb{P}_{\mathcal{A}}(w_0) \cdots \mathbb{P}_{\mathcal{A}}(w_n)m_{w_0 \cdots w_n}.$$

Suppose that $w \in \mathcal{A}$ with $|w| = n + 1$. For every $w_0, \dots, w_n \in \mathcal{A}$, either $Y_{w_0 \cdots w_n} \subset Y_w$ (when the word $w_0 \cdots w_n$ starts with w) or $Y_{w_0 \cdots w_n} \cap Y_w = \emptyset$ (otherwise). Hence

$$m(Y_w) = \sum_{\substack{w_0,\dots,w_n \in \mathcal{A}: \\ Y_{w_0 \cdots w_n} \subset Y_w}} \mathbb{P}_{\mathcal{A}}(w_0) \cdots \mathbb{P}_{\mathcal{A}}(w_n). \tag{2.5}$$

For $w_0, \dots, w_n$, let $\Omega_{w_0,\dots,w_n}$ denote the subset of Ω_0 with the first coordinates

$$(w_0, 0), \dots, (w_0, h(w_0) - 1), \dots, (w_n, 0), \dots, (w_n, h(w_n) - 1).$$

Note that $\pi(\Omega_{w_0,\dots,w_n}) = Y_{w_0 \cdots w_n}$ and by (2.1),

$$\mathbb{P}_{\Omega_0}(\Omega_{w_0,\dots,w_n}) = \mathbb{P}_{\mathcal{A}}(w_0) \cdots \mathbb{P}_{\mathcal{A}}(w_n).$$

Then

$$\mathbb{P}_{\Omega_0}(\pi^{-1}(Y_w)) = \sum_{\substack{w_0,\dots,w_n \in \mathcal{A}: \\ Y_{w_0 \cdots w_n} \subset Y_w}} \mathbb{P}_{\Omega_0}(\Omega_{w_0,\dots,w_n}) = \sum_{\substack{w_0,\dots,w_n \in \mathcal{A}: \\ Y_{w_0 \cdots w_n} \subset Y_w}} \mathbb{P}_{\mathcal{A}}(w_0) \cdots \mathbb{P}_{\mathcal{A}}(w_n). \tag{2.6}$$

Combining (2.5) and (2.6), we obtain that $\mathbb{P}_{\Omega_0}(\pi^{-1}(Y_w)) = m(Y_w)$, as required. □

3. *Meeting time*

In §2 we constructed the stationary and aperiodic Markov chain $(g_n)_{n \geq 0}$. In this section we introduce a meeting time on it and use it to prove a number of statements, which shall play a central role in the proof of the ASIP.

We work with the notation of §2. Without changing the distribution, we redefine the Markov chain $g_0, g_1, \dots$ on a new probability space as follows. Let $g_0 \in S$ be distributed according to ν (the stationary distribution defined by (2.2)). Let $\varepsilon_1, \varepsilon_2, \dots$ be a sequence of iid random variables with values in $\mathcal{A}$, distribution $\mathbb{P}_{\mathcal{A}}$, independent from g_0. For $n \geq 0$ let

$$g_{n+1} = U(g_n, \varepsilon_{n+1}), \tag{3.1}$$

where

$$U((w, \ell), \varepsilon) = \begin{cases} (w, \ell + 1), & \ell < h(w) - 1, \\ (\varepsilon, 0), & \ell = h(w) - 1. \end{cases} \tag{3.2}$$

We refer to $(\varepsilon_n)_{n \geq 1}$ as *innovations*.

Let g_0^* be a random variable in S with distribution ν, independent from g_0 and $(\varepsilon_n)_{n \geq 1}$. Let $g_0^*, g_1^*, g_2^*, \dots$ be a Markov chain given by

$$g_{n+1}^* = U(g_n^*, \varepsilon_{n+1}) \quad \text{for } n \geq 0. \tag{3.3}$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

Thus the chains $(g_n)_{n\geq 0}$ and $(g_n^*)_{n\geq 0}$ have independent initial states, but share the same innovations. Define the meeting time:

$$T = \inf\{n \geq 0 \colon g_n = g_n^*\}. \tag{3.4}$$

For $\beta, \eta > 1$, define $\psi_{\beta,\eta}, \widetilde{\psi}_{\beta,\eta} \colon [0, \infty) \to [0, \infty)$,

$$\psi_{\beta,\eta}(x) = x^{\beta}(\log(1+x))^{-\eta}, \quad \widetilde{\psi}_{\beta,\eta}(x) = x^{\beta-1}(\log(1+x))^{-\eta}$$

for $x > 0$ and $\psi_{\beta,\eta}(0) = \widetilde{\psi}_{\beta,\eta}(0) = 0$.

For the maps (1.1) and (1.5), moments of T can be estimated by Proposition 2.2 and the following lemma.

LEMMA 3.1. *Suppose that $\beta > 1$.*
(a) *If $\mathbb{P}_{\mathcal{A}}(h \geq k) \ll k^{-\beta}$, then $\mathbb{E}(\widetilde{\psi}_{\beta,\eta}(T)) < \infty$ for all $\eta > 1$.*
(b) *If $\int h^{\beta}\, d\mathbb{P}_{\mathcal{A}} < \infty$, then $\mathbb{E}(T^{\beta-1}) < \infty$.*

Proof. Let $S_c = \{(w, \ell) \in S \colon \ell = h(w) - 1\}$ be the 'ceiling' of S and

$$T^* = \inf\{n \geq 0 \colon g_n \in S_c \text{ and } g_n^* \in S_c\}.$$

From the representation (3.1), it is clear that $T \leq T^* + 1$.

Now, the segments $(g_0, g_1, \ldots, g_{T^*})$ and $(g_0^*, g_1^*, \ldots, g_{T^*}^*)$ never use the same innovations and behave independently. In addition, $g_{T^*+1} = g_{T^*+1}^* = (\varepsilon_{T^*+1}, 0)$ and $g_{n+T^*} = g_{n+T^*}^*$ for any $n \geq 1$.

Consider $(\varepsilon_n')_{n\geq 1}$, an independent copy of $(\varepsilon_n)_{n\geq 1}$, independent also from g_0. Let g_0' be a random variable in S with distribution ν, independent from $(g_0, (\varepsilon_n)_{n\geq 1}, (\varepsilon_n')_{n\geq 1})$. Define the Markov chain $(g_n')_{n\geq 0}$ by

$$g_{n+1}' = U(g_n', \varepsilon_{n+1}') \quad \text{for } n \geq 0.$$

Let

$$T' = \inf\{n \geq 0 \colon g_n \in S_c \text{ and } g_n' \in S_c\}.$$

Due to the previous considerations, T' is equal to T^* in law.

Note that S_c is a recurrent atom for the Markov chain $(g_n)_{n\geq 0}$ (see, for instance, [**24**, Ch. 4.2] for the definition of an atom). Let

$$\tau_0 = \inf\{n \geq 0 \colon g_n \in S_c\}$$

be the first renewal time. If $\mathbb{P}_{\mathcal{A}}(h \geq k) \ll k^{-\beta}$, we claim that for all $\eta > 1$,

$$\mathbb{E}(\widetilde{\psi}_{\beta,\eta}(\tau_0)) < \infty.$$

Then, according to Lindvall [**18**] (see also Rio [**28**, Proposition 9.6]), since the chain $(g_n)_{n\geq 0}$ is aperiodic (see Remark 2.3), $\mathbb{E}(\widetilde{\psi}_{\beta,\eta}(T')) < \infty$ and (a) follows. For (b), the argument is similar, with x^{β} instead of $\psi_{\beta,\eta}(x)$ and $x^{\beta-1}$ instead of $\widetilde{\psi}_{\beta,\eta}(x)$.

It remains to verify the claim. Note that if $g_0 = (w, \ell)$, then $\tau_0 = h(w) - \ell - 1$ and

$$\widetilde{\psi}_{\beta,\eta}(\tau_0) = \frac{(h(w) - \ell - 1)^{\beta-1}}{(\log(h(w) - \ell))^{\eta}} \leq C_{\beta,\eta} \frac{h(w)^{\beta-1}}{(\log h(w))^{\eta}}.$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

For any $\eta > 1$, using that $\nu(w, \ell) \le \mathbb{P}_{\mathcal{A}}(w)/\mathbb{E}_{\mathcal{A}}(h)$, write

$$\begin{aligned}\mathbb{E}(\widetilde{\psi}_{\beta,\eta}(\tau_0)) &= \sum_{\substack{w\in\mathcal{A},\\ 0\le\ell<h(w)}} \mathbb{E}_{g_0=(w,\ell)}(\widetilde{\psi}_{\beta,\eta}(\tau_0))\nu(w,\ell)\\ &\le C_{\beta,\eta}(\mathbb{E}_{\mathcal{A}}(h))^{-1}\sum_{w\in\mathcal{A}}\frac{h(w)^{\beta}}{(\log h(w))^{\eta}}\mathbb{P}_{\mathcal{A}}(w)\\ &= C_{\beta,\eta}(\mathbb{E}_{\mathcal{A}}(h))^{-1}\mathbb{E}_{\mathcal{A}}(\psi_{\beta,\eta}(h)) < \infty,\end{aligned}$$

by taking into account Proposition 2.2. □

Let $\psi\colon \Omega \to \mathbb{R}$ be a Hölder continuous observable with $\int \psi\, d\mathbb{P}_\Omega = 0$. (Such as $\psi = \varphi \circ \pi$ in §2.) For $\ell \ge 0$, define $\delta_\ell\colon \Omega \to \mathbb{R}$,

$$\delta_\ell(g_0, g_1, \ldots) = \sup|\psi(g_0, g_1, \ldots, g_{\ell+1}, g_{\ell+2}, \ldots) - \psi(g_0, g_1, \ldots, \tilde{g}_{\ell+1}, \tilde{g}_{\ell+2}, \ldots)|,$$

where the supremum is taken over all possible trajectories $(\tilde{g}_{\ell+1}, \tilde{g}_{\ell+2}, \ldots)$.

PROPOSITION 3.2. *Assume that* $\mathbb{E}(T) < \infty$. *For all* $r \ge 1$,

$$\mathbb{E}(\delta_\ell) \ll \ell^{-r/2} + \mathbb{P}(T \ge [\ell/r]).$$

Proof. By (2.4) and the first item of Lemma 2.4, there exist $C > 0$ (depending on the Hölder norm of ψ) and $\theta \in (0, 1)$ (depending on λ and on the Hölder exponent of ψ) such that $\delta_\ell \le C\theta^{s_\ell}$, where $s_\ell = \#\{k \le \ell\colon g_k \in S_0\}$. Write

$$\begin{aligned}C^{-1}\mathbb{E}(\delta_\ell) \le \mathbb{E}(\theta^{s_\ell}) &\le \theta^{\frac{1}{2}(\ell+1)\mathbb{P}(g_0\in S_0)} + \mathbb{E}(\theta^{s_\ell}\mathbf{1}_{s_\ell < \frac{1}{2}(\ell+1)\mathbb{P}(g_0\in S_0)})\\ &\le \theta^{\frac{1}{2}(\ell+1)\mathbb{P}(g_0\in S_0)} + \mathbb{P}(s_\ell < \tfrac{1}{2}(\ell+1)\mathbb{P}(g_0 \in S_0)).\end{aligned} \tag{3.5}$$

Next,

$$\mathbb{P}\left(s_\ell < \frac{1}{2}(\ell+1)\mathbb{P}(g_0\in S_0)\right) \le \mathbb{P}\left(\left|\sum_{i=0}^{\ell}\mathbf{1}_{\{g_i\in S_0\}} - (\ell+1)\nu(S_0)\right| > \frac{1}{2}(\ell+1)\nu(S_0)\right).$$

Recall now the definition (3.4) of the meeting time T and the following coupling inequality: for all $n \ge 1$,

$$\beta(n) := \frac{1}{2}\int \|\delta_{(x,y)}(P\times P)^n - \nu\times\nu\|_v\, d(\nu\times\nu)(x,y) \le \mathbb{P}(T \ge n), \tag{3.6}$$

where $\|\cdot\|_v$ denotes the total variation norm of a signed measure and P is the transition function of the Markov chain $(g_k)_{k\ge0}$ (see [**27**, Lemma 2.6]). From $\mathbb{E}(T) < \infty$, it follows that $\sum_{n\ge1}\beta(n) < \infty$. Applying [**28**, Theorem 6.2] and using that $\alpha(n) \le \beta(n)$, where $(\alpha(n))_{n\ge1}$ is the sequence of strong mixing coefficients defined in [**28**, (2.1)], we infer that for all $r \ge 1$,

$$\mathbb{P}\left(\left|\sum_{i=0}^{\ell}\mathbf{1}_{\{g_i\in S_0\}} - (\ell+1)\nu(S_0)\right| > \frac{1}{2}(\ell+1)\nu(S_0)\right) \le c_1\ell^{-r/2} + c_2\mathbb{P}(T \ge [\ell/r]), \tag{3.7}$$

where c_1 and c_2 are positive constant independent of ℓ. The result follows. □

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

For $n \geq 0$, let

$$X_n = \psi \circ \sigma^n = \psi(g_n, g_{n+1}, \ldots).$$

Then $(X_n)_{n\geq 0}$ is a stationary random process. It is straightforward to use the meeting time to estimate correlations.

LEMMA 3.3. *Assume that* $\mathbb{E}(T) < \infty$. *Then for all* $k \geq 1$ *and* $\alpha \geq 1$,

$$|\mathrm{Cov}\,(X_0, X_k)| \ll k^{-\alpha/2} + \mathbb{P}(T \geq [k/4\alpha]).$$

Proof. Let $k \geq 2$. Let $(\varepsilon_i')_{i\geq 1}$ be an independent copy of the innovations $(\varepsilon_i)_{i\geq 1}$, independent also from g_0. Define $(g_i')_{i\geq k-[k/2]+1}$ by $g'_{k-[k/2]+1} = U(g_{k-[k/2]}, \varepsilon'_{k-[k/2]+1})$ and $g'_{i+1} = U(g'_i, \varepsilon'_{i+1})$ for $i > k - [k/2]$.

Let

$$X_{0,k} = \mathbb{E}_g(\psi(g_0, g_1, \ldots, g_{k-[k/2]}, (g_i')_{i\geq k-[k/2]+1})),$$

where $\mathbb{E}_g$ denotes the conditional expectation given $g := (g_n)_{n\geq 0}$. Write

$$|\mathrm{Cov}\,(X_0, X_k)| \leq \|X_k\|_\infty \|X_0 - X_{0,k}\|_1 + |\mathbb{E}(X_{0,k}X_k)|.$$

Note that $\|X_k\|_\infty \leq |\psi|_\infty < \infty$. By Proposition 3.2, for any $\alpha \geq 1$,

$$\|X_0 - X_{0,k}\|_1 \ll k^{-\alpha/2} + \mathbb{P}(T \geq [k/(4\alpha)]).$$

Hence it is enough to show that

$$|\mathbb{E}(X_{0,k}X_k)| \ll \mathbb{P}(T \geq [k/2]). \tag{3.8}$$

With this aim, note that by the Markovian property and stationarity,

$$|\mathbb{E}(X_{0,k}X_k)| \leq \|X_{0,k}\|_\infty \|\mathbb{E}(X_k \mid g_{k-[k/2]})\|_1 \leq |\psi|_\infty \|\mathbb{E}(X_{[k/2]} \mid g_0)\|_1.$$

Recall the definition of the Markov chain $(g_n^*)_{n\geq 0}$. For all $n \geq 0$, let $X_n^* = \psi((g_k^*)_{k\geq n})$. Since $\mathbb{E}(X_{[k/2]}^*) = 0$ and $X_{[k/2]}^*$ is independent from g_0,

$$\|\mathbb{E}(X_{[k/2]} \mid g_0)\|_1 \leq \|X_{[k/2]} - X_{[k/2]}^*\|_1.$$

Note now that $X_{[k/2]} \neq X_{[k/2]}^*$ only if $T > [k/2]$. Hence

$$\|X_{[k/2]} - X_{[k/2]}^*\|_1 \leq 2|\psi|_\infty \mathbb{P}(T > [k/2]),$$

which proves (3.8) and thus completes the proof of the lemma. □

For $n \geq 1$, let $S_n = \sum_{k=1}^n X_k$. From Lemma 3.3, we get

COROLLARY 3.4. *Assume that* $\mathbb{E}(T) < \infty$. *Then the limit*

$$c^2 = \lim_{n\to\infty} \frac{1}{n} \|S_n\|_2^2$$

exists and

$$c^2 = \|X_0\|_2^2 + 2\sum_{n=1}^{\infty} \mathrm{Cov}\,(X_0, X_n).$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

LEMMA 3.5. *Assume that* $\mathbb{E}(T) < \infty$. *Then, for any* $x > 0$ *and any* $r \geq 1$,

$$\mathbb{P}\big(\max_{k\leq n} |S_k| \geq 5x\big) \ll \frac{n}{x}(x^{-r} + \mathbb{P}(T \geq Cx)) + (1 + \kappa x^2/n)^{-r/2}, \tag{3.9}$$

where C *and* κ *are constants depending on* $|\psi|_\infty$ *and* r, *and the constant involved in* $\ll$ *does not depend on* (n, x).

Proof. Our proof is similar to that of [**28**, Theorem 6.1].

Let $(\varepsilon'_n)_{n\geq 1}$ be an independent copy of the innovations $(\varepsilon_n)_{n\geq 1}$, independent also of g_0. Fix $n \geq 1$ and $1 \leq q \leq n$. For $k \geq 0$, let

$$X'_k = \mathbb{E}_g(\psi(g_k, g_{k+1}, \ldots, g_{k+[q/2]}, (\tilde{g}_i)_{i\geq k+[q/2]+1})),$$

where $\mathbb{E}_g$ denotes the conditional expectation given $(g_n)_{n\geq 0}$, while $(\tilde{g}_i)_{i\geq k+[q/2]+1}$ is defined by $\tilde{g}_{k+[q/2]+1} = U(g_{k+[q/2]}, \varepsilon'_{k+[q/2]+1})$ and $\tilde{g}_{i+1} = U(\tilde{g}_i, \varepsilon'_{i+1})$ for $i > k + [q/2]$. The function U is given by (3.2).

Let

$$S'_n = \sum_{k=1}^{n} X'_k.$$

Observe that

$$\max_{k\leq n} |S_k| \leq \sum_{k=1}^{n} |X_k - X'_k| + \max_{1\leq k\leq n} |S'_k|.$$

Now, set $k_n = [n/q]$ and $U'_i = S'_{iq} - S'_{(i-1)q}$ for $1 \leq i \leq k_n$ and $U'_{k_n+1} = S'_n - S'_{k_n q}$. Since all integers j are on the distance of at most $[q/2]$ from $q\mathbb{N}$, we write

$$\begin{aligned}\max_{k\leq n} |S_k| \leq \sum_{k=1}^{n} |X_k - X'_k| + 2[q/2]|\psi|_\infty \\ + \max_{2j\leq k_n+1} \Big|\sum_{k=1}^{j} U'_{2k}\Big| + \max_{2j-1\leq k_n+1} \Big|\sum_{k=1}^{j} U'_{2k-1}\Big|.\end{aligned} \tag{3.10}$$

We shall now construct random variables $(U^*_i)_{1\leq i\leq k_n+1}$ such that (a) U^*_i has the same distribution as U'_i for all $1 \leq i \leq k_n + 1$, (b) the variables $(U^*_{2i})_{2\leq 2i\leq k_n+1}$ are independent as well as the random variables $(U^*_{2i-1})_{1\leq 2i-1\leq k_n+1}$ and (c) we can suitably control the quantity $\|U_i - U^*_i\|_1$.

This is done recursively as follows. Let $U^*_2 = U'_2$ and let us first construct U^*_4. With this aim, we note that

$$X'_k = h_q(g_k, g_{k+1}, \ldots, g_{k+[q/2]})$$

for some centered function h_q with $|h_q|_\infty \leq |\psi|_\infty$. Let $g^{(2)}_{2q+[q/2]}$ be a random variable in S with law ν and independent from $(g_0, (\varepsilon_k)_{k\geq 1})$ and define the Markov chain $(g^{(2)}_k)_{k\geq 2q+[q/2]}$ by

$$g^{(2)}_{k+1} = U(g^{(2)}_k, \varepsilon_{k+1}) \quad \text{for } k \geq 2q + [q/2].$$

Let

$$X^{(2)}_k = h_q(g^{(2)}_k, g^{(2)}_{k+1}, \ldots, g^{(2)}_{k+[q/2]}) \quad \text{for } k \geq 2q + [q/2]$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

and

$$U_4^* = \sum_{k=3q+1}^{4q} X_k^{(2)}.$$

It is clear that U_4^* is independent of U_2^* and equal to U_4' in law.

Now, for any $i \geq 3$, we define Markov chains $(g_k^{(i)})_{k\geq 2(i-1)q+[q/2]}$ in the following iterative way : $g_{2(i-1)q+[q/2]}^{(i)}$ is a random variable in S with law ν and independent from $(g_0, (\varepsilon_k)_{k\geq 1}, (g_{2(j-1)q+[q/2]}^{(j)})_{2\leq j<i})$ and we set

$$g_{k+1}^{(i)} = U(g_k^{(i)}, \varepsilon_{k+1}) \quad \text{for } k \geq 2(i-1)q + [q/2].$$

Next,

$$X_k^{(i)} = h_q(g_k^{(i)}, g_{k+1}^{(i)}, \ldots, g_{k+[q/2]}^{(i)}) \quad \text{for } k \geq 2(i-1)q + [q/2]$$

and

$$U_{2i}^* = \sum_{k=(2i-1)q+1}^{2iq} X_k^{(i)}.$$

It is clear that the so-constructed $(U_{2i}^*)_{2\leq 2i\leq k_n+1}$ are independent and that U_{2i}^* is equal in law to U_{2i}' for all i.

By stationarity, for all $1 \leq i \leq [(k_n+1)/2]$,

$$\|U_{2i}^* - U_{2i}'\|_1 \leq \|U_4^* - U_4'\|_1 \leq \sum_{k=3q+1}^{4q} \|X_k' - X_k^{(2)}\|_1.$$

But, by stationarity again,

$$\sum_{k=3q+1}^{4q} \|X_k - X_k^{(2)}\|_1 = \sum_{k=q-[q/2]+1}^{2q-[q/2]} \|h_q(g_k, g_{k+1}, \ldots, g_{k+[q/2]}) - h_q(g_k^*, g_{k+1}^*, \ldots, g_{k+[q/2]}^*)\|_1,$$

where $(g_k^*)_{k\geq 0}$ is the Markov chain defined in (3.3). Hence, for all $1 \leq i \leq [(k_n+1)/2]$,

$$\|U_{2i}^* - U_{2i}'\|_1 \leq 2|\psi|_\infty \sum_{k=q-[q/2]+1}^{2q-[q/2]} \mathbb{P}(T \geq k) \leq 2q|\psi|_\infty \mathbb{P}(T \geq [q/2]). \tag{3.11}$$

Similarly for the odd blocks, we can construct random variables $(U_{2i-1}^*)_{1\leq 2i-1\leq k_n+1}$, which are independent and such that U_{2i-1}^* equals in law to U_{2i-1}' for all i and

$$\|U_{2i-1}^* - U_{2i-1}'\|_1 \leq 2q|\psi|_\infty \mathbb{P}(T \geq [q/2]). \tag{3.12}$$

Overall, from (3.10), (3.11) and (3.12), we deduce that for all $x > 1$ and $1 \leq q \leq n$ such that $q|\psi|_\infty \leq x$,

$$\mathbb{P}\Big(\max_{k\leq n} |S_k| \geq 5x\Big) \leq x^{-1} \sum_{k=0}^{n-1} \|X_k - X_k'\|_1 + 2nx^{-1}|\psi|_\infty \mathbb{P}(T \geq [q/2]) + \mathbb{P}\Big(\max_{2j\leq k_n+1} \Big|\sum_{k=1}^{j} U_{2k}^*\Big| \geq x\Big) + \mathbb{P}\Big(\max_{2j-1\leq k_n+1} \Big|\sum_{k=1}^{j} U_{2k-1}^*\Big| \geq x\Big). \tag{3.13}$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

By Proposition 3.2, for all $\alpha \geq 1$,

$$\|X_k - X'_k\|_1 \ll q^{-\alpha/2} + \mathbb{P}(T \geq [q/2]/\alpha), \tag{3.14}$$

where the constant involved in $\ll$ does not depend on k or q. Using that $\|U^*_{2i}\|_\infty \leq q|\psi|_\infty$, we apply Bennet's inequality [**2**] and derive

$$\mathbb{P}\Big(\max_{2j\leq k_n+1}\Big|\sum_{k=1}^{j} U^*_{2k}\Big| \geq x\Big) \leq 2\exp\Big(-\frac{x}{2q|\psi|_\infty}\log(1+xq|\psi|_\infty/v_q)\Big),$$

where one can take v_q any real such that

$$v_q \geq \sum_{i=1}^{[(k_n+1)/2]} \|U^*_{2i}\|_2^2 = \sum_{i=1}^{[(k_n+1)/2]} \|U'_{2i}\|_2^2.$$

But, by stationarity,

$$\|U'_{2i}\|_2 = \|S'_q\|_2 \leq \|S_q\|_2 + (2|\psi|_\infty)^{1/2}\sum_{k=1}^{q}\|X_k - X'_k\|_1^{1/2}.$$

By Corollary 3.4, $\|S_q\|_2^2 \ll q$. Since $n\mathbb{P}(T \geq n) \ll 1$, we infer that

$$\sum_{k=1}^{q}\|X_k - X'_k\|_1^{1/2} \ll q^{1/2}.$$

Therefore, $\|U'_{2i}\|_2^2 \ll q$. Hence, taking $v_q = n/\kappa'$ where κ' is a sufficiently small positive constant not depending on x, n and q, we get

$$\mathbb{P}\Big(\max_{2j\leq k_n+1}\Big|\sum_{k=1}^{j} U^*_{2k}\Big| \geq x\Big) \leq 2\exp\Big(-\frac{x}{2q|\psi|_\infty}\log(1+\kappa' xq|\psi|_\infty/n)\Big). \tag{3.15}$$

It follows from (3.13)–(3.15), that for all $\alpha \geq 1$, $x > 0$ and $1 \leq q < n$ with $q|\psi|_\infty \leq x$,

$$\mathbb{P}\Big(\max_{k\leq n}|S_k| \geq 5x\Big) \ll nx^{-1}(q^{-\alpha/2} + \mathbb{P}(T \geq [q/2]/\alpha)) + \exp\Big(-\frac{x}{2q|\psi|_\infty}\log(1+\kappa' xq|\psi|_\infty/n)\Big).$$

Let now $r \geq 1$. Then, for $x \in [r|\psi|_\infty, n|\psi|_\infty/5]$, choose $q = [x/(r|\psi|_\infty)]$ and $\alpha = 2r$ in the previous inequality and the result follows. To end the proof, note that if $x > n|\psi|_\infty/5$, the deviation probability obviously equals zero and if $0 < x < r|\psi|_\infty$, the inequality follows easily from Markov's inequality at order 1. □

The following Rosenthal-type inequality relates T to the moments of S_n.

PROPOSITION 3.6. *Assume that* $\mathbb{E}(T) < \infty$. *Then, for each* $p \geq 2$, *there exist* κ_1, κ_2, $\kappa_3 > 0$ *such that for all* $n \geq 1$,

$$\mathbb{E}\Big(\max_{k\leq n}|S_k|^p\Big) \leq \kappa_1 n^{p/2} + \kappa_2 n\sum_{i=1}^{[\kappa_3 n]} i^{p-2}\mathbb{P}(T \geq i).$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

Proof. Write

$$\mathbb{E}\Big(\max_{k\le n}|S_k|^p\Big) = p5^p\int_0^{n|\psi|_\infty/5} x^{p-1}\mathbb{P}\Big(\max_{k\le n}|S_k|>5x\Big)\,dx. \tag{3.16}$$

Using Lemma 3.5 with $r = p+1$, we get that for $p \ge 2$,

$$\int_{r|\psi|_\infty}^{n|\psi|_\infty/5} x^{p-1}\mathbb{P}\Big(\max_{k\le n}|S_k|\ge 5x\Big)\,dx \ll n^{p/2} + n\int_{r|\psi|_\infty}^{n|\psi|_\infty/5} x^{p-2}\mathbb{P}(T\ge Cx)\,dx.$$

Together with (3.16), the above implies that for any $p \ge 2$,

$$\mathbb{E}\Big(\max_{k\le n}|S_k|^p\Big) \ll n^{p/2} + n\int_0^{Cn|\psi|_\infty/5} x^{p-2}\mathbb{P}(T\ge x)\,dx,$$

where the constant involved in $\ll$ depends on p but not on n. The result follows. □

4. *Proof of Theorem 1.3*

4.1. *Outline.* Let $g_0, g_1, \ldots$ be the stationary Markov chain constructed in §2. Suppose that $\psi : \Omega \to \mathbb{R}$ is a Hölder continuous observable with $\int \psi\, d\mathbb{P}_\Omega = 0$. Let

$$X_n = \psi\circ\sigma^n = \psi(g_n, g_{n+1}, \ldots) \quad\text{and}\quad S_n = \sum_{k=1}^n X_k.$$

By Corollary 2.5, the proof of Theorem 1.3 reduces to proving ASIP with the same rates for the process $(S_n)_{n\ge1}$. This is the aim of this section. Our strategy is to adapt the argument in [**4**].

Remark 4.1. We restrict to the case when the variance c^2, given by (1.4), is positive. The case $c^2 = 0$ requires a different approach, and it is addressed by Remark 1.5.

The Markov chain $(g_n)_{n\ge0}$ behaves similarly to the Markov chain $(W_n)_{n\ge0}$ on the state space $\mathbb{N}$, studied in [**4**, §3.3.1]. Let us briefly recall [**4**, Corollary 5]: for any bounded and centered function $h : \mathbb{N}\to\mathbb{R}$, the process $(\sum_{k=1}^n h(W_k))_{n\ge1}$ satisfies the ASIP with rate $o(n^{1/p})$, $p>2$, provided that $\sum_{k\ge1} k^{p-2}\mathbb{P}(T\ge k) < \infty$, where ν is the stationary distribution of $(W_n)_{n\in\mathbb{N}}$ and T is the meeting time of the Markov chain.

Remark 4.2. By [**4**, Proposition 15], the condition $\sum_{k\ge1} k^{p-2}\mathbb{P}(T\ge k) < \infty$ is sharp to get the rate $o(n^{1/p})$ in the ASIP.

The strategy of [**4**] was to adapt the method of Berkes, Liu and Wu [**3**] for functions of a sequence of iid random variables to instantaneous functions of Markov chains, in order to obtain sufficient conditions for the ASIP with rate $o(n^{1/p})$ in terms of an $\mathbb{L}^1$-coupling coefficient. For the Markov chain $(W_n)_{n\in\mathbb{N}}$, this $\mathbb{L}^1$-coupling condition can be obtained from the tails of the meeting time. More precisely, writing $W_n = F(W_{n-1}, \varepsilon_n)$, where F is a measurable function and $(\varepsilon_n)_{n\ge1}$ is a sequence of iid random variables independent of W_0, following [**3**], the approach developed in [**4**] is first to approximate the partial sums by a sum of big blocks and then to approximate the random variables in the kth big block by m_k dependent random variables with the help of conditional expectations (here m_k is much less than the size of the big block). This approximation can be then evaluated using

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

the tail of the meeting time which is defined as the first time when two chains starting from two independent random variables and sharing the same innovations meet and then glue. Next, as in [**3**], each big block is divided in smaller blocks of size $3m_k$, which become independent once we work conditionally to a fixed part of the sequence $(\varepsilon_n)_{n\geq 1}$ of innovations. Then, conditionally to this fixed part of the sequence $(\varepsilon_n)_{n\geq 1}$, it is possible to couple these small independent blocks with independent gaussian random variables and to evaluate the error between their partial sums by using [**29**, Theorem 1] (note that once a part of the innovations is fixed, the underlying random variables do not have the same law anymore). The error in the approximation depends once again of the tail of the meeting time. Next, as in [**3**], it remains to proceed to an unconditional gaussian approximation.

The main difference between our situation and the one considered in [**4**] is that X_n's are functions not only of g_n, but of the whole future $g_n, g_{n+1}, \ldots$ However, using the regularity of our observables, we shall see that it is possible to approximate X_n by a measurable function of a finite number of coordinates. Then the proof in [**3**] can be adapted also to our situation, and the rate in the ASIP is, as in [**4**], related to the tail of the meeting time of the chain $(g_n)_{n\geq 0}$ (see §3).

4.2. *The proof.* Let c^2 be given by (1.4). From Corollaries 2.5 and 3.4, $c^2 = \lim_{n\to\infty} n^{-1}\|S_n\|_2^2 = \|X_0\|_2^2 + 2\sum_{n=1}^{\infty} \mathrm{Cov}\,(X_0, X_n)$. If the process $(S_n)_{n\geq 0}$ satisfies the ASIP, this has to be the variance of the limiting Brownian motion. Recall that we suppose that $c^2 > 0$.

All along the proof, we set $\beta = 1/\gamma$ (so $\beta > 2$ since $\gamma < 1/2$), and η will designate a constant, which is equal either to 1 in the case of the map (1.1) or to 0 in the case of the map (1.5).

It suffices to prove the following strong approximation: one can redefine $(S_n)_{n\geq 1}$ without changing its distribution on a probability space (possibly richer than $(\Omega, \mathbb{P}_\Omega)$) on which there exists a sequence $(N_i)_{i\geq 1}$ of iid centered Gaussian random variables with variance c^2 such that for all $\kappa > 1/\beta$,

$$\sup_{k\leq n}\left|S_n - \sum_{i=1}^{k} N_i\right| = o(n^{1/\beta}(\log n)^{\eta\kappa}) \quad \text{a.s.} \tag{4.1}$$

The proof of (4.1) is divided in several steps. Throughout, we use the notation $b_n = \lceil (\log n)/(\log 3)\rceil$ for $n \geq 2$ (so that b_n is the unique integer such that $3^{b_n-1} < n \leq 3^{b_n}$), and fix $\kappa > 1/\beta$.

Step 1. For $\ell \geq 0$, let

$$m_\ell = [3^{\ell/\beta}\ell^{\eta\kappa}] \tag{4.2}$$

and define, for $k \geq 0$,

$$X_{\ell,k} = \mathbb{E}_g(\psi(g_k, g_{k+1}, \ldots, g_{k+m_\ell}, (\tilde{g}_i)_{i\geq k+m_\ell+1})),$$

where $\mathbb{E}_g$ denotes the conditional expectation given $g := (g_n)_{n\geq 0}$. Here $(\tilde{g}_i)_{i\geq k+m_\ell+1}$ is defined as follows: $\tilde{g}_{k+m_\ell+1} = U(g_{k+m_\ell}, \varepsilon'_{k+m_\ell+1})$ and $\tilde{g}_{i+1} = U(\tilde{g}_i, \varepsilon'_{i+1})$ for any $i > k + m_\ell$, where $(\varepsilon'_i)_{i\geq 1}$ is an independent copy of $(\varepsilon_i)_{i\geq 1}$, independent of g_0, and U

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

is given by (3.2). Note that the $X_{\ell,k}$'s are centered. Define

$$W_{\ell,i} = \sum_{k=1+3^{\ell-1}}^{i+3^{\ell-1}} X_k, \quad \overline{W}_{\ell,i} = \sum_{k=1+3^{\ell-1}}^{i+3^{\ell-1}} X_{\ell,k} \quad \text{and} \quad W'_{\ell,i} = W_{\ell,i} - \overline{W}_{\ell,i}.$$

The first step is to prove that

$$\sum_{\ell=1}^{b_n-1} W'_{\ell,3^\ell-3^{\ell-1}} + W'_{b_n,n-3^{b_n-1}} = o(n^{1/\beta}(\log n)^{\eta\kappa}) \quad \text{a.s.} \tag{4.3}$$

This will hold provided that for all $\varepsilon > 0$,

$$\sum_{j\geq 1} \mathbb{P}\Bigg(\sum_{\ell=1}^{j} \sum_{k=3^{\ell-1}+1}^{3^\ell} |X_k - X_{\ell,k}| > \varepsilon 3^{j/\beta} j^{\eta\kappa} \Bigg) < \infty. \tag{4.4}$$

By Proposition 3.2, for all $k \geq 0$, $\ell \geq 1$ and $r \geq 1$,

$$\|X_k - X_{\ell,k}\|_1 \ll m_\ell^{-r/2} + \mathbb{P}(T \geq [m_\ell/r]), \tag{4.5}$$

where the constant involved in $\ll$ does not depend on k and ℓ. By Markov inequality at order 1, for all $\varepsilon > 0$ and $r \geq 1$,

$$\begin{aligned} &\sum_{j\geq 1} \mathbb{P}\Bigg(\sum_{\ell=1}^{j} \sum_{k=3^{\ell-1}+1}^{3^\ell} |X_k - X_{\ell,k}| > \varepsilon 3^{j/\beta} j^{\eta\kappa} \Bigg) \\ &\quad \ll \sum_{j\geq 1} \frac{1}{\varepsilon 3^{j/\beta} j^{\eta\kappa}} \sum_{\ell=1}^{j} 3^\ell m_\ell^{-r/2} + \sum_{j\geq 1} \frac{1}{\varepsilon 3^{j/\beta} j^{\eta\kappa}} \sum_{\ell=1}^{j} 3^\ell \mathbb{P}(T \geq [m_\ell/r]). \end{aligned}$$

Taking into account the fact that $m_\ell = [3^{\ell/\beta} \ell^{\eta\kappa}]$, the first term in the right-hand side is finite provided we take $r > 2(\beta - 1)$ whereas, by a change of variables, we have, for any $r \geq 1$,

$$\sum_{j\geq 1} \frac{1}{3^{j/\beta} j^{\eta\kappa}} \sum_{\ell=1}^{j} 3^\ell \mathbb{P}(T \geq [m_\ell/r]) \leq C \sum_{n\geq 2} \frac{n^{\beta-2}}{(\log n)^{\eta\kappa\beta}} \mathbb{P}(T \geq n) \tag{4.6}$$

where C is a constant depending on r, β, κ and η. In the case of the map (1.1), $\eta = 1$ and the series above converge iff $\mathbb{E}(\widetilde{\psi}_{\beta,\kappa\beta}(T)) < \infty$, which holds by Lemma 3.1(a) and the fact that $\kappa\beta > 1$. Now in the case of the map (1.5), $\eta = 0$ and then, again from Lemma 3.1, the series above converges since $\mathbb{E}(T^{\beta-1}) < \infty$. It follows that (4.4) is satisfied and then (4.3) holds.

This completes the proof of step 1.

Step 2. Let

$$\tilde{X}_{\ell,k} = \mathbb{E}(X_{\ell,k} | \varepsilon_{k-m_\ell}, \ldots, \varepsilon_{k+m_\ell}). \tag{4.7}$$

Let $\widetilde{W}_{\ell,i} = \sum_{k=1+3^{\ell-1}}^{i+3^{\ell-1}} \tilde{X}_{\ell,k}$ and $W''_{\ell,i} = \overline{W}_{\ell,i} - \widetilde{W}_{\ell,i}$. The second step consists of proving that

$$\sum_{\ell=1}^{b_n-1} W''_{\ell,3^\ell-3^{\ell-1}} + W''_{b_n,n-3^{b_n-1}} = o(n^{1/\beta}(\log n)^{\eta\kappa}) \quad \text{a.s.} \tag{4.8}$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

Clearly, (4.8) will follow from the Kronecker lemma, if one can prove that

$$\sum_{\ell\geq 1}\frac{1}{3^{\ell/\beta}\ell^{\eta\kappa}}\sum_{k=3^{\ell-1}+1}^{3^\ell}\|X_{\ell,k}-\tilde{X}_{\ell,k}\|_1<\infty. \tag{4.9}$$

We claim that

$$\|X_{\ell,k}-\tilde{X}_{\ell,k}\|_1\leq 2|\psi|_\infty\mathbb{P}(T\geq m_\ell). \tag{4.10}$$

Then, using (4.10),

$$\sum_{\ell\geq 1}\frac{1}{3^{\ell/\beta}\ell^{\eta\kappa}}\sum_{k=3^{\ell-1}+1}^{3^\ell}\|X_{\ell,k}-\tilde{X}_{\ell,k}\|_1\leq 2|\psi|_\infty\sum_{\ell\geq 1}\frac{3^\ell}{3^{\ell/\beta}\ell^{\eta\kappa}}\mathbb{P}(T\geq m_\ell).$$

Therefore, (4.9) holds by using (4.6) and Lemma 3.1 (as quoted right after (4.6)).

It remains to prove the claim (4.10). This follows closely the proof of [**4**, Lemma 24]. Indeed, we can write

$$X_{\ell,k}=h_\ell(g_k,g_{k+1},\dots,g_{k+m_\ell}),$$

where h_ℓ is a measurable function such that $|h_\ell|_\infty\leq|\psi|_\infty$ and $\mathbb{P}_\Omega(h_\ell)=0$. Hence

$$\begin{aligned}X_{\ell,k}-\tilde{X}_{\ell,k}&=h_\ell(g_k,g_{k+1},\dots,g_{k+m_\ell})\\&\quad-\mathbb{E}(h_\ell(g_k,g_{k+1},\dots,g_{k+m_\ell})|\varepsilon_{k-m_\ell},\dots,\varepsilon_{k+m_\ell}).\end{aligned}$$

Recall that for all $k\geq 1$, $g_k=U(g_{k-1},\varepsilon_k)$ where U is a measurable function from $S\times\mathcal{A}$ to S. For any $i\geq 1$, let then U_i be the function from $S\times\mathcal{A}^{\otimes i}$ to S defined in the following iterative way:

$$U_1=U\quad\text{and}\quad U_i(a,x_1,x_2,\dots,x_i)=U_{i-1}(U(a,x_1),x_2,\dots,x_i),\quad i\geq 2.$$

Then for all $i\geq 0$ and $k\geq m_\ell+1$,

$$g_{k+i}=U_{i+m_\ell+1}(g_{k-m_\ell-1},\varepsilon_{k-m_\ell},\dots,\varepsilon_{k+i}).$$

Hence,

$$\begin{aligned}&h_\ell(g_k,g_{k+1},\dots,g_{k+m_\ell})\\&\quad=h_\ell(U_{m_\ell+1}(g_{k-m_\ell-1},\varepsilon_{k-m_\ell},\dots,\varepsilon_k),\dots,U_{2m_\ell+1}(g_{k-m_\ell-1},\varepsilon_{k-m_\ell},\dots,\varepsilon_{k+m_\ell}))\\&\quad=:H_{\ell,m_\ell}(g_{k-m_\ell-1},\varepsilon_{k-m_\ell},\dots,\varepsilon_{k+m_\ell}).\end{aligned}$$

Let now $(\varepsilon'_k)_{k\geq 1}$ be an independent copy of $(\varepsilon_k)_{k\geq 1}$, independent of g_0. Let g'_0 be a random variable in S with distribution ν and independent from $(g_0,(\varepsilon_k)_{k\geq 1},(\varepsilon'_k)_{k\geq 1})$. Define a Markov chain $(g'_n)_{n\geq 0}$ by

$$g'_{n+1}=U(g'_n,\varepsilon'_{n+1})\quad\text{for } n\geq 0.$$

Denoting $V_{k,m_\ell}=(g_0,\varepsilon_1,\dots,\varepsilon_{k+m_\ell})$ and $\mathbb{E}_{V_{k,m_\ell}}(\cdot)=\mathbb{E}(\cdot|V_{k,m_\ell})$, we have

$$\begin{aligned}X_{\ell,k}-\tilde{X}_{\ell,k}&=\mathbb{E}_{V_{k,m_\ell}}(H_{\ell,m_\ell}(g_{k-m_\ell-1},\varepsilon_{k-m_\ell},\dots,\varepsilon_{k+m_\ell}))\\&\quad-\mathbb{E}_{V_{k,m_\ell}}(H_{\ell,m_\ell}(g'_{k-m_\ell-1},\varepsilon_{k-m_\ell},\dots,\varepsilon_{k+m_\ell})).\end{aligned}$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

Hence, using the stationarity,

$$\begin{aligned}\|X_{\ell,k}-\tilde{X}_{\ell,k}\|_1 &\le \|H_{\ell,m_\ell}(g_{k-m_\ell-1},\varepsilon_{k-m_\ell},\dots,\varepsilon_{k+m_\ell})\\ &\quad - H_{\ell,m_\ell}(g'_{k-m_\ell-1},\varepsilon_{k-m_\ell},\dots,\varepsilon_{k+m_\ell})\|_1\\ &= \|H_{\ell,m_\ell}(g_0,\varepsilon_1,\dots,\varepsilon_{2m_\ell+1}) - H_{\ell,m_\ell}(g'_0,\varepsilon_1,\dots,\varepsilon_{2m_\ell+1})\|_1.\end{aligned}$$

Let $(g_n^*)_{n\ge 0}$ be the Markov chain in the definition of the meeting time, see (3.3). Then

$$\begin{aligned}\|X_{\ell,k}-\tilde{X}_{\ell,k}\|_1 &\le \|H_{\ell,m_\ell}(g_0,\varepsilon_1,\dots,\varepsilon_{2m_\ell+1}) - H_{\ell,m_\ell}(g_0^*,\varepsilon_1,\dots,\varepsilon_{2m_\ell+1})\|_1\\ &= \|h_\ell(g_{m_\ell+1},g_{m_\ell+2},\dots,g_{2m_\ell+1}) - h_\ell(g^*_{m_\ell+1},g^*_{m_\ell+2},\dots,g^*_{2m_\ell+1})\|_1.\end{aligned}$$

Recall that for every $k\ge T$, $g_k=g_k^*$. Therefore,

$$\|X_{\ell,k}-\tilde{X}_{\ell,k}\|_1 \le 2|h_\ell|_\infty \mathbb{P}(T\ge m_\ell),$$

proving (4.10). This ends the proof of step 2.

Step 3. Setting $\tilde{S}_n := \sum_{\ell=1}^{b_n-1}\widetilde{W}_{\ell,3^\ell-3^{\ell-1}} + \widetilde{W}_{b_n,n-3^{b_n-1}}$, the rest of the proof consists in showing that, enlarging the underlying probability space if necessary, there exists a sequence $(N_i)_{i\ge 1}$ of iid centered Gaussian random variables with variance c^2 such that

$$\sup_{k\le n}\left|\tilde{S}_k - \sum_{i=1}^{k} N_i\right| = o(n^\gamma(\log n)^{\eta\kappa}) \quad \text{a.s.} \tag{4.11}$$

This can be achieved using the method of [**3**]. Indeed the constructed $\tilde{X}_{\ell,k}$ can be rewritten as

$$\tilde{X}_{\ell,k} := G_\ell(\varepsilon_{k-m_\ell},\dots,\varepsilon_{k+m_\ell}),$$

where G_ℓ is a measurable function. So $\tilde{X}_{\ell,k}$ is a measurable function of $(\varepsilon_{k-m_\ell},\dots,\varepsilon_{k+m_\ell})$ instead of $(\varepsilon_{k-m_\ell},\dots,\varepsilon_k)$ as in [**3**]. However, this difference can be handled by only minor adjustments, mainly taking $2m_\ell$ instead of m_ℓ in [**3**]. More precisely, the blocks $B_{\ell,j}$ in [**3**] can be defined as follows: for $\ell\ge k_0 := \inf\{k\ge 1 : m_k \le 4^{-1}3^{k-2}\}$ and $j=1,\dots,q_\ell := \lceil 3^{\ell-2}/m_\ell\rceil - 2$,

$$B_{\ell,j} = \sum_{i=1+(6j-1)m_\ell}^{(6j+5)m_\ell} \tilde{X}_{\ell,i+m_\ell+3^{\ell-1}}.$$

Define, for $j\ge 1$,

$$\mathcal{J}_{\ell,j} = \{3^{\ell-1}+(6j-1)m_\ell+k,\ k=1,2,\dots,2m_\ell\},$$

$$\mathbf{U}_{\ell,j} = (\varepsilon_i,\ i\in\mathcal{J}_{\ell,j}) \quad \text{and} \quad \mathbf{U} = (\mathbf{U}_{\ell,j},\ j=1,\dots,q_\ell+1)_{\ell=k_0}^{\infty}.$$

Then

$$\begin{aligned}B_{\ell,j} &= \sum_{i=1+(6j-1)m_\ell}^{(6j+1)m_\ell} \tilde{X}_{\ell,i+m_\ell+3^{\ell-1}} + \sum_{i=1+(6j+1)m_\ell}^{(6j+3)m_\ell} \tilde{X}_{\ell,i+m_\ell+3^{\ell-1}}\\ &\quad + \sum_{i=1+(6j+3)m_\ell}^{(6j+5)m_\ell} \tilde{X}_{\ell,i+m_\ell+3^{\ell-1}}\\ &:= H_\ell(\mathbf{U}_{\ell,j},\{\varepsilon_{i+3^{\ell-1}}\}_{1+(6j+1)m_\ell\le i\le(6j+5)m_\ell},\mathbf{U}_{\ell,j+1}).\end{aligned}$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

On the set $\{\mathbf{U}=\mathbf{u}\}$, $(B_{\ell,j}(\mathbf{u}))_{j=1,\dots,q_\ell}$ are then independent between them. Then, following [**3**], we use Sakhanenko's strong approximation [**29**] to get a bound for the approximation error between $\tilde{S}_n(\mathbf{u})$ and a Wiener process with variance depending on $\mathbf{u}$. To get the unconditional ASIP, we use the arguments given in [**3**, step 3.4]. So, as it is summarized in [**4**, Proposition 21], we infer that (4.11) will follow if one can prove that there exists $r \in (2,\infty)$ such that

$$\sum_{\ell\ge k_0} \frac{3^\ell}{3^{\ell r/\beta}\ell^{\eta\kappa r}m_\ell}\mathbb{E}\Big(\max_{1\le k\le 6m_\ell}|\widetilde{W}_{\ell,k}|^r\Big)<\infty, \tag{4.12}$$

and

$$3^\ell(v_\ell^{1/2}-c)^2 = o(3^{2\ell/\beta}\ell^{2\eta\kappa}(\log\ell)^{-1}) \quad \text{as } \ell\to\infty, \tag{4.13}$$

where

$$v_\ell = (2m_\ell)^{-1}\big\{\mathbb{E}(\widetilde{W}^2_{\ell,2m_\ell}) + 2\mathbb{E}(\widetilde{W}_{\ell,2m_\ell}(\widetilde{W}_{\ell,4m_\ell}-\widetilde{W}_{\ell,2m_\ell}))\big\}. \tag{4.14}$$

To end the proof, it remains to prove the two conditions above. We start with (4.12). Note first that for all $r\ge 1$,

$$\Big\|\max_{1\le k\le 6m_\ell}|W_k-\widetilde{W}_{\ell,k}|\Big\|_r \le \sum_{k=1+3^{\ell-1}}^{6m_\ell+3^{\ell-1}}\|X_k-\tilde{X}_{\ell,k}\|_r.$$

Using that $\|X_k\|_\infty\le|\psi|_\infty$ and $\|\tilde{X}_{\ell,k}\|_\infty\le 2|\psi|_\infty$, we get

$$\begin{aligned}&\Big\|\max_{1\le k\le 6m_\ell}|W_k-\widetilde{W}_{\ell,k}|\Big\|_r\\ &\quad\le (3|\psi|_\infty)^{(r-1)/r}\sum_{k=1+3^{\ell-1}}^{6m_\ell+3^{\ell-1}}(\|X_k-X_{\ell,k}\|_1^{1/r}+\|X_{\ell,k}-\tilde{X}_{\ell,k}\|_1^{1/r}).\end{aligned}$$

But according to (4.5) and (4.10), for all $\alpha\ge 1$,

$$\|X_k-\tilde{X}_{\ell,k}\|_1 \ll m_\ell^{-\alpha/2}+\mathbb{P}(T\ge[m_\ell/\alpha]), \tag{4.15}$$

where the constant involved in $\ll$ does not depend on k and ℓ. Therefore, for all $r\ge 1$ and $\alpha\ge 1$,

$$\begin{aligned}&\sum_{\ell\ge k_0}\frac{3^\ell}{3^{\ell r/\beta}\ell^{\eta\kappa r}m_\ell}\mathbb{E}\Big(\max_{1\le k\le 6m_\ell}|W_k-\widetilde{W}_{\ell,k}|^r\Big)\\ &\quad\ll\sum_{\ell\ge k_0}\frac{3^\ell m_\ell^r}{3^{\ell r/\beta}\ell^{\eta\kappa r}m_\ell}\big(m_\ell^{-\alpha/2}+\mathbb{P}(T\ge[m_\ell/\alpha])\big)\\ &\quad\ll\sum_{\ell\ge k_0}\frac{3^{\ell(\beta-1)/\beta}}{\ell^{\eta\kappa}}3^{-\alpha\ell/(2\beta)}\ell^{-\alpha\eta\kappa/2}+\sum_{\ell\ge k_0}\frac{3^{\ell(\beta-1)/\beta}}{\ell^{\eta\kappa}}\mathbb{P}(T\ge 3^{\ell/\beta}\ell^{\eta\kappa}/\alpha).\end{aligned}$$

The first term in the right-hand side is finite provided that we take $\alpha>2(\beta-1)$ whereas, the second series converge for any $\alpha\ge 1$, by using once again (4.6) and Lemma 3.1. Therefore, to prove (4.12), it suffices to show that there exists $r\in\,]2,\infty[$ such that

$$\sum_{\ell\ge k_0}\frac{3^\ell}{3^{\ell r/\beta}\ell^{\eta\kappa r}m_\ell}\mathbb{E}\Big(\max_{1\le k\le 6m_\ell}|W_k|^r\Big)<\infty. \tag{4.16}$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

By Lemma 3.1, $\mathbb{E}(T) < \infty$ since $\beta > 2$ for both maps. Using stationarity and Proposition 3.6, we get that for any $r \geq 2$,

$$\sum_{\ell \geq k_0} \frac{3^\ell}{3^{\ell r/\beta} \ell^{\eta\kappa r} m_\ell} \mathbb{E}\Big(\max_{1 \leq k \leq 6m_\ell} |W_k|^r \Big)$$
$$\ll \sum_{\ell \geq k_0} \frac{3^\ell}{3^{\ell r/\beta} \ell^{\eta\kappa r}} m_\ell^{r/2-1} + \sum_{\ell \geq k_0} \frac{3^\ell}{3^{\ell r/\beta} \ell^{\eta\kappa r}} \sum_{i=1}^{[6\kappa_3 m_\ell]} i^{r-2} \mathbb{P}(T \geq i).$$

Since $m_\ell = [3^{\ell/\beta} \ell^{\eta\kappa}]$, the first term of the right-hand side is finite provided that we take $r > 2(\beta - 1)$. To control the second term, we note that for any $r > \beta$, by a change of variables,

$$\sum_{\ell \geq k_0} \frac{3^\ell}{3^{\ell r/\beta} \ell^{(1+\eta)r}} \sum_{i=1}^{[6\kappa_3 m_\ell]} i^{r-2} \mathbb{P}(T \geq i) \ll \sum_{i \geq 2} \frac{i^{\beta-2}}{(\log i)^{\eta\kappa r}} \mathbb{P}(T \geq i),$$

which is finite by Lemma 3.1 as it was quoted after (4.6). So, provided that we take $r > 2(\beta - 1)$, since $\beta > 2$, (4.16) holds (and then (4.12)).

We turn now to the proof of (4.13). Proceeding as to get the relation [**4**, (66)], we have

$$v_\ell = \tilde{c}_{\ell,0} + 2 \sum_{k=1}^{2m_\ell} \tilde{c}_{\ell,k},$$

where, for any $i \geq 0$,

$$\tilde{c}_{\ell,i} = \mathrm{Cov}\,(\tilde{X}_{\ell,m_\ell+1}, \tilde{X}_{\ell,i+m_\ell+1}).$$

Note also that since c^2 is assumed to be positive, to prove (4.13), it suffices to prove that

$$3^\ell (v_\ell - c^2)^2 = o(3^{2\ell/\beta} \ell^{2\eta\kappa} (\log \ell)^{-1}) \quad \text{as } \ell \to \infty. \tag{4.17}$$

To show that (4.17) is satisfied, we first note that, by stationarity, for all $i \geq 0$,

$$\begin{aligned} \big|\tilde{c}_{\ell,i} - \mathrm{Cov}\,(X_0, X_i)\big| &= \big|\mathrm{Cov}\,(\tilde{X}_{\ell,m_\ell+1} - X_{m_\ell+1}, \tilde{X}_{\ell,i+m_\ell+1}) \\ &\quad + \mathrm{Cov}\,(X_{m_\ell+1}, \tilde{X}_{\ell,i+m_\ell+1} - X_{i+m_\ell+1})\big| \\ &\leq 2|\psi|_\infty (\|\tilde{X}_{\ell,m_\ell+1} - X_{m_\ell+1}\|_1 + \|\tilde{X}_{\ell,i+m_\ell+1} - X_{i+m_\ell+1}\|_1). \end{aligned}$$

Let $\alpha \geq 1$. Then, according to (4.15), for all $i \geq 0$,

$$\big|\tilde{c}_{\ell,i} - \mathrm{Cov}\,(X_0, X_i)\big| \ll m_\ell^{-\alpha/2} + \mathbb{P}(T \geq [m_\ell/\alpha]).$$

It follows that

$$|v_\ell - c^2| \ll m_\ell^{1-\alpha/2} + m_\ell \mathbb{P}(T \geq [m_\ell/\alpha]) + 2 \sum_{i > 2m_\ell} \big|\mathrm{Cov}\,(X_0, X_i)\big|.$$

Recall that $\beta > 2$. By Lemma 3.1 (since $\kappa\beta > 1$),

$$\mathbb{P}(T \geq n) = o((\log n)^{\eta\kappa\beta} n^{1-\beta}) \quad \text{as } n \to \infty.$$

Using, in addition, Lemma 3.3, we derive that for all $\alpha \geq 1$,

$$|v_\ell - c^2| \ll 3^{\ell(2-\alpha)/(2\beta)} \ell^{\eta\kappa(2-\alpha)/2} + o(3^{\ell(2-\beta)/\beta} \ell^{2\eta\kappa}),$$

proving (4.17) (and then (4.13)) using the fact that $\beta > 2$ and taking $\alpha \geq 2\beta - 2$. This ends the proof of Theorem 1.3 when $c^2 > 0$.

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

4.3. *Extension to other observables.* As already mentioned in Remark 1.6, it is possible to relax the Hölder continuity assumption. For instance, if $m \geq 1$ is an integer, assume that φ is Hölder on the interior of $Y_{a_0 \cdots a_{m-1}}$ for every $a_0, \ldots, a_{m-1} \in \alpha$. Denote by $\alpha(a_0, \ldots, a_{m-1})$ the corresponding Hölder exponent and by $|\varphi|_{\alpha(a_0,\ldots,a_{m-1})}$ the corresponding Hölder norm. Assume further that $\alpha^* := \inf_{a_0,\ldots,a_{m-1} \in \alpha} \alpha(a_0, \ldots, a_{m-1}) > 0$ and that $|\varphi|_{\alpha *} := \sup_{a_0,\ldots,a_{m-1} \in \alpha} |\varphi|_{\alpha(a_0,\ldots,a_{m-1})} < \infty$. Under the above assumptions, the conclusion of Theorem 1.3 holds.

Let us briefly give the arguments explaining why such an extension is possible. We just give the necessary arguments to prove the estimate (4.5) (or more generally Proposition 3.2). Similar arguments may be used at each place where the Hölder property has been used to get similar estimates as (4.5). To do so one has to bound

$$|\psi(g_0, \ldots, g_n, g_{n+1}, \ldots) - \psi(g_0, \ldots, g_n, (\tilde{g}_k)_{k \geq n+1}, \ldots)| \tag{4.18}$$

If $\#\{k \leq n : g_k \in S_0\} < m$ we bound (4.18) by $2|\varphi|_\infty$.

Assume now that $\#\{k \leq n : g_k \in S_0\} \geq m$. Set $g_0 = (w_0, \ell_0)$. Assume that we can write that $w_0 = ww'$ with $h(w) = \ell_0$ and w may be an empty word (in which case $\ell_0 = 0$). Hence, $\pi(g_0, \ldots, g_n, g_{n+1}, \ldots)$ and $\pi(g_0, \ldots, g_n, (\tilde{g}_k)_{k \geq n+1}, \ldots)$ belong to Y and even, since $\#\{k \leq n : g_k \in S_0\} \geq m$, to some $Y_{a_0 \cdots a_{m-1}}$ (on which φ is Hölder). In particular one may bound (4.18) by $|\varphi|_{\alpha *} \lambda^{-\alpha * \#\{k \leq n : g_k \in S_0\}}$.

If w_0 cannot be written as above then, $\pi(g_0, \ldots, g_n, g_{n+1}, \ldots)$ and $\pi(g_0, \ldots, g_n, (\tilde{g}_k)_{k \geq n+1}, \ldots)$ belongs to $[0, 1/2)$ and we infer a similar bound.

So at the end, there exists $C > 0$ depending on $|\varphi|_\infty$ and $|\varphi|_{\alpha *}$, such that

$$|\psi(g_0, \ldots, g_n, g_{n+1}, \ldots) - \psi(g_0, \ldots, g_n, (\tilde{g}_k)_{k \geq n+1}, \ldots)| \leq C\theta^{\#\{k \leq n : g_k \in S_0\} - m}.$$

The end of the proof of Proposition 3.2 remains unchanged.

5. *Non-uniformly expanding dynamical systems*

We stated and proved Theorem 1.3 for two particular families of maps. In this section we extend our result to the class of non-uniformly expanding systems which admit inducing schemes as in Young [**32**] with polynomially decaying tails of return times.

5.1. *Non-uniformly expanding maps.* Let X be a complete bounded separable metric space with the Borel σ-algebra. Suppose that $f : X \to X$ is a measurable transformation which admits an inducing scheme consisting of:

- a closed subset Y of X with a *reference* probability measure m on Y;
- a finite or countable partition α of Y (up to a zero measure set) with $m(a) > 0$ for all $a \in \alpha$;
- an integrable *return time* function $\tau : Y \to \{1, 2, \ldots\}$ which is constant on each $a \in \alpha$ with value $\tau(a)$ and $f^{\tau(a)}(y) \in Y$ for all $y \in a$, $a \in \alpha$. (We do not require that τ is the *first* return time to Y.)

Define $F : Y \to Y$, $F(y) = f^{\tau(y)}(y)$. We assume that there are constants $\kappa > 1$, $K > 0$ and $\eta \in (0, 1]$ such that for each $a \in \alpha$ and all $x, y \in a$:

- F restricts to a (measure-theoretic) bijection from a to Y, non-singular with respect to the measure m;

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

- $d(F(x), F(y)) \geq \kappa d(x, y)$;
- $d(f^k(x), f^k(y)) \leq K d(F(x), F(y))$ for all $0 \leq k \leq \tau(a)$;
- the inverse Jacobian $\zeta_a = dm/(dm \circ F)$ of the restriction $F : a \to Y$ satisfies

$$|\log |\zeta_a(x)| - \log |\zeta_a(y)|| \leq K d(F(x), F(y))^\eta.$$

The map f as above is said to be non-uniformly expanding. It is standard [**1**, Corollary p. 199], [**32**, Proof of Theorem 1] that there is a unique absolutely continuous F-invariant probability measure μ_Y on Y with $1/c \leq d\mu_Y/dm \leq c$ for some $c > 0$, and the corresponding f-invariant probability measure μ on X.

We make an additional assumption, which is not part of the usual definition of non-uniformly expanding maps, but is straightforward to verify in examples. Denote by $\mathcal{A}$ the set of all finite words in the alphabet α (not including the empty word) and set $Y_w := \bigcap_{k=0}^n F^{-k}(a_k)$ for $w = a_0 \cdots a_n$ in $\mathcal{A}$. We assume that

$$m(Y_w) = m(\bar{Y}_w) \quad \text{for every } w \in \mathcal{A}. \tag{5.1}$$

We say that the return times of f have:

- a weak polynomial moment of order $\beta \geq 1$, if $m(\tau \geq n) \ll n^{-\beta}$;
- a strong polynomial moment of order $\beta \geq 1$, if $\int \tau^\beta \, dm < \infty$.

Remark 5.1. Intermittent maps (1.1) and (1.5) are non-uniformly expanding. Their return times have respective weak and strong moments of order $\beta = 1/\gamma$.

More generally, our results apply to non-uniformly expanding and non-uniformly hyperbolic dynamical systems which can be modelled by Young towers [**31**, **32**]. A notable example with polynomial return times is the class of non-Markov maps with indifferent fixed points in [**32**, §7]. (Cf. AFN maps in Zweimüller [**34**].)

5.2. *Rates in the ASIP.* Suppose that $\varphi : X \to \mathbb{R}$ is a Hölder continuous observable such that $\mu(\varphi) = 0$. Let $S_n(\varphi) = \sum_{k=0}^{n-1} \varphi \circ f^k$ be the corresponding random process, defined on the probability space (X, μ). Assume in addition that the return times of f have a polynomial moment of order $\beta > 2$ (weak or strong). Let

$$c^2 = \lim_{n\to\infty} \frac{1}{n} \int |S_n(\varphi)|^2 \, d\mu. \tag{5.2}$$

Remark 5.2. The limit above exists by e.g. [**17**, Corollary 2.12]. In the case of summable correlations, c^2 can be computed by the formula (1.4), but in the setup of this section, f may be non-mixing and the correlations may not decay.

The ASIP for $S_n(\varphi)$ with variance c^2 was first proved in [**20**]. Prior to our work, the best available rates were due to [**5**, **17**], formulated for the *strong* polynomial moment of order β:

$$S_n(\varphi) - W_n = \begin{cases} o(n^{1/\beta}(\log n)^{1/2}), & \beta \in (2, 4), \\ O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}), & \beta \geq 4. \end{cases}$$

Again those rates are not better than $O(n^{1/4})$. Our main result is the following.

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

THEOREM 5.3. *Suppose that the return times of f have a weak polynomial moment of order $\beta > 2$. Then $S_n(\varphi)$ satisfies the ASIP with variance c^2 given by (5.2) and rate $o(n^{1/\beta}(\log n)^{1/\beta+\varepsilon})$ for all $\varepsilon > 0$. Further, if the return times of f have a strong polynomial moment of order $\beta > 2$, then the rate is $o(n^{1/\beta})$.*

In the remainder of this section we prove Theorem 5.3. First we consider the special case when $c^2 = 0$.

PROPOSITION 5.4. *Suppose that the return times of f have a weak polynomial moment of order β. Then on the probability space (Y, μ_Y),*

$$\max_{k\le n} \tau \circ F^k = o(n^{1/\beta}(\log n)^{1/\beta+\varepsilon}) \quad \text{a.s. for all } \varepsilon > 0.$$

With a strong polynomial moment of order β,

$$\max_{k\le n} \tau \circ F^k = o(n^{1/\beta}) \quad \text{a.s.}$$

Proof. The sequence $(\tau \circ F^n)_{n\ge 0}$ is stationary and by the Borel–Cantelli lemma, it suffices to check that for all $\delta > 0$,

$$\sum_{n\ge 1} \mu_Y(\tau > \delta n^{1/\beta}(\log n)^{1/\beta+\varepsilon}) < \infty.$$

Since $d\mu_Y/d\mu$ is bounded, it is enough to verify that

$$\sum_{n\ge 1} m(\tau > \delta n^{1/\beta}(\log n)^{1/\beta+\varepsilon}) < \infty,$$

which follows immediately from our assumptions.

The proof for the strong polynomial moments is similar. (See also [**17**, Proposition 2.6].) □

COROLLARY 5.5. *Theorem 5.3 holds when $c^2 = 0$.*

Proof. In [**17**], the ASIP for non-uniformly expanding dynamical systems uses the martingale–coboundary decomposition. With $c^2 = 0$, the martingale part vanishes [**17**, Corollaries 2.12 and 3.4]. The estimates of the coboundary part are reduced to those in Proposition 5.4, see the proof of [**17**, Proposition 2.6]. □

From here on, we assume that $c^2 > 0$. We construct a Markov chain as in §2. The general setup of non-uniformly expanding maps brings in a few minor technical complications that we explain below.

- Proposition 2.2 is the basis of the Markov chain construction, providing a 'regenerative' decomposition of the reference measure. It is proved in the general setup in [**15**].
- Construction of the semiconjugacy $\pi : \Omega \to X$ needs additional work, because we may not be able to define it everywhere as we did for the intermittent maps. Nevertheless, using assumption (5.1) we define it almost everywhere as follows.

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

Let $\mathcal{A}_n \subset \mathcal{A}$ denote the set of all words with length $n+1$. Let

$$Z = \bigcup_{n\in\mathbb{N}} \bigcup_{w\neq w'\in\mathcal{A}_n} (\bar{Y}_w \cap \bar{Y}_{w'}),$$

$$\tilde{Y} = \left(\bigcap_{n\in\mathbb{N}} \bigcup_{w\in\mathcal{A}_n} Y_w\right) \Big\backslash Z.$$

Then $m(Z) = 0$ and $m(\tilde{Y}) = 1$, and for every $y \in \tilde{Y}$ there exists a *unique* sequence $(a_n)_{n\in\mathbb{N}} \in \alpha^{\mathbb{N}}$ such that $y \in \bigcap_{n\in\mathbb{N}} Y_{a_0\cdots a_n}$.

We endow $\alpha^{\mathbb{N}}$ with the metric $\delta((a_n)_{n\in\mathbb{N}}, (a'_n)_{n\in\mathbb{N}}) = \kappa^{-s}$, where $s \geq 0$ is the largest such that $a_0 \cdots a_{s-1} = a'_0 \cdots a'_{s-1}$. Define a map $\chi\colon \alpha^{\mathbb{N}} \to Y$ by $\chi((a_n)_{n\in\mathbb{N}}) = y$, where $\{y\} = \bigcap_{n\in\mathbb{N}} \bar{Y}_{w_n}$. It follows from completeness of X and expansion of F that χ is defined everywhere and is Lipschitz. Set $\mathcal{X} := \chi^{-1}(\tilde{Y})$. Then $\mathcal{X}$ is measurable and $\{\chi((a_n)_{n\in\mathbb{N}})\} = \bigcap_{n\in\mathbb{N}} Y_{a_0\cdots a_n}$ for every $(a_n)_{n\in\mathbb{N}} \in \mathcal{X}$.

Every $g \in \Omega$ can be written as

$$g = ((w_0, \ell_0), \ldots, (w_0, h(w_0)-1), (w_1, 0), \ldots, (w_1, h(w_1)-1), (w_2, 0), \ldots).$$

Let $\Omega_0 = \{g \in \Omega : \ell_0 = 0\}$. Define $\iota\colon \Omega_0 \to \alpha^{\mathbb{N}}$ by $\iota(g) = (a_0, a_1, \ldots)$ where $a_0 a_1 \cdots = w_0 w_1 \cdots$. Define $\Omega'_0 = \iota^{-1}(\mathcal{X})$ and $\pi_0\colon \Omega'_0 \to \tilde{Y}$, $\pi_0 = \chi \circ \iota$. Let

$$\Omega' = \{\sigma^\ell(g) : g \in \Omega'_0,\ 0 \leq \ell < h(w_0)\}.$$

Observe that $\mathbb{P}_\Omega(\Omega') = 1$. Define a projection $\pi : \Omega' \to X$ by $\pi(g) = f^{\ell_0}(\pi_0(g^{(0)}))$. Following the proof of Lemma 2.4 with straightforward changes, we see that π is Lipschitz on Ω'.

Remark 5.6. Construction of the Markov chain for non-uniformly expanding dynamical systems can be found in [**16**], done in different notation. There the space X is not assumed complete, and a more general, though less hands-on, assumption is used in place of (5.1): that the set

$$\{(a_0, a_1, \ldots) \in \alpha^{\mathbb{N}} : \text{there exists } y \in Y \text{ with } F^k(y) \in a_k \text{ for all } k\}$$

is measurable in $\alpha^{\mathbb{N}}$ (in the product topology with Borel sigma algebra).

Further we work in notation of §2. Let

$$p = \gcd\{h(w)\colon w \in \mathcal{A}\}.$$

For the maps (1.1) and (1.5) we showed that $p = 1$. This means that the Markov chain $g_0, g_1, \ldots$ is aperiodic, which was necessary to control the moments of the meeting time in §3. In the general case, however, it could be that $p \geq 2$. This is typical for example for logistic maps with Collet–Eckmann parameters.

For $p = 1$, our proof proceeds without changes. Below we treat the periodic case $p \geq 2$. For $0 \leq k < p$, define

$$\tilde{S}_k = \{(w, \ell) \in S\colon \ell \equiv k \pmod p\}$$

and

$$\Omega_k = \{(g_0, g_1, \ldots) \in \Omega\colon g_0 \in \tilde{S}_k\}.$$

The sets Ω_k partition Ω, and they are cyclically permuted by $\sigma\colon \sigma(\Omega_k) = \Omega_{k+1 \bmod p}$.

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

Note that if $g_{np} = (w, \ell) \in \tilde{S}_0$ for some $n \geq 0$, then $g_{np+k} = (w, \ell + k)$ for $0 \leq k < p$. Thus we can identify Ω_0 with

$$\tilde{\Omega} = \{(g_0, g_p, g_{2p} \ldots) \in \Omega \colon g_0 \in \tilde{S}_0\}.$$

Let now $(\tilde{g}_0, \tilde{g}_1, \ldots)$ be a Markov chain with state space $\tilde{S}_0$ and transition probabilities

$$\mathbb{P}(\tilde{g}_{n+1} = (w, \ell p) \mid \tilde{g}_n = (w', \ell' p)) = \begin{cases} 1, & \ell = \ell' + 1 \text{ and } \ell' + 1 < h(w)/p \text{ and } w = w', \\ \mathbb{P}_{\mathcal{A}}(w), & \ell = 0 \text{ and } \ell' + 1 = h(w')/p, \\ 0 & \text{else.} \end{cases}$$

This Markov chain admits a unique (ergodic) invariant probability measure $\tilde{\nu}$ on $\tilde{S}_0$ given by

$$\tilde{\nu}(w, \ell p) = p \frac{\mathbb{P}_{\mathcal{A}}(w) \mathbf{1}_{\{0 \leq \ell < h(w)/p\}}}{\mathbb{E}_{\mathcal{A}}(h)}. \tag{5.3}$$

The Markov chain $(\tilde{g}_n)_{n \geq 0}$ starting from $\tilde{\nu}$ defines a probability measure $\mathbb{P}_{\tilde{\Omega}}$ on the space $\tilde{\Omega}$. Note that $\mathbb{P}_{\tilde{\Omega}}$ corresponds to $\mathbb{P}_{\Omega}$ conditioned on Ω_0. Note also that $\tilde{g}_0, \tilde{g}_1, \ldots$ is a Markov chain, identical to $g_0, g_1, \ldots$ in structure except that it is aperiodic and the return times to $S_0 = \{(w, \ell) \in S \colon \ell = 0\}$ are divided by p.

Following §2, we define the separation time $\tilde{s}$ and the separation metric $\tilde{d}$ on $\tilde{\Omega}$, using the same constant $\lambda > 1$. Suppose that $\tilde{a}, \tilde{b} \in \tilde{\Omega}$ with the corresponding $a, b \in \Omega_0$. The separation time is measured in terms of returns to S_0, hence

$$\tilde{s}(\tilde{a}, \tilde{b}) = s(a, b) \quad \text{and} \quad \tilde{d}(\tilde{a}, \tilde{b}) = d(a, b).$$

Further, $d(\sigma^k(a), \sigma^k(b)) = d(a, b)$ for $0 \leq k < p$. Let $\widetilde{\psi} \colon \tilde{\Omega} \to \mathbb{R}$,

$$\widetilde{\psi}(\tilde{a}) = \sum_{k=0}^{p-1} \psi(\sigma^k(a)).$$

It follows that $\operatorname{Lip} \widetilde{\psi} \leq p \operatorname{Lip} \psi$. Also, $\widetilde{\psi}$ is mean zero with respect to $\mathbb{P}_{\tilde{\Omega}}$.

In Corollary A.2 of Appendix A we show that the ASIP for $\sum_{k=0}^{n-1} \psi \circ \sigma^k$ on $(\Omega, \mathbb{P}_{\Omega})$ (and hence the ASIP for $\sum_{k=0}^{n-1} \varphi \circ f^k$ on (X, μ)) follows from the ASIP for $\sum_{k=0}^{n-1} \widetilde{\psi} \circ \tilde{\sigma}^k$ on $(\tilde{\Omega}, \mathbb{P}_{\tilde{\Omega}})$ with the same rates and variance v^2/p, where v^2 is the variance of the Wiener process on $(\tilde{\Omega}, \mathbb{P}_{\tilde{\Omega}})$.

Now, as in §4, the ASIP for $\sum_{k=0}^{n-1} \widetilde{\psi} \circ \tilde{\sigma}^k$ on $(\tilde{\Omega}, \mathbb{P}_{\tilde{\Omega}})$ follows from the ASIP for $\sum_{k=0}^{n-1} \tilde{X}_k$ where, for any $k \geq 0$,

$$\tilde{X}_k = \widetilde{\psi}((\tilde{g}_\ell)_{\ell \geq k}),$$

and $(\tilde{g}_n)_{n \geq 0}$ is the stationary Markov chain defined above with the state space $\tilde{S}_0$ and stationary distribution $\tilde{\nu}$. The proof of the ASIP for $\sum_{k=1}^{n} \tilde{X}_k$ with the adequate rates is, as in §4, mainly based on suitable bounds for the tails of the meeting time $\tilde{T}$ for the Markov chain $(\tilde{g}_n)_{n \in \mathbb{N}}$, which is defined as follows. First, without changing the distribution, we redefine $(\tilde{g}_n)_{n \in \mathbb{N}}$ on a new probability space as follows. Let $\tilde{g}_0 \in \tilde{S}_0$ be distributed according to $\tilde{\nu}$ (the stationary distribution defined by (5.3)). Let $\varepsilon_1, \varepsilon_2, \ldots$ be a sequence

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

of iid random variables with values in $\mathcal{A}$, distribution $\mathbb{P}_{\mathcal{A}}$ and independent from $\tilde{g}_0$. For $n \geq 0$, let

$$\tilde{g}_{n+1} = \tilde{U}(\tilde{g}_n, \varepsilon_{n+1}),$$

where, for any $\ell \in \mathbb{N}$,

$$\tilde{U}((w, \ell p), \varepsilon) = \begin{cases} (w, (\ell+1)p), & \ell p < h(w) - p, \\ (\varepsilon, 0), & \ell p = h(w) - p. \end{cases} \tag{5.4}$$

The meeting time $\tilde{T}$ of the Markov chain $(\tilde{g}_n)_{n\in\mathbb{N}}$ is then defined by

$$\tilde{T} = \inf\{n \geq 0 \colon \tilde{g}_n = \tilde{g}_n^*\}, \tag{5.5}$$

where $(\tilde{g}_n^*, n \in \mathbb{N})$ is the Markov chain defined as follows: $\tilde{g}_0^*$ is a random variable in $\tilde{S}_0$ with distribution $\tilde{\nu}$ and independent from $(\tilde{g}_0, \{\varepsilon_n\}_{n\geq 1})$ and, for $n \geq 0$, $\tilde{g}_{n+1}^* = \tilde{U}(\tilde{g}_n^*, \varepsilon_{n+1})$. Proceeding as in the proof of Lemma 3.1 and taking into account the bounds on the tails of h proved in [**15**], we infer that the following lemma holds.

LEMMA 5.7.

- *If the return times of f have weak polynomial moment of order $\beta > 1$, then, for any $\eta > 1$, $\mathbb{E}(\widetilde{\psi}_{\beta,\eta}(\tilde{T})) < \infty$, where $\widetilde{\psi}_{\beta,\eta}(x) = x^{\beta-1}(\log(1+x))^{-\eta}$ for $x > 0$.*
- *If the return times of f have strong polynomial moment of order $\beta > 1$, then $\mathbb{E}(\tilde{T}^{\beta-1}) < \infty$.*

In addition, proceeding as in the proof of Lemma 3.3, we also get the following bound.

LEMMA 5.8. *Assume that $\mathbb{E}(\tilde{T}) < \infty$. Then, for any $k \geq 1$ and any $\alpha \geq 1$,*

$$|\mathrm{Cov}\,(\tilde{X}_0, \tilde{X}_k)| \ll k^{-\alpha/2} + \mathbb{P}(\tilde{T} \geq [k/4\alpha]).$$

Now, with the same arguments as those developed in §4 and taking into account Lemmas 5.7 and 5.8, we infer that, enlarging the underlying probability space if necessary, there exists a sequence $(N_i)_{i\geq 1}$ of iid centered Gaussian random variables with variance

$$v^2 = \mathrm{Var}\,(\tilde{X}_0) + 2\sum_{k\geq 1} \mathrm{Cov}\,(\tilde{X}_0, \tilde{X}_k) \tag{5.6}$$

such that, for any $\kappa > 1/\beta$,

$$\sup_{k\leq n} \left| \sum_{i=0}^{k-1} \tilde{X}_i - \sum_{i=1}^{k} N_i \right| = o(n^{1/\beta}(\log n)^{\eta\kappa}) \quad \text{a.s.},$$

where $\eta = 0$ if the return times of f have strong polynomial moment of order $\beta > 2$ and $\eta = 1$ if the return times of f have weak polynomial moment of order $\beta > 2$.

Now, according to Corollary A.2, the ASIP for $\sum_{k=0}^{n-1} \psi \circ \sigma^k$ on $(\Omega, \mathbb{P}_\Omega)$ (and then the ASIP for $\sum_{k=0}^{n-1} \varphi \circ f^k$ on (X, μ)) holds with variance v^2/p and rate $o(n^{1/\beta}(\log n)^{\eta\kappa})$. It remains to check that $v^2/p = c^2$, with c^2 given by (5.2). Since by Lemmas 5.8 and 5.7, the series defined in (5.6) is absolutely convergent, we have

$$v^2 = \lim_{n\to\infty} \frac{1}{n} \mathrm{Var}\left(\sum_{k=0}^{n-1} \tilde{X}_n\right) = \lim_{n\to\infty} \frac{1}{n} \mathbb{E}_{\mathbb{P}_{\tilde{\Omega}}}\left(\left(\sum_{k=0}^{n-1} \widetilde{\psi} \circ \tilde{\sigma}^k\right)^2\right).$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

Hence, according to Lemma A.1,

$$v^2 = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_{\mathbb{P}_\Omega}\Big(\Big(\sum_{k=0}^{np-1} \psi\circ\sigma^k\Big)^2\Big) = \lim_{n\to\infty} \frac{1}{n}\|S_{np}(\varphi)\|_{2,\mu}^2 = pc^2.$$

This ends the proof of Theorem 5.3 when $c^2 > 0$.

5.3. *Optimality of the rates.* In this subsection we prove Proposition 1.4. In fact we prove a stronger statement as follows. We consider a non-uniformly expanding map $f\colon X \to X$ as above. We assume that:

- τ is the first return time to Y;
- for some $\beta > 2$, $\kappa > 0$ and all $n \geq 1$,

$$m(\tau \geq n) \geq \frac{\kappa}{n^\beta}.$$

These assumptions are verified for the map (1.1) with $\beta = 1/\gamma$ and $Y = [1/2, 1]$.

PROPOSITION 5.9. *Let ψ be a bounded observable such that $\psi \equiv 0$ on $X\backslash Y$ and $\mu(\psi) > 0$, and let $\varphi = \psi - \mu(\psi)$. Then for every process $(Z_n)_{n\in\mathbb{N}}$ with the same law as $(\varphi\circ f^n)_{n\in\mathbb{N}}$ and every stationary and Gaussian centered sequence $(g_k)_{k\in\mathbb{Z}}$ such that $n^{-1}\mathrm{Var}(\sum_{i=1}^n g_i)$ converges, living on a same probability space,*

$$\limsup_{n\to\infty} (n\log n)^{-1/\beta}\Big|\sum_{k=1}^n Z_k - \sum_{k=1}^n g_k\Big| > 0 \quad a.s.$$

Remark 5.10. Under relaxed assumptions, there exist Lipshitz observables φ with $\int \varphi\, d\mu = 0$ satisfying the hypotheses of Proposition 5.9. Indeed, if m is regular and $\mu(\mathring{Y}) \neq 0$, then there is a compact $K \subset \mathring{Y}$ such that $\mu(K) > 0$. Note that K and $\overline{X\backslash Y}$ are closed disjoint sets. Thus $\psi\colon X \to \mathbb{R}$,

$$\psi(x) := \frac{d(x, \overline{X\backslash Y})}{d(x, \overline{X\backslash Y}) + d(x, K)}$$

is Lipschitz, and so is $\varphi = \psi - \int \psi\, d\mu$.

Remark 5.11. If $f\colon X \to X$ is a Young tower [**32**], then one can take $\varphi = \mathbf{1}_Y - \mu(Y)$. Then φ is Lipschitz with respect to the distance on the tower.

Proof of Proposition 5.9. Recall that μ_Y is the F-invariant probability measure on Y. For $n \geq 0$, let $\tau_n = \tau\circ F^n$. We claim that μ_Y-a.s., $n^{-1}\sum_{i=1}^n \tau_i \to \int \tau\, d\mu$ as $n \to \infty$ and $\tau_n \geq (n\log n)^{1/\beta}$ infinitely often. Then our result follows as in the proof of [**4**, Proposition 15].

It remains to verify the claim. Its first part is provided by the pointwise ergodic theorem, so further we verify the second part. We follow Gouëzel [**9**].

For $n \geq 0$, let $A_n = \{y \in Y\colon \tau_n(y) \geq (n\log n)^{1/\beta}\}$. Recall that there is a constant $c > 0$ such that for all $n, k \geq 0$,

$$\mu_Y(\tau_n = k) \geq c\, m(\tau = k).$$

Thus

$$\sum_{n=0}^\infty \mu_Y(A_n) \geq c\sum_{n=0}^\infty m(\tau \geq (n\log n)^{1/\beta}) = \infty. \tag{5.7}$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

Next, there are constants $C > 0$ and $\theta \in]0, 1[$ such that for all $k \neq n \geq 0$,

$$|\mu_Y(A_k \cap A_n) - \mu_Y(A_k)\mu_Y(A_n)| \leq C\theta^{|n-k|}\mu_Y(A_k)\mu_Y(A_n).$$

(See, for instance, the last line of [**1**, §1].) Therefore,

$$\begin{aligned}\left|\sum_{1\leq k,\ell\leq n} \mu_Y(A_k \cap A_\ell) - \sum_{1\leq k,\ell\leq n} \mu_Y(A_k)\mu_Y(A_\ell)\right| &\leq \sum_{1\leq k,\ell\leq n} |\mu_Y(A_k \cap A_\ell) - \mu_Y(A_k)\mu_Y(A_\ell)| \\ &\ll \sum_{k=1}^{n} \mu_Y(A_k) + \sum_{1\leq k,\ell\leq n} \theta^{|\ell-k|}\mu_Y(A_k)\mu_Y(A_\ell) \ll \sum_{k=1}^{n} \mu_Y(A_k).\end{aligned}$$

Taking into account (5.7), we obtain

$$\lim_{n\to\infty} \frac{\sum_{1\leq k,\ell\leq n} \mu_Y(A_k \cap A_\ell)}{(\sum_{k=1}^{n} \mu_Y(A_k))^2} = 1.$$

By [**6**, Lemma C], we verify a criterion for the second Borel–Cantelli lemma and prove that $\mu_Y(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) = 1$, i.e. that μ_Y-a.s., $\tau_n \geq (n \log n)^{1/\beta}$ infinitely often. This completes the proof of the claim. □

A. *Appendix. ASIP for periodic dynamical systems*

Suppose that $(\Omega, \mathbb{P})$ is a probability space and $\sigma\colon \Omega \to \Omega$ is a measure preserving transformation.

Suppose that $p \geq 2$ is an integer and σ is p-periodic in the sense that Ω can be partitioned into disjoint subsets $\Omega_0, \ldots, \Omega_{p-1}$, which are permuted by σ cyclically: $\sigma(\Omega_k) = \Omega_{k+1 \bmod p}$. In particular, $\mathbb{P}(\Omega_k) = 1/p$ for any $k = 0, \ldots, p-1$.

Let $\tilde{\sigma}\colon \Omega_0 \to \Omega_0$, $\tilde{\sigma} = \sigma^p$. We refer to $\tilde{\sigma}$ as the *induced map.* The space Ω_0 is endowed with a probability measure $\mathbb{P}_0$, which is $\mathbb{P}$ conditioned on Ω_0. Note that $\mathbb{P}_0$ is invariant under $\tilde{\sigma}$.

Suppose that $\psi\colon \Omega \to \mathbb{R}$ is an observable with $|\psi|_\infty = \sup_\Omega |\psi| < \infty$. Define the *induced observable* $\widetilde{\psi}\colon \Omega_0 \to \mathbb{R}$,

$$\widetilde{\psi}(x) = \sum_{k=0}^{p-1} \psi(\sigma^k(x)).$$

Denote

$$\psi_n = \sum_{k=0}^{n-1} \psi \circ \sigma^k \quad \text{and} \quad \widetilde{\psi}_n = \sum_{k=0}^{n-1} \widetilde{\psi} \circ \tilde{\sigma}^k.$$

We consider ψ_n and $\widetilde{\psi}_n$ as random processes, defined on probability spaces $(\Omega, \mathbb{P})$ and $(\Omega, \mathbb{P}_0)$, respectively. Define a projection $\pi_0\colon \Omega \to \Omega_0$ by

$$\pi_0(x) = \begin{cases} x & \text{if } x \in \Omega_0, \\ \sigma^{p-k}(x) & \text{if } x \in \Omega_k,\ k = 1, \ldots, p-1. \end{cases} \tag{A.1}$$

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

LEMMA A.1. *We have*

$$|\psi_n - \widetilde{\psi}_{[n/p]} \circ \pi_0|_\infty \le 2p|\psi|_\infty. \tag{A.2}$$

Moreover, if $\lim_{n\to\infty} n^{-1} \int_\Omega \psi_n^2 \mathbb{P}(\omega)\, d\omega = c^2$, *then* $\lim_{n\to\infty} n^{-1} \int_{\Omega_0} \widetilde{\psi}^2_{[n/p]} \mathbb{P}_0(\omega)\, d\omega = c^2$.

Proof. The bound (A.2) is obvious. Indeed, for instance, if $x \in \Omega_1$, it suffices to write

$$|\psi_n - \widetilde{\psi}_{[n/p]} \circ \pi_0| = \left| \sum_{k=0}^{n-1} \psi \circ \sigma^k - \sum_{k=p-1}^{p[n/p]+p-2} \psi \circ \sigma^k \right| \le 2(p-1)|\psi|_\infty.$$

To end the proof of the lemma, note that $(\pi_0)_* \mathbb{P} = \mathbb{P}_0$, thus $\widetilde{\psi}_n \circ \pi_0$, defined on the probability space $(\Omega, \mathbb{P})$, has the same distribution as $\widetilde{\psi}_n$ on $(\Omega_0, \mathbb{P}_0)$. □

COROLLARY A.2. *Let* $(b_n)_{n\ge1}$ *be a regularly varying sequence with values in* $\mathbb{R}^+$, *and such that* $b_n(\log n)^{-1/2} \to \infty$ *as* $n \to \infty$. *Assume that* Ω *can be enlarged in such a way that there exists a Brownian motion* $\tilde{W}_t$ *(with variance* v^2*) such that*

$$\widetilde{\psi}_n \circ \pi_0 = \tilde{W}_n + o(b_n) \quad a.s.$$

Then, on the same probability space, there is a Brownian motion W_t *(with variance* $c^2 = v^2/p$*) such that*

$$\psi_n = W_n + o(b_n) \quad a.s.$$

Proof. By assumption and Lemma A.1,

$$\psi_n = \tilde{W}_{[n/p]} + o(b_n) \quad \text{a.s.}$$

Then $W_t = \tilde{W}_{t/p}$ is a Brownian motion (with variance $c^2 = v^2/p$), and

$$\sup_{s\le t} |\tilde{W}_{s/p} - \tilde{W}_{[s/p]}| = O((\log t)^{1/2}) \quad \text{a.s.}$$

(See, for instance, [**10**, Theorem 3.2A].) The result follows. □

Acknowledgements. A.K. is partially supported by a European Advanced Grant *StochExtHomog* (ERC AdG 320977); is thankful to Ian Melbourne for helpful discussions; acknowledges the warm hospitality of Paris-Est and Paris Descartes universities.

REFERENCES

[**1**] J. Aaronson and M. Denker. Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps. *Stoch. Dyn.* **1**(2) (2001), 193–237.
[**2**] G. Bennett. Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* **57** (1962), 33–45.
[**3**] I. Berkes, W. Liu and W. Wu. Komlós–Major–Tusnády approximation under dependence. *Ann. Probab.* **42** (2014), 794–817.
[**4**] C. Cuny, J. Dedecker and F. Merlevède. On the Komlós, Major and Tusnády strong approximation for some classes of random iterates. *Stochastic Process. Appl.* **128**(4) (2018), 1347–1385.
[**5**] C. Cuny and F. Merlevède. Strong invariance principles with rate for 'reverse' martingales and applications. *J. Theoret. Probab.* **28**(1) (2015), 137–183.

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press

[**6**] P. Erdős and A. Rényi. On Cantor's series with convergent $\sum 1/q_n$. *Ann. Univ. Sci. Budapest. Eötvös. Sect. Math.* **2** (1959), 93–109.
[**7**] S. Gouëzel. Sharp polynomial estimates for the decay of correlations. *Israel J. Math.* **139** (2004), 29–65.
[**8**] S. Gouëzel. Vitesse de décorrélation et théorèmes limites pour les applications non uniformément dilatantes. *PhD Thesis*, Orsay, 2004.
[**9**] S. Gouëzel. A Borel–Cantelli lemma for intermittent interval maps. *Nonlinearity* **20** (2007), 1491–1497.
[**10**] D. L. Hanson and R. P. Russo. Some results on increments of the Wiener process with applications to lag sums of i.i.d. random variables. *Ann. Probab.* **11**(3) (1983), 609–623.
[**11**] M. Holland. Slowly mixing systems and intermittency maps. *Ergod. Th. & Dynam. Sys.* **25** (2005), 133–159.
[**12**] H. Hu. Decay of correlations for piecewise smooth maps with indifferent fixed points. *Ergod. Th. & Dynam. Sys.* **24** (2004), 495–524.
[**13**] J. Komlós, P. Major and G. Tusnády. An approximation of partial sums of independent RV'-s and the sample DF. I. *Z. Wahrscheinlichkeitstheor. verw. Geb.* **32** (1975), 111–131.
[**14**] J. Komlós, P. Major and G. Tusnády. An approximation of partial sums of independent RV'-s and the sample DF. II. *Z. Wahrscheinlichkeitstheor. verw. Geb.* **34** (1976), 34–58.
[**15**] A. Korepanov. Equidistribution for nonuniformly expanding systems. *Comm. Math. Phys.* **359** (2018), 1123–1138.
[**16**] A. Korepanov. Rates in almost sure invariance principle for dynamical systems with some hyperbolicity. *Comm. Math. Phys.* **363** (2018), 173–190.
[**17**] A. Korepanov, Z. Kosloff and I. Melbourne. Martingale-coboundary decomposition for families of dynamical systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **35** (2018), 859–885.
[**18**] T. Lindvall. On coupling of discrete renewal processes. *Z. Wahrscheinlichkeitstheor. verw. Geb.* **48** (1979), 57–70.
[**19**] C. Liverani, B. Saussol and S. Vaienti. A probabilistic approach to intermittency. *Ergod. Th. & Dynam. Sys.* **19** (1999), 671–685.
[**20**] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems. *Comm. Math. Phys.* **260** (2005), 131–146.
[**21**] I. Melbourne and M. Nicol. A vector-valued almost sure invariance principle for hyperbolic dynamical systems. *Ann. Probab.* **37**(2) (2009), 478–505.
[**22**] F. Merlevède and E. Rio. Strong approximation of partial sums under dependence conditions with application to dynamical systems. *Stochastic Process. Appl.* **122** (2012), 386–417.
[**23**] F. Merlevède and E. Rio. Strong approximation for additive functionals of geometrically ergodic Markov chains. *Electron. J. Probab.* **20** (2015), 1–27.
[**24**] E. Nummelin. *General Irreducible Markov Chains and Nonnegative Operators (Cambridge Tracts in Mathematics, 83)*. Cambridge University Press, Cambridge, 1984.
[**25**] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Comm. Math. Phys.* **74** (1980), 189–197.
[**26**] W. Philipp and W. F. Stout. Almost sure invariance principle for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc.* **161** (1975), iv+140pp.
[**27**] J. W. Pitman. Uniform rates of convergence for Markov chain transition probabilities. *Z. Wahrscheinlichkeitstheor. verw. Geb.* **29** (1974), 193–227.
[**28**] E. Rio. Théorie asymptotique des processus aléatoires faiblement dépendants. *Math. Appl. (Berlin)* **31** (2000), x+169pp.
[**29**] A. I. Sakhanenko. Estimates in the invariance principle in terms of truncated power moments. *Sib. Math. J.* **47** (2006), 1355–1371.
[**30**] O. Sarig. Subexponential decay of correlations. *Invent. Math.* **150** (2002), 629–653.
[**31**] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math. (2)* **147** (1998), 585–650.
[**32**] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.* **110** (1999), 153–188.
[**33**] A. Yu. Zaitsev. Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments. *ESAIM Probab. Stat.* **2** (1998), 41–108.
[**34**] R. Zweimüller. Ergodic structure and invariant densities of non-Markovian interval maps with indifferent fixed points. *Nonlinearity* **11** (1998), 1263–1276.
[**35**] R. Zweimüller. Measure preserving transformations similar to Markov shifts. *Israel J. Math.* **173** (2009), 421–443.

https://doi.org/10.1017/etds.2019.2 Published online by Cambridge University Press