Rates in almost sure invariance principle for non uniformly expanding maps

Christophe Cuny

joint work with J. Dedecker, A. Korepanov and F. Merlevède

Université de Bretagne Occidentale

Luminy, november 2018

Strong invariance principles

- Let (X, Σ, μ) be a probability space. For $(X_n)_{n\geq 1} \subset L^2(\mu)$ set $S_n = X_1 + \ldots + X_n$. If T is an ergodic transformation preserving μ and $f \in L^2(\mu)$ we shall take $X_n = f \circ T^{n-1}$.

Strong invariance principles

- Let (X, Σ, μ) be a probability space. For $(X_n)_{n\geq 1} \subset L^2(\mu)$ set $S_n = X_1 + \ldots + X_n$. If T is an ergodic transformation preserving μ and $f \in L^2(\mu)$ we shall take $X_n = f \circ T^{n-1}$.
- Let $(r_n)_{n\geq 1}$ be a non decreasing sequence with $r_n = O((n \log\log n)^{1/2})$. We say that $(S_n)_{n\geq 1}$ satisfies the almost sure invariance principle (ASIP) with rate $(r_n)_{n\geq 1}$, if one can redefine $(S_n)_{n\geq 1}$ without changing its distribution on a (richer) probability space on which there exists a sequence $(Z_i)_{i\geq 1}$ of iid centered Gaussian variables such that

$$\max_{k\leq n} |S_k - B_k| = o(b_n) \text{ almost surely,}$$

where $B_k = \sum_{i=1}^{k} Z_i$.

First results: the iid case

Assume that $(X_n)_{n\geq 1}$ is iid. Then, one has the following ASIPs with rate $(b_n)_{n\geq 1}$.

- Strassen ('64): $b_n = (n \log\log n)^{1/2}$ when $X_1 \in L^2$;
- Major ('76): $b_n = n^{1/p}$ when $X_1 \in L^p$, $2 < p \leq 3$;
- Komlós-Major-Tusnády ('75): $b_n = n^{1/p}$ when $X_1 \in L^p$, $p > 3$;
- Komlós-Major-Tusnády ('76): $b_n = \log n$ when $\mathbb{E}(\mathrm{e}^{\varepsilon |X_1|}) < \infty$, for some $\varepsilon > 0$.

First results: the iid case

Assume that $(X_n)_{n\geq 1}$ is iid. Then, one has the following ASIPs with rate $(b_n)_{n\geq 1}$.

- Strassen ('64): $b_n = (n \log\log n)^{1/2}$ when $X_1 \in L^2$;
- Major ('76): $b_n = n^{1/p}$ when $X_1 \in L^p$, $2 < p \leq 3$;
- Komlós-Major-Tusnády ('75): $b_n = n^{1/p}$ when $X_1 \in L^p$, $p > 3$;
- Komlós-Major-Tusnády ('76): $b_n = \log n$ when $\mathbb{E}(e^{\varepsilon |X_1|}) < \infty$, for some $\varepsilon > 0$.

When $(X_n)_{n\geq 1}$ is a stationary and ergodic sequence of (reverse) martingale differences and $X_1 \in L^p$, $2 \leq p \leq 4$, the above rates have been obtained up to some logarithmic factors.

A coupling inequality allowing to recover KMT's resullts

- Sakhanenko ('06). Let $(X_i)_{i\geq 1}$ be a sequence of independent, non necessarily identically distibuted, r.v.'s centered and in $\mathbb{L}^2$. Let $r>2$. On a richer probability space, one can construct a sequence $(Z_i)_{i\geq 1}$ of independent centered gaussian r.v.'s with $\mathrm{Var}(Z_n)=\mathrm{Var}(X_n)$ and such that for all $x>0$ and all $n\geq 1$,

$$\mathbb{P}\Big(\max_{1\leq k\leq n}\big|S_k-B_k\big|>c(r)x\Big)\leq\sum_{i=1}^{n}\mathbb{E}\min\Big(\frac{|X_i|^r}{x^r},\frac{|X_i|^2}{x^2}\Big)\,.$$

A coupling inequality allowing to recover KMT's resullts

- Sakhanenko ('06). Let $(X_i)_{i\geq 1}$ be a sequence of independent, non necessarily identically distibuted, r.v.'s centered and in $\mathbb{L}^2$. Let $r > 2$. On a richer probability space, one can construct a sequence $(Z_i)_{i\geq 1}$ of independent centered gaussian r.v.'s with $\mathrm{Var}(Z_n) = \mathrm{Var}(X_n)$ and such that for all $x > 0$ and all $n \geq 1$,

$$\mathbb{P}\Big(\max_{1\leq k\leq n}\big|S_k - B_k\big| > c(r)x\Big) \leq \sum_{i=1}^{n} \mathbb{E}\min\Big(\frac{|X_i|^r}{x^r}, \frac{|X_i|^2}{x^2}\Big)\,.$$

- Let $p > 2$. Applying the above result in the iid setting with $n = 2^m$, $x = 2^{m/p}$ and $r > p$, one can prove that for every $\varepsilon > 0$

$$\sum_{m\geq 0} \mathbb{P}\Big(\max_{1\leq k\leq n}\big|S_k - B_k\big| > \varepsilon 2^{m/p}\Big)\,,$$

which implies KMT's results for the corresponding $p > 2$.

A coupling inequality allowing to recover KMT's resullts

- Sakhanenko ('06). Let $(X_i)_{i\geq 1}$ be a sequence of independent, non necessarily identically distibuted, r.v.'s centered and in $\mathbb{L}^2$. Let $r>2$. On a richer probability space, one can construct a sequence $(Z_i)_{i\geq 1}$ of independent centered gaussian r.v.'s with $\mathrm{Var}(Z_n)=\mathrm{Var}(X_n)$ and such that for all $x>0$ and all $n\geq 1$,

$$\mathbb{P}\Big(\max_{1\leq k\leq n}\big|S_k-B_k\big|>c(r)x\Big)\leq\sum_{i=1}^n\mathbb{E}\min\Big(\frac{|X_i|^r}{x^r},\frac{|X_i|^2}{x^2}\Big)\,.$$

- Let $p>2$. Applying the above result in the iid setting with $n=2^m$, $x=2^{m/p}$ and $r>p$, one can prove that for every $\varepsilon>0$

$$\sum_{m\geq 0}\mathbb{P}\Big(\max_{1\leq k\leq n}\big|S_k-B_k\big|>\varepsilon 2^{m/p}\Big)\,,$$

  which implies KMT's results for the corresponding $p>2$.

- Sakhanenko ('84) proved a version of the above result for variables with exponential moments that implies KMT's result under an exponential moment.

A first dynamical example: the doubling map

Let $(X, \Sigma, \mu, T) = ([0,1), \mathcal{B}([0,1)), \lambda, T)$ where λ is the Lebesgue measure and $Tx = 2x \mod 1$. Let f be an Hölder observable on X.

- Until 2014, the best available rate in the ASIP has been $b_n = n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}$. This can be proved by mean of martingale approximation.

A first dynamical example: the doubling map

Let $(X, \Sigma, \mu, T) = ([0,1), \mathcal{B}([0,1)), \lambda, T)$ where λ is the Lebesgue measure and $Tx = 2x \mod 1$. Let f be an Hölder observable on X.

- Until 2014, the best available rate in the ASIP has been $b_n = n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}$. This can be proved by mean of martingale approximation.

  Notice that to study $(f \circ T^{n-1})_{n\geq 1}$ (under λ) it is enough to study $(f(Z_n)_{n\geq 1}$ where $Z_n = \sum_{k\geq 0} \frac{\varepsilon_{k+n}}{2^{k+1}}$ and $(\varepsilon_n)_{n\geq 1}$ is iid with $\mathbb{P}(\varepsilon_0 = 0) = \mathbb{P}(\varepsilon_0 = 1) = 1/2$.

- In 2014, Berkes-Liu-Wu proved the ASIP with rate $o(n^{1/p})$, $p > 2$, when $X_k = g(\varepsilon_k, , \varepsilon_{k+1}, \ldots)$ with $(\varepsilon_k)_{k\geq 1}$ are iid r.v.' s, $\|X_0\|_p < \infty$ and assuming some weak dependence conditions.

A first dynamical example: the doubling map

Let $(X, \Sigma, \mu, T) = ([0,1), \mathcal{B}([0,1)), \lambda, T)$ where λ is the Lebesgue measure and $Tx = 2x \mod 1$. Let f be an Hölder observable on X.

- Until 2014, the best available rate in the ASIP has been $b_n = n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}$. This can be proved by mean of martingale approximation.

  Notice that to study $(f \circ T^{n-1})_{n\geq 1}$ (under λ) it is enough to study $(f(Z_n)_{n\geq 1}$ where $Z_n = \sum_{k\geq 0} \frac{\varepsilon_{k+n}}{2^{k+1}}$ and $(\varepsilon_n)_{n\geq 1}$ is iid with $\mathbb{P}(\varepsilon_0 = 0) = \mathbb{P}(\varepsilon_0 = 1) = 1/2$.

- In 2014, Berkes-Liu-Wu proved the ASIP with rate $o(n^{1/p})$, $p > 2$, when $X_k = g(\varepsilon_k, , \varepsilon_{k+1}, \ldots)$ with $(\varepsilon_k)_{k\geq 1}$ are iid r.v.' s, $\|X_0\|_p < \infty$ and assuming some weak dependence conditions.
- They assume a polynomial decay of convergence of $\|X_1 - X^*_{1,k}\|_p$ where $X^*_{1,k} = g(\varepsilon_1, \ldots, \varepsilon_{k-1}, \varepsilon^*_k, \varepsilon_{k+1}, \ldots)$ and $(\varepsilon^*_k)_{k\geq 1}$ is an independent copy of $(\varepsilon_k)_{k\geq 1}$.

Extensions of the work of BLW

It follows from the results of BLW that an Hölder observable f on $[0, 1)$, for every $\varepsilon > 0$, $(f \circ T^n)_{n \geq 1}$ satisfies the ASIP with rate $b_n = n^{\varepsilon}$.

It follows from the results of BLW that an Hölder observable f on $[0,1)$, for every $\varepsilon > 0$, $(f \circ T^n)_{n \geq 1}$ satisfies the ASIP with rate $b_n = n^{\varepsilon}$.

The dependence coefficients used by BLW are well suited to linear processes and their functionals. However they are not well adapted to more general processes such as Markov processes or processes arising from dynamical systems.

Extensions of the work of BLW

It follows from the results of BLW that an Hölder observable f on $[0,1)$, for every $\varepsilon > 0$, $(f \circ T^n)_{n\geq 1}$ satisfies the ASIP with rate $b_n = n^\varepsilon$.

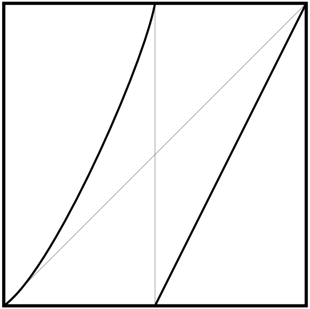
The dependence coefficients used by BLW are well suited to linear processes and their functionals. However they are not well adapted to more general processes such as Markov processes or processes arising from dynamical systems.

In 2018, C.-Dedecker-Merlevède adapted the ideas developed in BLW to obtain sharp conditions for the ASIP with rate $o(n^{1/p})$. For instance, when $X_k = g(\varepsilon_k, , \varepsilon_{k+1}, \ldots)$ with $(\varepsilon_k)_{k\geq 1}$ they obtained conditions relying on $\|X_1 - \tilde{X}^*_{1,k}\|_p$ where $\tilde{X}_{1,k} = g(\varepsilon_1, \ldots, \varepsilon_{k-1}, \varepsilon^*_k, \varepsilon^*_{k+1}, \ldots)$ and $(\varepsilon^*_k)_{k\geq 1}$ is an independent copy of $(\varepsilon_k)_{k\geq 1}$.

A second dynamical example: the LSV map

Let consider the LSV map (Liverani, Saussol and Vaienti, ('99)):

$$\text{for } 0 < \gamma < 1, \quad T(x) = \begin{cases} x(1 + 2^\gamma x^\gamma) & \text{if } x \in [0, 1/2[\\ 2x - 1 & \text{if } x \in [1/2, 1] \end{cases}$$



Graph of f

- There exists a unique absolutely continuous f-invariant probability measure μ on $[0, 1]$, which is equivalent to the Lebesgue measure and whose density h satisfies $0 < c \leq x^\gamma h(x) \leq C < \infty$.

- There exists a unique absolutely continuous f-invariant probability measure μ on $[0,1]$, which is equivalent to the Lebesgue measure and whose density h satisfies $0 < c \le x^\gamma h(x) \le C < \infty$.
- The intermittent behaviour comes from the fact that 0 is a fixed point with $f'(0) = 1$.

- There exists a unique absolutely continuous f-invariant probability measure μ on $[0,1]$, which is equivalent to the Lebesgue measure and whose density h satisfies $0 < c \leq x^\gamma h(x) \leq C < \infty$.
- 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 $(T^n(x))_{n\geq 0}$ stays around 0 for a *long* time.

- There exists a unique absolutely continuous f-invariant probability measure μ on $[0,1]$, which is equivalent to the Lebesgue measure and whose density h satisfies $0 < c \leq x^\gamma h(x) \leq C < \infty$.
- 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 $(T^n(x))_{n\geq 0}$ 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 $\tau\colon Y \to \mathbb{N}$,

$$\tau(x) = \min\{n \geq 1\colon T^n(x) \in Y\}\,.$$

- There exists a unique absolutely continuous f-invariant probability measure μ on $[0,1]$, which is equivalent to the Lebesgue measure and whose density h satisfies $0 < c \le x^\gamma h(x) \le C < \infty$.
- 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 $(T^n(x))_{n\ge 0}$ 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 $\tau\colon Y \to \mathbb{N}$,

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

- We have $C^{-1} n^{-1/\gamma} \le \mathrm{Leb}\,(\tau \ge n) \le C n^{-1/\gamma}$ (Gouezel'04 or Young'99)

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

- Suppose that $f\colon [0,1] \to \mathbb{R}$ is a Hölder continuous observable with $\int f \, d\mu = 0$.
- We have the following polynomial decay for the correlations associated with $(f \circ T^{n-1})_{n\geq 1}$

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

- Suppose that $f\colon [0,1] \to \mathbb{R}$ is a Hölder continuous observable with $\int f \, d\mu = 0$.
- We have the following polynomial decay for the correlations associated with $(f \circ T^{n-1})_{n \geq 1}$

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

- If $\gamma < 1/2$, $n^{-1/2} S_n \to^d N(0, c^2)$ with

$$c^2 = \int f^2 \, d\mu + 2 \sum_{n=1}^{\infty} \int f \, f \circ T^n \, d\mu \qquad (*)$$

- Suppose that $f\colon [0,1] \to \mathbb{R}$ is a Hölder continuous observable with $\int f\, d\mu = 0$.
- We have the following polynomial decay for the correlations associated with $(f \circ T^{n-1})_{n \geq 1}$

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

- If $\gamma < 1/2$, $n^{-1/2} S_n \to^d N(0, c^2)$ with

$$c^2 = \int f^2 \, d\mu + 2 \sum_{n=1}^{\infty} \int f\, f \circ T^n \, d\mu \qquad (*)$$

- What about rates in the ASIP when $\gamma < 1/2$?

Previous results for the LSV map

- The first results were obtained by Melbourne and Nicol ('05) but without explicit rates by using a coupling method due to Philipp and Stout ('75).

Previous results for the LSV map

- The first results were obtained by Melbourne and Nicol ('05) but without explicit rates by using a coupling method due to Philipp and Stout ('75).
- For Hölder continuous or bounded variation observables, using a conditional quantile method, Merlevède-Rio ('12), proved the ASIP with rates

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

for all $\varepsilon > 0$, where $\gamma' = \max\{\gamma, 1/3\}$.

Previous results for the LSV map

- The first results were obtained by Melbourne and Nicol ('05) but without explicit rates by using a coupling method due to Philipp and Stout ('75).
- For Hölder continuous or bounded variation observables, using a conditional quantile method, Merlevède-Rio ('12), proved the ASIP with rates

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

  for all $\varepsilon > 0$, where $\gamma' = \max\{\gamma, 1/3\}$.
- Using an approximation via reverse martingale difference sequences and an ASIP for reverse MDS due to C.-Merlevède ('15), Korepanov-Kosloff-Melbourne ('18) 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 $\varepsilon > 0$ (No way to get better bounds with this method!)

Our results

Theorem (C.-Dedecker-Korepanov-Merlevède (submitted))

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

Theorem (C.-Dedecker-Korepanov-Merlevède (submitted))

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

If $c^2 = 0$, the rate in the ASIP can be improved to $O(1)$. Indeed,in this case φ is a *coboundary* in the sense that $f = u - u \circ T$ with u bounded.

Our results

Theorem (C.-Dedecker-Korepanov-Merlevède (submitted))

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

If $c^2 = 0$, the rate in the ASIP can be improved to $O(1)$. Indeed,in this case φ is a *coboundary* in the sense that $f = u - u \circ T$ with u bounded.

However in general the rates are optimal in the following sense :

Our results

Theorem (C.-Dedecker-Korepanov-Merlevède (submitted))

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

If $c^2 = 0$, the rate in the ASIP can be improved to $O(1)$. Indeed,in this case φ is a *coboundary* in the sense that $f = u - u \circ T$ with u bounded.

However in general the rates are optimal in the following sense :

Proposition (C-D-K-M. ('18))

There exists a Hölder continuous observable f with $\int f \, d\mu = 0$ such that

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

for all Brownian motions $(W_t)_{t\geq 0}$ defined on the same (possibly enlarged) probability space as $(S_n)_{n\geq 0}$.

A third dynamical example: definition and properties

- Let $\kappa \geq 0$. Let $T\colon [0,1] \to [0,1]$ be defined by

$$T(x) = \begin{cases} x(1 + \frac{c}{|\log x|^\kappa}), & x \leq 1/2 \\ 2x - 1, & x > 1/2 \end{cases} \tag{1}$$

with $c = (\log 2)^\kappa$ so that $f(1/2) = 1$.

Define Y and τ as in the LSV map.

A third dynamical example: definition and properties

- Let $\kappa \geq 0$. Let $T\colon [0,1] \to [0,1]$ be defined by

$$T(x) = \begin{cases} x\left(1 + \frac{c}{|\log x|^{\kappa}}\right), & x \leq 1/2 \\ 2x - 1, & x > 1/2 \end{cases} \tag{1}$$

  with $c = (\log 2)^{\kappa}$ so that $f(1/2) = 1$.

  Define Y and τ as in the LSV map.

- There exists $\eta_1, \eta_2 > 0$ such that

$$\mathrm{e}^{-\eta_2 n^{1/(\kappa+1)}} \leq \mathrm{Leb}(\tau \geq n) \leq \mathrm{e}^{-\eta_1 n^{1/(\kappa+1)}} .$$

A third dynamical example: definition and properties

- Let $\kappa \geq 0$. Let $T\colon [0,1] \to [0,1]$ be defined by

$$T(x) = \begin{cases} x(1 + \frac{c}{|\log x|^\kappa}), & x \leq 1/2 \\ 2x - 1, & x > 1/2 \end{cases} \tag{1}$$

  with $c = (\log 2)^\kappa$ so that $f(1/2) = 1$.

  Define Y and τ as in the LSV map.

- There exists $\eta_1, \eta_2 > 0$ such that

$$\mathrm{e}^{-\eta_2 n^{1/(\kappa+1)}} \leq \mathrm{Leb}(\tau \geq n) \leq \mathrm{e}^{-\eta_1 n^{1/(\kappa+1)}}.$$

- There exists a unique absolutely continuous invariant probablity measure μ that is bounded and bounded away from 0 on Y.

A third dynamical example: asip with rate

Theorem (C-D-K-M, in progress)

Let $\kappa > 0$. Let f be a Hölder continuous observable. Then, $(f \circ T^{n-1})_{n\geq 1}$ satisfies the ASIP with rate $b_n = (\log n)^{2+\kappa}$.

A third dynamical example: asip with rate

Theorem (C-D-K-M, in progress)

Let $\kappa > 0$. Let f be a Hölder continuous observable. Then, $(f \circ T^{n-1})_{n\geq 1}$ satisfies the ASIP with rate $b_n = (\log n)^{2+\kappa}$.

Unfortunately, the obtained rate does not seem to be the best possible. To see that first notice that when $\kappa = 0$ one recovers the doubling map. In that case, our process is f applied to a linear process. Taking $f = \mathrm{Identity}$, it is possible to prove by a direct argument that $b_n = \log n$ in the ASIP.

A third dynamical example: asip with rate

Theorem (C-D-K-M, in progress)

Let $\kappa > 0$. Let f be a Hölder continuous observable. Then, $(f \circ T^{n-1})_{n \geq 1}$ satisfies the ASIP with rate $b_n = (\log n)^{2+\kappa}$.

Unfortunately, the obtained rate does not seem to be the best possible. To see that first notice that when $\kappa = 0$ one recovers the doubling map. In that case, our process is f applied to a linear process. Taking $f = \mathrm{Identity}$, it is possible to prove by a direct argument that $b_n = \log n$ in the ASIP.

Also, let us mention that for Harris recurrent geometrically ergodic Markov chains and (only) bounded observables, Merlevède et Rio obtained the rate $b_n = \log n$.

Main ideas of the proof in the case of the LSV

- The main idea is to construct a stationary Markov chain $(g_n, n \in \mathbb{N})$ on a countable space S and an observable $\psi : \Omega \to \mathbb{R}$ (here $\Omega \subset S^{\mathbb{N}}$) such that $\int \psi \, d\mathbb{P}_\Omega = 0$ and setting

$$X_k = \psi(g_{k-1}, g_k, \ldots) \, , \; k \geq 1,$$

the process $(X_k)_{k\geq 1}$ on the probability space $(\Omega, \mathbb{P}_\Omega)$ has the same law as $(f \circ T^{k-1})_{k\geq 1}$ on $([0,1], \mu)$.

Main ideas of the proof in the case of the LSV

- The main idea is to construct a stationary Markov chain $(g_n, n \in \mathbb{N})$ on a countable space S and an observable $\psi : \Omega \to \mathbb{R}$ (here $\Omega \subset S^{\mathbb{N}}$) such that $\int \psi \, d\mathbb{P}_\Omega = 0$ and setting

$$X_k = \psi(g_{k-1}, g_k, \ldots)\,,\ k \geq 1,$$

  the process $(X_k)_{k\geq 1}$ on the probability space $(\Omega, \mathbb{P}_\Omega)$ has the same law as $(f \circ T^{k-1})_{k\geq 1}$ on $([0,1], \mu)$.
- The Markov chain is related to the classical Young towers.

Main ideas of the proof in the case of the LSV

- The main idea is to construct a stationary Markov chain $(g_n, n \in \mathbb{N})$ on a countable space S and an observable $\psi : \Omega \to \mathbb{R}$ (here $\Omega \subset S^{\mathbb{N}}$) such that $\int \psi \, d\mathbb{P}_\Omega = 0$ and setting

$$X_k = \psi(g_{k-1}, g_k, \ldots) \, , \; k \geq 1,$$

the process $(X_k)_{k\geq 1}$ on the probability space $(\Omega, \mathbb{P}_\Omega)$ has the same law as $(f \circ T^{k-1})_{k\geq 1}$ on $([0,1], \mu)$.

- The Markov chain is related to the classical Young towers.
- Recall that $Y =]1/2, 1]$ and $\tau \colon Y \to \mathbb{N}$ be the inducing time $\tau(x) = \min\{n \geq 1 \colon f^n(x) \in Y\}$. Let $F : Y \to Y$ be the induced map: $F(x) = f^{\tau(x)}(x)$. Let α be the partition of Y into the intervals where τ is constant.

Main ideas of the proof in the case of the LSV

- The main idea is to construct a stationary Markov chain $(g_n, n \in \mathbb{N})$ on a countable space S and an observable $\psi : \Omega \to \mathbb{R}$ (here $\Omega \subset S^{\mathbb{N}}$) such that $\int \psi \, d\mathbb{P}_\Omega = 0$ and setting

$$X_k = \psi(g_{k-1}, g_k, \dots) \, , \; k \geq 1,$$

  the process $(X_k)_{k\geq 1}$ on the probability space $(\Omega, \mathbb{P}_\Omega)$ has the same law as $(f \circ T^{k-1})_{k\geq 1}$ on $([0,1], \mu)$.

- The Markov chain is related to the classical Young towers.

- Recall that $Y =]1/2, 1]$ and $\tau \colon Y \to \mathbb{N}$ be the inducing time $\tau(x) = \min\{n \geq 1 \colon f^n(x) \in Y\}$. Let $F : Y \to Y$ 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 $\mathcal{A}$ denote the set of all finite words in the alphabet α, not including the empty word. Denote by $w = a_0 \cdots a_{n-1}$ an element of $\mathcal{A}$. Let also $h \colon \mathcal{A} \to \mathbb{N}$, $h(w) = \tau(a_0) + \cdots + \tau(a_{n-1})$

On the Markov chain

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

On the Markov chain

- Let $S = \{(w, \ell) \in \mathcal{A} \times \mathbb{Z} \colon 0 \leq \ell < h(w)\}$.
- Let $g_0 \in S$ be distributed according to a certain ν and (ε_k) be a sequence of iid r.v. with values in $\mathcal{A}$, distribution $\mathbb{P}_{\mathcal{A}}$ and independent from g_0.

On the Markov chain

- Let $S = \{(w, \ell) \in \mathcal{A} \times \mathbb{Z} \colon 0 \le \ell < h(w)\}$.
- Let $g_0 \in S$ be distributed according to a certain ν and (ε_k) be a sequence of iid r.v. with values in $\mathcal{A}$, distribution $\mathbb{P}_{\mathcal{A}}$ and independent from g_0.
- Then, for any $n \ge 0$, we define

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

On the Markov chain

- Let $S = \{(w, \ell) \in \mathcal{A} \times \mathbb{Z} \colon 0 \leq \ell < h(w)\}$.
- Let $g_0 \in S$ be distributed according to a certain ν and (ε_k) be a sequence of iid r.v. with values in $\mathcal{A}$, distribution $\mathbb{P}_{\mathcal{A}}$ and independent from g_0.
- Then, for any $n \geq 0$, we define

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

- where

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

On the Markov chain

- Let $S = \{(w, \ell) \in \mathcal{A} \times \mathbb{Z} \colon 0 \le \ell < h(w)\}$.
- Let $g_0 \in S$ be distributed according to a certain ν and (ε_k) be a sequence of iid r.v. with values in $\mathcal{A}$, distribution $\mathbb{P}_{\mathcal{A}}$ and independent from g_0.
- Then, for any $n \ge 0$, we define

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

- where

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

- For the LSV maps the constructed Markov chain is aperiodic.

- We are back to prove an ASIP for the partial sums associated with $(X_k)_{k\geq 0}$ where $X_k = \psi(g_k, g_{k+1}, \ldots)$.

- We are back to prove an ASIP for the partial sums associated with $(X_k)_{k\geq 0}$ where $X_k = \psi(g_k, g_{k+1}, \ldots)$.
- The function ψ satisfies the following property: for $a = (g_0, \ldots, g_N, g_{N+1}, \ldots)$ and $b = (g_0, \ldots, g_N, g'_{N+1}, \ldots)$ with $g_{N+1} \neq g'_{N+1}$,

$$|\psi(a) - \psi(b)| \leq C\theta^{\sum_{k=0}^{N} \mathbf{1}_{\{g_k \in S_0\}}} ,$$

where $S_0 = \{(w, 0) : w \in \mathcal{A}\}$ and $\theta \in]0, 1[$.

- We are back to prove an ASIP for the partial sums associated with $(X_k)_{k\geq 0}$ where $X_k = \psi(g_k, g_{k+1}, \ldots)$.
- The function ψ satisfies the following property: for $a = (g_0, \ldots, g_N, g_{N+1}, \ldots)$ and $b = (g_0, \ldots, g_N, g'_{N+1}, \ldots)$ with $g_{N+1} \neq g'_{N+1}$,

$$|\psi(a) - \psi(b)| \leq C\theta^{\sum_{k=0}^{N} \mathbf{1}_{\{g_k \in S_0\}}},$$

where $S_0 = \{(w, 0) : w \in \mathcal{A}\}$ and $\theta \in]0, 1[$.

- With the help of the above property one can prove that there exists a bounded measurable function G_m such that, for any $r \geq 1$,

$$\|X_k - G_m(\varepsilon_{k-m}, \ldots, \varepsilon_{k+m})\|_1 \ll \mathbb{P}(S \geq m) + m^{-r/2}$$

where S is the meeting time

$$S = \inf\{n \geq 0 : g_n = g_n^*\}$$

here g_0^* has distribution ν and is independent of $(g_0, (\varepsilon_k)_{k\geq 1})$ and $g_{n+1}^* = U(g_n^*, \varepsilon_{n+1})$.

- The moments of S can be handled thanks to a result of Lindvall ('79) combined with a result of Korepanov ('18): Since $\text{Leb}\,(\tau \geq n) \leq Cn^{-1/\gamma}$ then $\mathbb{E}(h_{\gamma,\eta}(S)) < \infty$ for any $\eta > 1$ where $h_{\beta,\eta}(x) = x^{(1-\gamma)/\gamma}(log(1+x))^{-\eta}$.

- The moments of S can be handled thanks to a result of Lindvall ('79) combined with a result of Korepanov ('18):
  Since $\mathrm{Leb}\,(\tau \geq n) \leq Cn^{-1/\gamma}$ then $\mathbb{E}(h_{\gamma,\eta}(S)) < \infty$ for any $\eta > 1$ where $h_{\beta,\eta}(x) = x^{(1-\gamma)/\gamma}(log(1+x))^{-\eta}$.
- The $2m$-dependent approximation

$$\|X_k - G_m(\varepsilon_{k-m}, \ldots, \varepsilon_{k+m})\|_1 \ll \mathbb{P}(S \geq m) + m^{-r/2}$$

  allows to adapt the scheme of proof developped by Berkes-Liu-Wu ('14) to prove KMT with rate $o(n^{1/p})$ for functions of iid having a moment of order p, under a weak dependence condition.

- The moments of S can be handled thanks to a result of Lindvall ('79) combined with a result of Korepanov ('18):
  Since $\text{Leb}\,(\tau \geq n) \leq Cn^{-1/\gamma}$ then $\mathbb{E}(h_{\gamma,\eta}(S)) < \infty$ for any $\eta > 1$ where $h_{\beta,\eta}(x) = x^{(1-\gamma)/\gamma}(log(1+x))^{-\eta}$.
- The $2m$-dependent approximation

$$\|X_k - G_m(\varepsilon_{k-m}, \ldots, \varepsilon_{k+m})\|_1 \ll \mathbb{P}(S \geq m) + m^{-r/2}$$

  allows to adapt the scheme of proof developped by Berkes-Liu-Wu ('14) to prove KMT with rate $o(n^{1/p})$ for functions of iid having a moment of order p, under a weak dependence condition.
- Their proof consists first in providing a conditional Gaussian approximation by freezing some part of the $(\varepsilon_k)_k$, making suitable blocks and applying Sakhanenko's '06 result, and after of proceeding to a unconditional Gaussian approximation.

Thank you for your attention!