

INVARIANCE PRINCIPLES AND GAUSSIAN APPROXIMATION FOR STRICTLY STATIONARY PROCESSES

DALIBOR VOLNÝ

ABSTRACT. We show that in any aperiodic and ergodic dynamical system there exists a square integrable process $(f \circ T^i)$ the partial sums of which can be closely approximated by the partial sums of Gaussian i.i.d. random variables. For $(f \circ T^i)$ both weak and strong invariance principles hold.

1. INTRODUCTION AND RESULTS

By $(\Omega, \mathcal{A}, T, \mu)$ we denote an aperiodic and ergodic dynamical system where T is an automorphism of the probability space $(\Omega, \mathcal{A}, T, \mu)$.

If the dynamical system is of positive entropy, we can find a measurable function f such that the random variables $f \circ T^i$ are independent, hence the process $(f \circ T^i)$ is a sequence of i.i.d.; if the entropy is infinite, $f \circ T^i$ can be Gaussian. In the case of positive entropy we thus can find processes $(f \circ T^i)$ for which the central limit theorem or a functional central limit theorem (invariance principle) holds. The process need not be independent: e.g., there exists a vast literature on the CLT for (strictly) stationary processes (cf. [G], [Ha-He], [Vo2])—but the assumptions imply positive entropy.

From the properties of joinings (see [Ru], Thm. 6.11) it follows that if (X_i, Y_i) is a strictly stationary sequence of random vectors where the process (X_i) has zero entropy and the Y_i are independent (identically distributed), then the processes (X_i) and (Y_i) are independent.

The central limit theorem for processes of zero entropy has for some time been an open problem: Denker and Keane quote ([De-K]) a problem of J.-P. Conze: whether there exists a process $(f \circ T^i)$ in a dynamical system of zero entropy, for which the CLT would hold. The question was answered by Burton and Denker ([Bu-De]), who proved that in any aperiodic and ergodic dynamical system there exists a process $(f \circ T^i)$ for which the central limit theorem holds. The result was surprising—the CLT can hold even in such dynamical systems as irrational rotations. This means that even in such dynamical systems, the normalized partial sums of random variables $f \circ T^i$ can be “almost” Gaussian. The CLT was proved in the sense of the convergence of the distributions of $S_n(f)/\|S_n(f)\|_2$, $S_n(f)$ denoting

Received by the editors February 21, 1997.

1991 *Mathematics Subject Classification*. Primary 28D05, 60G10, 60F17, 60F05, 28D20.

Key words and phrases. Zero entropy stationary process, weak invariance principle, strong invariance principle, approximation by Gaussian random variables.

This research has been partially supported by the Grant Agency of the Charles University (Prague), grant #GAUK 6191.

the sum $\sum_{i=0}^{n-1} f \circ T^i$. In [D-L-P-W] a minor error in [Bu-De] is corrected, and the ideas are further developed.

We shall prove that in the aperiodic and ergodic dynamical system $(\Omega, \mathcal{A}, T, \mu)$ we can find a square integrable function f such that the partial sums $S_n(f) = \sum_{i=0}^{n-1} f \circ T$ can be approximated by partial sums of independent and identically distributed Gaussian random variables in a much stronger sense:

Theorem 1. *For every $n = 1, 2, \dots$ there exist independent random variables $Z_{n,1}, \dots, Z_{n,n-1}$ with normal distributions $N(0, 2(\log \log 3 - \log \log 2))$ such that*

$$\max_{1 \leq \ell \leq n} \frac{1}{\sqrt{n}} \|S_\ell(f) - \sum_{j=0}^{\ell-1} Z_{n,j}\|_2 = O\left(\frac{1}{\log n}\right)^{1/2}.$$

Theorem 2. *There exist independent random variables Z_j with normal distributions $N(0, 2(\log \log 3 - \log \log 2))$ such that*

$$\max_{1 \leq \ell \leq n} \frac{1}{\sqrt{n}} \|S_\ell(f) - \sum_{j=0}^{\ell-1} Z_j\|_2 = O\left(\frac{\log \log \log \log n}{\log n}\right)^{1/2}.$$

Using an adequate multiple of f , in both theorems we can get for the functions $Z_{n,j}$, Z_j the standard normal distribution $N(0, 1)$. In order to shorten the formulae in the proofs, however, we use the given formulations.

Let us recall a few definitions (cf. [B] and [Cs-Re]). Suppose that f is a measurable function on Ω . For $t \in [0, 1]$ we define

$$\varphi_n(t) = \frac{1}{\sqrt{n}} \sum_{i=0}^{[tn]-1} f \circ T^i + \frac{tn - [tn]}{\sqrt{n}} f \circ T^{[tn]},$$

$n = 1, 2, \dots$

If the mappings $\varphi_n : \Omega \rightarrow C([0, 1])$ converge in distribution to the Brownian motion process $W(t)$, we say that (for the process $(f \circ T^i)$) the weak invariance principle holds.

If for any vector $(t_1, \dots, t_k) \in [0, 1]^k$

$$(\varphi_n(t_1), \dots, \varphi_n(t_k)) \text{ converge in distribution to } (W(t_1), \dots, W(t_k)),$$

we say that the finite dimensional distributions of φ_n converge to those of W .

In [La1], M. Lacey proved the invariance principle in the sense of the convergence of finite dimensional distributions to the distributions of self-similar processes, but the proofs need a correction. For irrational rotations of the circle the result was given in [La2] (the author also showed that f can be Hölder continuous).

Theorem 3. *For the process $(f \circ T^i)$ the (weak) invariance principle holds.*

If there exists a probability space with a Brownian motion process $W(t)$ and a sequence of random variables (X_i) equally distributed as $(f \circ T^i)$ such that for $S_n = \sum_{i=0}^{n-1} X_i$,

$$\frac{|S_n - W(n)|}{\sqrt{n \log \log n}} \rightarrow 0 \text{ a.s.,}$$

we say that the strong invariance principle holds.

In the paper [Bu-De] the authors conjectured that in any aperiodic and ergodic dynamical system the strong invariance principle with $n^{\delta+1/2}$ replacing the $\sqrt{n \log \log n}$ holds for some process $(f \circ T^i)$, too.

We shall prove the following version of the strong invariance principle:

Theorem 4. *There exist Gaussian independent and identically distributed random variables Z_i such that*

$$\frac{|S_n(f) - \sum_{i=0}^{n-1} Z_i|}{\sqrt{n \log \log n}} \rightarrow 0 \text{ a.s.}$$

2. PROOFS

Using Proposition 1, we shall construct the function f .

Proposition 1. *Let $\epsilon_1, \epsilon_2, \dots$ be positive real numbers, $\alpha_1, \alpha_2, \dots$ real numbers, and d_1, d_2, \dots positive integers.*

Then there exist functions $\bar{f}_1, \bar{f}_2, \dots \in L^2$ with zero means, and mutually independent random variables $X_{k,i}$, $i = 0, \dots, 2d_k$, $k = 1, 2, \dots$, with distributions $N(0, \alpha_k^2)$, such that for each $k = 1, 2, \dots$ and $0 \leq i \leq 2d_k$

$$\|X_{k,i} - \bar{f}_k \circ T^i\|_2 < \epsilon_k.$$

The random variables $X_{k,i}$ are thus i.i.d. in each row, and the rows are mutually independent. The random variables $\bar{f}_k, \bar{f}_k \circ T, \dots, \bar{f}_k \circ T^{2d_k}$ are thus “almost” i.i.d. normally distributed and “almost” independent of the σ -algebra generated by $f_j, f_j \circ T, \dots, f_j \circ T^{2d_j}$, $1 \leq j \leq k-1$.

Proposition 1 is an immediate consequence of

Proposition 2. *Let n_k be positive integers. Let $(X_{k,i})_{i \in \mathbb{Z}}$ be ergodic strictly stationary processes (defined on different probability spaces), and let $\epsilon_k > 0$, $k = 1, 2, \dots$*

Then there exist measurable finite valued functions f_k such that

1. for every $k = 1, 2, \dots$ there exists a random vector $(X'_{k,0}, \dots, X'_{k,n_k})$ with the same distribution as $(X_{k,0}, \dots, X_{k,n_k})$ on Ω ,

$$\mu\{\exists 0 \leq i \leq n_k, |f_k \circ T^i - X'_{k,i}| > \epsilon_k\} < \epsilon_k,$$

and

2. the partitions ξ generated by $f_j \circ T^i$, $0 \leq i \leq n_j$, $1 \leq j \leq k-1$, and η generated by $f_k \circ T^i$, $0 \leq i \leq n_k$, are ϵ -independent in the sense that

$$\sum_{A \in \eta, B \in \xi, \mu(B) > 0} |\mu(A) - \mu(A|B)| < \epsilon_k.$$

Remark. Usually, ϵ -independence is defined in a weaker sense—the sum is taken over a set of $B \in \xi$ whose union has measure greater than $1 - \epsilon_k$.

Proof of Proposition 2. For proving Proposition 2 it is enough to show that when given a finite measurable partition ξ_1 of Ω , $\epsilon > 0$, $n \in \mathbb{N}$, and an ergodic strictly stationary process (X_i) , there exist a finite valued function f and a random vector (X'_0, \dots, X'_n) equally distributed as (X_0, \dots, X_n) such that

$$(1) \quad \mu\{\exists 0 \leq i \leq n, |f \circ T^i - X'_i| > \epsilon\} < \epsilon$$

and

$$(2) \quad \sum_{A \in \eta_1, B \in \xi_1, \mu(B) > 0} |\mu(A) - \mu(A|B)| < \epsilon,$$

where η_1 is the partition generated by $f, f \circ T, \dots, f \circ T^n$.

Without loss of generality we can assume that the X_i are finitely valued.

Let $\xi_1 = \{A_1, \dots, A_K\}$. Let N be a positive integer, $F, TF, \dots, T^N F$ a Rokhlin tower. By π_1 we denote the partition generated by $F, TF, \dots, T^N F$.

Let $(\Omega', \mathcal{C}, S, \nu)$ be another ergodic dynamical system with a measurable function \bar{X} on Ω' such that the distributions of the random vectors $(\bar{X}, \bar{X} \circ S, \dots, \bar{X} \circ S^n)$ and (X_0, \dots, X_n) are the same; without loss of generality we can assume that ν is aperiodic. By η_2 we denote the partition of Ω' generated by $\bar{X}, \dots, \bar{X} \circ S^n$.

In the dynamical system $(\Omega', \mathcal{C}, S, \nu)$ there is a Rokhlin tower $E, SE, \dots, S^N E$ such that $\nu(E) = \mu(F)$, generating a partition π_2 , and a partition ξ_2 such that the partitions $\xi_1 \vee \pi_1$ generated by ξ_1, π_1 , and $\xi_2 \vee \pi_2$ generated by ξ_2, π_2 , respectively, are equally distributed.

On $(\Omega, \mathcal{A}, T, \mu)$ there exists a partition η_1 such that the distributions of the partitions $\xi_2 \vee \pi_2 \vee \eta_2$ and $\xi_1 \vee \pi_1 \vee \eta_1$ are the same. By f we denote the function which corresponds to \bar{X} . If we choose the Rokhlin tower so that $\mu(\bigcup_{i=n}^{N-n} T^i F) > 1 - \epsilon$, the distribution of $f, f \circ T, \dots, f \circ T^n$ will be close enough to that of $\bar{X}, \bar{X} \circ S, \dots, \bar{X} \circ S^n$, and (1) will be guaranteed. Similarly we can guarantee (2). \square

By U we denote the unitary operator on L^2 defined by $Uf = f \circ T$; I denotes the identity operator $If = f$. For every function f and positive integer n we denote

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

For $k = 1, 2, \dots$ we define

$$\begin{aligned} d_k &= 3^k, \\ p_k &= 2^k, \\ \alpha_k &= 1/(p_k \sqrt{k}), \\ \epsilon_k &= 6^{-3k}. \end{aligned} \tag{3}$$

By Proposition 1 there exist functions $\bar{f}_1, \bar{f}_2, \dots \in L^2$ with zero means such that for each $k = 1, 2, \dots$ and $0 \leq i \leq 2d_k$

$$\|X_{k,i} - \bar{f}_k \circ T^i\|_2 < \epsilon_k,$$

where $X_{k,i}$, $i = 0, \dots, 2d_k$, $k = 1, 2, \dots$, are mutually independent random variables with distributions $N(0, \alpha_k^2)$. Put

$$\begin{aligned} f_k &= \sum_{i=0}^{p_k-1} U^i \bar{f}_k - U^{d_k} \sum_{i=0}^{p_k-1} U^i \bar{f}_k, \\ f &= \sum_{k=1}^{\infty} f_k. \end{aligned}$$

We define

$$\tilde{U}X_{k,i} = X_{k,i+1}$$

whenever $X_{k,i}, X_{k,i+1}$ are defined.

For $j = 0, 1, \dots, n$ let us denote

$$Y_{n,j} = \sum_{k: p_k < n < d_k} p_k (X_{k,j} - X_{k,d_k+j}),$$

$$Z_{n,j} = Y_{n,j} \frac{\sqrt{2(\log \log 3 - \log \log 2)}}{\|Y_{n,j}\|_2}.$$

The random variables $Y_{n,j}$, $0 \leq j \leq n-1$, are mutually independent and Gaussian.

Lemma 1.

$$(i) \quad \left| \|Y_{n,j}\|_2 - \sqrt{2(\log \log 3 - \log \log 2)} \right| = O\left(\frac{1}{\log n}\right),$$

$$(ii) \quad \left| \|Y_{n,j}\|_2 - \sqrt{2(\log \log 3 - \log \log 2)} \right| = O\left(\frac{1}{\log n}\right)$$

for every $j = 0, 1, \dots, n$.

First let us notice that

$$(4) \quad \left| \sum_{\log n / \log 3 < k < \log n / \log 2} \frac{1}{k} - \left(\log \left(\frac{\log n}{\log 2} \right) - \log \left(\frac{\log n}{\log 3} \right) \right) \right| \leq \frac{3 \log 3}{\log n}.$$

(4) follows from the fact that for any positive integers $a \leq b$ we have

$$\left| \sum_{k=a}^b \frac{1}{k} - \int_a^b \frac{1}{x} dx \right| \leq \sum_{k=a}^b \left(\frac{1}{k} - \frac{1}{k+1} \right) \leq \frac{1}{a}.$$

Proof of Lemma 1. From the independence of the $X_{k,i}$ it follows that

$$\left\| \sum_{j=0}^{n-1} Y_{n,j} \right\|_2^2 = 2n \sum_{k: p_k < n < d_k} \alpha_k^2 p_k^2 = 2n \sum_{\log n / \log 3 < k < \log n / \log 2} \frac{1}{k}.$$

From (4) and

$$\log \left(\frac{\log n}{\log 2} \right) - \log \left(\frac{\log n}{\log 3} \right) = \log \log 3 - \log \log 2$$

it follows that

$$\left| \frac{1}{n} \left\| \sum_{j=0}^{n-1} Y_{n,j} \right\|_2^2 - 2(\log \log 3 - \log \log 2) \right| = O\left(\frac{1}{\log n}\right).$$

Because $Y_{n,0}, \dots, Y_{n,n}$ are independent and identically distributed,

$$\left| \|Y_{n,j}\|_2^2 - 2(\log \log 3 - \log \log 2) \right| = O\left(\frac{1}{\log n}\right)$$

and

$$(i) \quad \left| \|Y_{n,j}\|_2 - \sqrt{2(\log \log 3 - \log \log 2)} \right| = O\left(\frac{1}{\log n}\right)$$

for every $j = 0, 1, \dots, n$. From

$$\begin{aligned} \left| \|Y_{n,j}\| - \|Z_{n,j}\|_2 \right| &\leq \|Y_{n,j} - Z_{n,j}\|_2 \leq \left| \|Y_{n,j}\|_2 - \sqrt{2(\log \log 3 - \log \log 2)} \right| \\ &= O(1/\log n) \end{aligned}$$

we get (ii). □

We denote

$$f' = \sum_{k: d_k \leq n} f_k, \quad f'' = \sum_{k: p_k < n < d_k} f_k, \quad f''' = \sum_{k: n \leq p_k} f_k.$$

For $j = 1, 2, \dots$ we thus have

$$S_j(f) = S_j(f') + S_j(f'') + S_j(f''').$$

In Lemmas 4-6 we shall approximate the partial sums $S_j(f')$, $S_j(f'')$, $S_j(f''')$ by Gaussian random variables. We shall prove Theorems 1-4 when using a double approximation: first we'll approximate $S_n(f)$ by Gaussian random variables constructed by $X_{n,i}$ and $Y_{n,i}$, and the Gaussian character of the approximating variables will then let us estimate their difference from the partial sums of the $Z_{n,i}$ and the Z_i (the Z_i will be defined later).

In the computations we shall use

Lemma 2. *Let $a > 1$, $p > 0$, $c > 0$. Then there exists a constant $K < \infty$ such that*

$$\sum_{k: a^{ck} \leq n} \frac{a^k}{k^p} \leq K \frac{n^{1/c}}{(\log n)^p}.$$

Proof. In fact,

$$\begin{aligned} \sum_{k: a^{ck} \leq n} \frac{a^k}{k^p} &\leq \sum_{k=1}^{\lfloor \frac{1}{2c} \log_a n \rfloor} a^k + \sum_{k=\lfloor \frac{1}{2c} \log_a n \rfloor + 1}^{\lfloor \frac{1}{c} \log_a n \rfloor} \frac{a^k}{(\frac{1}{2c} \log_a n)^p} \\ &\leq \frac{a^{1+\lfloor \frac{1}{2c} \log_a n \rfloor} - 1}{a - 1} + \frac{a^{1+\lfloor \frac{1}{c} \log_a n \rfloor} - 1}{a - 1} \frac{1}{(\frac{1}{2c} \log_a n)^p} \\ &\leq \frac{an^{\frac{1}{2c}} - 1}{a - 1} + \frac{an^{\frac{1}{c}} - 1}{a - 1} \frac{1}{(\frac{1}{2c} \log_a n)^p}. \end{aligned}$$

($[x]$ denotes the integer part of x .) □

The validity of the invariance principle is equivalent (see [B], Thm. 8.1) to the convergence of finite dimensional distributions of the $C([0, 1])$ -valued random variables φ_n to that of the Brownian motion process W , and to the tightness of the sequence of the distributions of the φ_n .

The convergence of finite dimensional distributions follows from Theorem 1. For proving the tightness it is sufficient (see [B], Thm. 8.4) to find for every $\epsilon > 0$ a $\lambda > 1$ such that for all n sufficiently big,

$$(5) \quad \mu \left(\max_{1 \leq j \leq n} |S_j(f)| > \lambda \sqrt{n} \right) < \frac{\epsilon}{\lambda^2}.$$

Lemma 3. *Let $\lambda > 0$, and let Z_0, \dots, Z_{n-1} be square integrable random variables. Then*

$$(i) \quad \lambda^2 \mu \left(\max_{0 \leq j \leq n-1} |Z_j| > \lambda \sqrt{n} \right) \leq \frac{1}{n} \sum_{j=0}^{n-1} \|Z_j\|_2^2.$$

Let $h_{n,j} = g_{n,j} - U g_{n,j}$, $1 \leq j \leq n$, and let each $g_{n,j}$ be normally distributed with zero mean and variance less than or equal to $K n / \log n$, $n = 1, 2, \dots$. Then,

uniformly for all n ,

$$(ii) \quad \lim_{\lambda \rightarrow \infty} \sup_n \lambda^2 \mu \left(\max_{1 \leq j \leq n} |S_j(h_{n,j})| > \lambda \sqrt{n} \right) = 0.$$

In the case when for all $1 \leq j \leq n$, $g_{n,j} = g_n$ and $h_{n,j} = h_n$ we have:

Let $h_n = g_n - U g_n$, and let each g_n be normally distributed with zero mean and variance less than or equal to $K n / \log n$, $n = 1, 2, \dots$. Then, uniformly for all n ,

$$\lim_{\lambda \rightarrow \infty} \sup_n \lambda^2 \mu \left(\max_{1 \leq j \leq n} |S_j(h_n)| > \lambda \sqrt{n} \right) = 0.$$

Proof of Lemma 3. By the Chebyshev inequality we have

$$\mu \left(\max_{0 \leq j \leq n-1} |Z_j| > \lambda \sqrt{n} \right) \leq \sum_{j=0}^{n-1} \mu(|Z_j| > \lambda \sqrt{n}) \leq \sum_{j=0}^{n-1} \frac{\|Z_j\|_2^2}{n \lambda^2}.$$

From this we get (i). (ii) follows from

$$\begin{aligned} \mu \left(\max_{1 \leq j \leq n} |S_j(h_{n,j})| > \lambda \sqrt{n} \right) &= \mu \left(\max_{1 \leq j \leq n} |g_{n,j} - U^j g_{n,j}| > \lambda \sqrt{n} \right) \\ &\leq \mu \left(\max_{1 \leq j \leq n} |U^j g_{n,j}| > \lambda \sqrt{n}/2 \right) \leq n \mu(|g_{n,j}| > \lambda \sqrt{n}/2) \\ &\leq n \sqrt{2/\pi} \int_{\lambda \sqrt{\log n}/(2\sqrt{K})}^{\infty} e^{-x^2/2} dx \leq n n^{-\lambda^2/4K}. \end{aligned}$$

□

Lemma 4. (i) For every $n = 1, 2, \dots$ there exist random variables $\tilde{g}'_{(n)}, \hat{g}'_{(n)}$ such that for every $1 \leq j \leq n$

$$\begin{aligned} S_j(f') &= \tilde{g}'_{(n)} + \hat{g}'_{(n)} - (\tilde{g}'_{(n)} + \hat{g}'_{(n)}) \circ T^j, \\ \|\tilde{g}'_{(n)}\|_2 &= O \left(\frac{n}{\log n} \right)^{1/2}, \quad \|\hat{g}'_{(n)}\|_2 \leq 1, \end{aligned}$$

$\tilde{g}'_{(n)}$ is Gaussian. Thus,

$$\|S_j(f')\|_2^2 = O \left(\frac{n}{\log n} \right).$$

(ii) For every $\epsilon > 0$ there exists a $\lambda > 1$ such that, for all n sufficiently big,

$$\mu \left(\max_{1 \leq j \leq n} |S_j(f')| > \lambda \sqrt{n} \right) < \frac{\epsilon}{\lambda^2}.$$

Proof. Each of the functions f_k is a coboundary with a transfer function

$$g_k = \sum_{j=0}^{d_k-1} \sum_{i=0}^{p_k-1} U^{i+j} \bar{f}_k.$$

Let us denote

$$\tilde{g}_k = \sum_{j=0}^{d_k-1} \sum_{i=0}^{p_k-1} X_{k,i+j}, \quad \hat{g}_k = g_k - \tilde{g}_k;$$

we have

$$\|\hat{g}_k\|_2 \leq d_k p_k \epsilon_k$$

and, as $p_k \leq d_k$,

$$\|\tilde{g}_k\|_2^2 \leq \alpha_k^2 p_k^2 (d_k + p_k) = \frac{d_k + p_k}{k}.$$

Let us define

$$\tilde{g}' = \tilde{g}'_{(n)} = \sum_{k: d_k \leq n} \tilde{g}_k, \quad \hat{g}' = \hat{g}'_{(n)} = \sum_{k: d_k \leq n} \hat{g}_k.$$

We have

$$\|\hat{g}'\|_2 \leq \sum_{k: d_k \leq n} p_k d_k \epsilon_k \leq \sum_{k=1}^{\infty} \frac{1}{6^{2k}} < 1$$

(recall that by (3) $\epsilon_k = 6^{-3k}$), and from Lemma 2 (with $a = \sqrt{3}$, $c = 2$, $p = 1/2$) it follows that

$$\|\tilde{g}'\|_2 = O\left(\frac{n}{\log n}\right)^{1/2}.$$

Because $S_j(f') = \tilde{g}' + \hat{g}' - (\tilde{g}' + \hat{g}') \circ T^j$, we can estimate the L^2 norm of $S_j(f')$ by

$$\|S_j(f')\|_2 = O\left(\frac{n}{\log n}\right)^{1/2} + 2 \sum_{k: d_k \leq n} d_k p_k \epsilon_k = O\left(\frac{n}{\log n}\right)^{1/2},$$

which finishes the proof of (i).

Let us prove the tightness (ii).

The functions f_k are coboundaries with bounded transfer functions. For any fixed k_0 , the tightness condition (ii) is fulfilled for $\sum_{k=1}^{k_0} f_k$.

We replace f' by $\sum_{k: k_0 \geq k, d_k \leq n} f_k$ (and denote it \tilde{f}'). In the same way we replace g' , \tilde{f}' , and \tilde{g}' . Let us denote

$$S_j(\tilde{f}') = \sum_{i=0}^{j-1} \tilde{U}^i \tilde{f}' = \tilde{g}' - \tilde{U}^j \tilde{g}'.$$

By Lemma 3(iii) we have

$$\lim_{\lambda \rightarrow \infty} \sup_n \lambda^2 \mu\left(\max_{1 \leq j \leq n} |S_j(\tilde{f}')| > \lambda \sqrt{n}\right) = 0,$$

and by Lemma 3(i) we get

$$\begin{aligned} & \lambda^2 \mu\left(\max_{1 \leq j \leq n} |S_j(f') - S_j(\tilde{f}')| > \lambda \sqrt{n}\right) \\ & \leq \frac{1}{n} \sum_{i=0}^{j-1} \|(g' - \tilde{g}') - (U^j g' - \tilde{U}^j \tilde{g}')\|_2^2 \leq 4 \|\hat{g}'\|_2^2. \end{aligned}$$

We have

$$\|\hat{g}'\|_2 \leq \sum_{k: k \geq k_0, d_k \leq n} p_k d_k \epsilon_k < 26^{-2k_0}.$$

Hence by Lemma 3 for every $\lambda > 0$

$$\lim_{k_0 \rightarrow \infty} \lambda^2 \mu\left(\max_{1 \leq j \leq n} |S_j(f') - S_j(\tilde{f}')| > \lambda \sqrt{n}\right) = 0.$$

This finishes the proof of (ii). \square

Lemma 5. (i) For every $n = 1, 2, \dots$ and $1 \leq j \leq n$ there exist Gaussian random variables $\tilde{C}_{*,j}$, $\tilde{D}_{*,j}$, $\tilde{E}_{*,j}$ and random variables $\hat{C}_{*,j}$, $\hat{D}_{*,j}$, $\hat{E}_{*,j}$ such that

$$\begin{aligned} S_j(f''') &= \tilde{C}_{*,j} + \tilde{D}_{*,j} + \tilde{E}_{*,j} + \hat{C}_{*,j} + \hat{D}_{*,j} + \hat{E}_{*,j}, \\ \|\tilde{C}_{*,j}\|_2^2 &= O\left(\frac{1}{\log n}\right), \quad \|\tilde{D}_{*,j} + \tilde{E}_{*,j}\|_2^2 = O\left(\frac{n}{\log n}\right), \\ \|\hat{C}_{*,1}\|_2 &= O\left(\frac{1}{n^2}\right), \quad \|\tilde{D}_{*,j} + \tilde{E}_{*,j}\|_2 = O\left(\frac{1}{n^2}\right). \end{aligned}$$

We thus (for $1 \leq j \leq n$) have

$$\|S_j(f''')\|_2^2 = O\left(\frac{n}{\log n}\right).$$

(ii) For every $\epsilon > 0$ there exists a $\lambda > 1$ such that, for all n sufficiently big,

$$\mu\left(\max_{1 \leq j \leq n} |S_j(f''')| > \lambda\sqrt{n}\right) < \frac{\epsilon}{\lambda^2}.$$

Proof. The first statement can be proved easily. In order to prove the tightness at the same time, we'll adopt a more complicated approach.

For $p_k \geq n$ and $1 \leq j \leq n$ we have

$$S_j(f_k) = \sum_{\ell=0}^{j-1} \sum_{i=0}^{p_k-1} U^{i+\ell}(\bar{f}_k - U^{d_k} \bar{f}_k) = C_{k,j} + D_{k,j} + E_{k,j}$$

where

$$\begin{aligned} C_{k,j} &= jC_{k,1} = j \sum_{i=n}^{p_k-1} U^i(\bar{f}_k - U^{d_k} \bar{f}_k), \\ D_{k,j} &= \sum_{i=0}^{j-2} (i+1)U^i(\bar{f}_k - U^{d_k} \bar{f}_k) + \sum_{i=p_k}^{p_k+j-2} (j+p_k-1-i)U^i(\bar{f}_k - U^{d_k} \bar{f}_k), \\ E_{k,j} &= j \sum_{i=j-1}^{n-1} U^i(\bar{f}_k - U^{d_k} \bar{f}_k). \end{aligned}$$

Let us denote

$$C_{*,j} = jC_{*,1} = \sum_{k: p_k \geq n} C_{k,j}, \quad D_{*,j} = \sum_{k: p_k \geq n} D_{k,j}, \quad E_{*,j} = \sum_{k: p_k \geq n} E_{k,j};$$

we thus have

$$S_j(f''') = C_{*,j} + D_{*,j} + E_{*,j}.$$

We shall prove both statements of the lemma for $C_{*,j}$, $D_{*,j}$, $E_{*,j}$, separately.

Replacing $U^i \bar{f}_k$ by $X_{k,i}$, we define $\tilde{C}_{k,j}$, $\tilde{D}_{k,j}$, $\tilde{E}_{k,j}$, and $\tilde{C}_{*,j}$, $\tilde{D}_{*,j}$, $\tilde{E}_{*,j}$. We denote $\hat{C}_{k,j} = C_{k,j} - \tilde{C}_{k,j}$, $\hat{D}_{k,j} = D_{k,j} - \tilde{D}_{k,j}$, $\hat{E}_{k,j} = E_{k,j} - \tilde{E}_{k,j}$, and $\hat{C}_{*,j} = C_{*,j} - \tilde{C}_{*,j}$, $\hat{D}_{*,j} = D_{*,j} - \tilde{D}_{*,j}$, $\hat{E}_{*,j} = E_{*,j} - \tilde{E}_{*,j}$.

From the independence of the $X_{k,i}$ it follows that $\|\tilde{C}_{k,j}\|_2^2 \leq 2j^2 p_k \alpha_k^2 = 2j^2 2^{-k}/k$. Hence

$$\|\tilde{C}_{*,1}\|_2^2 = \sum_{k \geq \log_2 n} \|\tilde{C}_{k,1}\|_2^2 \leq \frac{4}{n \log_2 n}.$$

We have

$$\|\hat{C}_{*,1}\|_2 \leq 2 \sum_{k \geq \log_2 n} p_k \epsilon_k = O\left(\frac{1}{n^5}\right).$$

Hence,

$$\|C_{*,1}\|_2 \leq \frac{2}{\sqrt{n \log n}} + 2 \sum_{k \geq \log_2 n} p_k \epsilon_k < \frac{4}{\sqrt{n \log n}}.$$

Therefore,

$$(6) \quad \|C_{*,j}\|_2 < \frac{4j}{\sqrt{n \log n}}$$

and

$$(7) \quad \lambda^2 \mu\left(\max_{1 \leq j \leq n} |C_{*,j}| > \lambda \sqrt{n}\right) = \lambda^2 \mu(n|C_{*,1}| > \lambda \sqrt{n}) \leq \frac{16}{\log n} \xrightarrow{n \rightarrow \infty} 0.$$

$\tilde{D}_{k,j} + \tilde{E}_{k,j}$ has the same distribution as $\sum_{\ell=0}^{j-1} \tilde{U}^\ell \tilde{h}_k$, where

$$\tilde{h}_k = \sum_{i=0}^{n-1} X_{k,i} - \tilde{U}^{2n} \sum_{i=0}^{n-1} X_{k,i},$$

and $\tilde{D}_{*,j} + \tilde{E}_{*,j}$ has the same distribution as $\sum_{\ell=0}^{j-1} \sum_{k \geq \log_2 n} \tilde{U}^\ell \tilde{h}_k$.

The function \tilde{h}_k is a coboundary with a transfer function

$$\tilde{g}_k = \sum_{u=0}^{2n-1} \sum_{i=0}^{n-1} X_{k,u+i}.$$

Hence

$$\tilde{D}_{k,j} + \tilde{E}_{k,j} = \tilde{g}_k - \tilde{U}^j \tilde{g}_k.$$

Let us denote

$$\tilde{G} = \sum_{k: n \leq p_k} \tilde{g}_k.$$

We have $\|\tilde{g}_k\|_2^2 \leq 2n^3 \alpha_k^2$, and hence

$$\|\tilde{G}\|_2^2 \leq \sum_{k: n \leq p_k} 2n^3 \frac{1}{k 2^{2k}} \leq \sum_{k=\lceil \log_2 n \rceil}^{\infty} 2n^3 \frac{1}{k 2^{2k}} \leq \frac{4n}{\log_2 n}.$$

Therefore

$$(8) \quad \|\tilde{D}_{*,j} + \tilde{E}_{*,j}\|_2^2 = \|\tilde{G} - \tilde{U}^j \tilde{G}\|_2^2 \leq \frac{16n}{\log_2 n}.$$

Because

$$(9) \quad \|\hat{D}_{*,j} + \hat{E}_{*,j}\|_2 \leq 6 \sum_{k: n \leq p_k} p_k^2 \epsilon_k = 6 \sum_{k: n \leq p_k} 4^k \epsilon_k = O\left(\frac{1}{n^5}\right),$$

we have

$$\|D_{*,j} + E_{*,j}\|_2^2 = O\left(\frac{n}{\log n}\right).$$

This together with (6) proves (i).

Let us prove (ii).

From Lemma 3(iii) and (8) it follows that

$$\lim_{\lambda \rightarrow \infty} \sup_n \lambda^2 \mu \left(\max_{1 \leq j \leq n} |\tilde{D}_{*,j} + \tilde{E}_{*,j}| > \lambda \sqrt{n} \right) \rightarrow 0,$$

and by (9) and Lemma 3(i) we have

$$\lim_{\lambda \rightarrow \infty} \lim_{n \rightarrow \infty} \lambda^2 \mu \left(\max_{1 \leq j \leq n} |\hat{D}_{*,j} + \hat{E}_{*,j}| > \lambda \sqrt{n} \right) = 0.$$

Hence

$$\lim_{\lambda \rightarrow \infty} \lim_{n \rightarrow \infty} \lambda^2 \mu \left(\max_{1 \leq j \leq n} |D_{*,j} + E_{*,j}| > \lambda \sqrt{n} \right) = 0.$$

From this and (7) it follows that

$$\lim_{\lambda \rightarrow \infty} \lim_{n \rightarrow \infty} \lambda^2 \mu \left(\max_{1 \leq j \leq n} |S_j(f''')| > \lambda \sqrt{n} \right) = 0.$$

□

Lemma 6. (i) For every $n = 1, 2, \dots$ and $1 \leq m \leq n$ there exist Gaussian random variables S''_{m+} , $\tilde{S}_{m,1}$, $\tilde{S}_{m,2}$, and random variables \hat{S}''_{m+} , $\hat{S}_{m,1}$, $\hat{S}_{m,2}$ such that

$$\begin{aligned} S_m(f'') &= S''_{m+} + \tilde{S}_{m,1} + \tilde{S}_{m,2} + \hat{S}''_{m+} + \hat{S}_{m,1} + \hat{S}_{m,2}, \\ \|S''_{m+}\|_2^2 &= O\left(\frac{n}{\log_2 n}\right), \quad \|\tilde{S}_{m,2}\|_2^2 = O\left(\frac{n}{\log_2 n}\right), \\ \max_{1 \leq m \leq n} \frac{1}{n} \left\| \sum_{j=0}^{m-1} Y_{n,j} - \tilde{S}_{m,1} \right\|_2^2 &= O\left(\frac{n}{\log n}\right), \\ \|\hat{S}''_{m+}\|_2 &= O\left(\frac{1}{n^2}\right), \quad \|\hat{S}_{m,1}\|_2 = O\left(\frac{1}{n^2}\right), \quad \|\hat{S}_{m,2}\|_2 = O\left(\frac{1}{n^2}\right). \end{aligned}$$

We thus have

$$\|S_m(f'') - \sum_{j=0}^{m-1} Y_{n,j}\|_2^2 = O\left(\frac{n}{\log n}\right).$$

(ii) For every $\epsilon > 0$ there exists a $\lambda > 1$ such that, for all n sufficiently big,

$$\mu \left(\max_{1 \leq j \leq n} |S_j(f'')| > \lambda \sqrt{n} \right) < \frac{\epsilon}{\lambda^2}.$$

Proof. Let $p_k \leq n \leq d_k$ and $1 \leq m \leq n$. For $0 \leq m \leq n-1$ we define

$$\begin{aligned} f''_{m-} &= \sum_{k: p_k < n < d_k, p_k < m} \sum_{i=0}^{p_k-1} U^i(\bar{f}_k - U^{d_k} \bar{f}_k), \\ f''_{m+} &= f'' - f''_{m-}, \end{aligned}$$

so that

$$S_m(f'') = S_m(f''_{m-}) + S_m(f''_{m+}).$$

For proving

$$\|S_n(f'') - \sum_{j=0}^{n-1} Y_{n,j}\|_2^2 = O\left(\frac{n}{\log n}\right)$$

and

$$\lim_{\lambda \rightarrow \infty} \limsup_{n \rightarrow \infty} \lambda^2 \mu \left(\max_{1 \leq m \leq n} |S_m(f'')| > \lambda \sqrt{n} \right) = 0$$

it is sufficient to show corresponding statements for the sequences $S_m(f''_{m+})$, $S_m(f''_{m-}) - \sum_{j=0}^{m-1} Y_{m,j}$ and $S_m(f''_{m-})$, respectively.

Estimation for $S_m(f''_{m+})$. We have

$$S_m(f''_{m+}) = \sum_{j=0}^{m-1} \sum_{k: p_k < n < d_k, m \leq p_k} \sum_{i=0}^{p_k-1} U^{i+j} (\bar{f}_k - U^{d_k} \bar{f}_k).$$

Let us denote

$$S''_{m+} = \sum_{j=0}^{m-1} \sum_{k: p_k < n < d_k, m \leq p_k} \sum_{i=0}^{p_k-1} (X_{k,i+j} - X_{k,d_k+i+j}),$$

$$\hat{S}''_{m+} = S_m(f''_{m+}) - S''_{m+}.$$

S''_{m+} has the same distribution as

$$\tilde{S}''_{m+} = \sum_{j=0}^{m-1} \sum_{k: p_k < n < d_k, m \leq p_k} \sum_{i=0}^{p_k-1} (X_{k,i+j} - X_{k,2p_k+i+j}).$$

Recall that $\tilde{U}X_{k,i} = X_{k,i+1}$ whenever $X_{k,i}$ and $X_{k,i+1}$ are defined; we thus get

$$\sum_{k: p_k < n < d_k, m \leq p_k} \sum_{i=0}^{p_k-1} (X_{k,i+j} - X_{k,p_k+i+j}) = \tilde{H}_{m+} - \tilde{U}\tilde{H}_{m+},$$

where

$$\tilde{H}_{m+} = \sum_{k: p_k < n < d_k, m \leq p_k} \sum_{j=0}^{2p_k-1} \sum_{i=0}^{p_k-1} X_{k,i+j}.$$

Hence

$$\tilde{S}''_{m+} = \tilde{H}_{m+} - \tilde{U}^m \tilde{H}_{m+}.$$

We have

$$\|\tilde{H}_{m+}\|_2^2 \leq \sum_{k: p_k < n < d_k, m \leq p_k} 4p_k^3 \alpha_k^2 = 4 \sum_{k: p_k < n < d_k, m \leq p_k} \frac{2^k}{k}.$$

By Lemma 2 (with $a = 2$, $p = 1$, $c = 1$) there exists $K < \infty$ such that

$$\|\tilde{H}_{m+}\|_2^2 \leq K \frac{n}{\log_2 n}$$

for all $m = 1, 2, \dots$; hence

$$(10) \quad \|S''_{m+}\|_2^2 = \|\tilde{S}''_{m+}\|_2^2 = O\left(\frac{n}{\log_2 n}\right).$$

From Lemma 3(ii) it follows that

$$(11) \quad \lim_{\lambda \rightarrow \infty} \lim_{n \rightarrow \infty} \lambda^2 \mu \left(\max_{1 \leq m \leq n} |S''_{m+}| > \lambda \sqrt{n}/2 \right) = 0.$$

We have (cf. (1))

$$(12) \quad \|\hat{S}_{m+}''\|_2 \leq 2 \sum_{k: p_k < n < d_k, m \leq p_k} m p_k \epsilon_k \leq 2 \sum_{\log_3 n < k < \log_2 n} p_k^2 \epsilon_k = O\left(\frac{1}{n^5}\right);$$

hence by Lemma 3(i)

$$(13) \quad \lim_{\lambda \rightarrow \infty} \lim_{n \rightarrow \infty} \lambda^2 \mu\left(\max_{1 \leq m \leq n} |\hat{S}_{m+}''| > \lambda \sqrt{n}\right) = 0.$$

Estimation for $S_m(f_{m-}'')$. For $p_k < m$ we have

$$\begin{aligned} \sum_{j=0}^{m-1} \sum_{i=0}^{p_k-1} U^{i+j} \bar{f}_k &= \sum_{j=0}^{p_k-2} (j+1) U^j \bar{f}_k + p_k \sum_{j=p_k-1}^{m-1} U^j \bar{f}_k \\ &\quad + \sum_{j=m}^{m+p_k-2} (m+p_k-j-1) U^j \bar{f}_k. \end{aligned}$$

Let us denote

$$\begin{aligned} S_{m,1} &= \sum_{k: p_k < n < d_k, m > p_k} p_k \sum_{j=p_k-1}^{m-1} U^j (\bar{f}_k - U^{d_k} \bar{f}_k), \\ S_{m,2} &= \sum_{k: p_k < n < d_k, m > p_k} \left(\sum_{j=0}^{p_k-2} (j+1) U^j (\bar{f}_k - U^{d_k} \bar{f}_k) \right. \\ &\quad \left. + \sum_{j=m}^{m+p_k-2} (m+p_k-j-1) U^j (\bar{f}_k - U^{d_k} \bar{f}_k) \right). \end{aligned}$$

Replacing $U^j \bar{f}_k$ by $X_{k,j}$, we define $\tilde{S}_{m,1}$ and $\tilde{S}_{m,2}$:

$$\hat{S}_{m,1} = S_{m,1} - \tilde{S}_{m,1}, \quad \hat{S}_{m,2} = S_{m,2} - \tilde{S}_{m,2}.$$

For simplicity of notation (and without losing generality), in the definitions of $\tilde{S}_{m,2}$ and $\hat{S}_{m,2}$ we replace $U^j (\bar{f}_k - U^{d_k} \bar{f}_k)$ by $U^j \bar{f}_k$ and $X_{k,j} - X_{k,d_k+j}$ by $X_{k,j}$. We thus have e.g.

$$\tilde{S}_{m,2} = \sum_{k: p_k < n < d_k, m > p_k} \left(\sum_{j=0}^{p_k-2} (j+1) X_{k,j} + \sum_{j=m}^{m+p_k-2} (m+p_k-j-1) X_{k,j} \right).$$

The $\tilde{S}_{m,2}$, $1 \leq m \leq n$, have the same distributions as

$$\begin{aligned} \tilde{S}'_{m,2} &= \sum_{k: p_k < n < d_k, m > p_k} \left(\sum_{j=0}^{p_k-2} (p_k-j) X_{k,j} - \sum_{j=m}^{m+p_k-2} (m+p_k-j-1) X_{k,j} \right) \\ &= \sum_{k: p_k < n < d_k, m > p_k} \left(\sum_{j=0}^{p_k-2} (p_k-j) X_{k,j} - \tilde{U}^m \sum_{j=0}^{p_k-2} (p_k-j) X_{k,j} \right). \end{aligned}$$

Therefore,

$$\tilde{S}'_{m,2} = \tilde{G} - \tilde{U}^m \tilde{G}$$

with

$$\tilde{G} = \sum_{k: p_k < n < d_k, m > p_k} \tilde{g}_k, \quad \text{where } \tilde{g}_k = \sum_{j=0}^{p_k-2} (p_k - j) X_{k,j}.$$

From

$$\|\tilde{g}_k\|_2^2 \leq p_k^3 \alpha^2 = \frac{2^k}{k}$$

and Lemma 2 (with $a = 2$, $p = 1$, $c = 1$) it follows that there exists a $K < \infty$ such that

$$\|\tilde{G}\|_2^2 \leq \sum_{\log_3 n < k < \log_2 n} \|\tilde{g}_k\|_2^2 \leq K \frac{n}{\log n}.$$

Therefore

$$(14) \quad \|\tilde{S}_{m,2}\|_2^2 = O\left(\frac{n}{\log_2 n}\right)$$

and from Lemma 3(ii) it follows that

$$(15) \quad \lambda^2 \mu \left(\max_{0 \leq m \leq n-1} |\tilde{S}_{m,2}| > \lambda \sqrt{n} \right) \rightarrow 0 \quad \text{for } \lambda \rightarrow \infty, \quad \text{uniformly for all } n.$$

We have

$$(16) \quad \|\hat{S}_{m,2}\|_2 \leq 2 \sum_{k: p_k < n < d_k, m > p_k} p_k^2 \epsilon_k = O\left(\frac{1}{n^3}\right).$$

Hence by Lemma 3(i)

$$(17) \quad \lambda^2 \mu \left(\max_{0 \leq m \leq n-1} |\hat{S}_{m,2}| > \lambda \sqrt{n} \right) \rightarrow 0 \quad \text{for all } \lambda, \quad n \rightarrow \infty.$$

The same results are obtained if we replace \bar{f}_k by $\bar{f}_k - U^{d_k} \bar{f}_k$ and $X_{k,j}$ by $X_{k,j} - X_{k,d_k+j}$.

We have

$$\begin{aligned} S_{m,1} &= \sum_{k: p_k < n < d_k, m > p_k} \sum_{j=p_k-1}^{m-1} p_k U^j (\bar{f}_k - U^{d_k} \bar{f}_k), \\ \tilde{S}_{m,1} &= \sum_{k: p_k < n < d_k, m > p_k} \sum_{j=p_k-1}^{m-1} p_k (X_{k,j} - X_{k,d_k+j}), \\ \hat{S}_{m,1} &= \sum_{k: p_k < n < d_k, m > p_k} \sum_{j=p_k-1}^{m-1} p_k ((\bar{f}_k - X_{k,j}) - (U^{d_k} \bar{f}_k - X_{k,d_k+j})). \end{aligned}$$

We calculate

$$\begin{aligned}
 \left\| \sum_{j=0}^{m-1} Y_{n,j} - \tilde{S}_{m,1} \right\|_2^2 &= \left\| \sum_{j=0}^{m-1} \sum_{k: p_k < n < d_k} p_k X_{k,j} - \sum_{k: p_k < n < d_k, m > p_k} \sum_{j=p_k-1}^{m-1} p_k X_{k,j} \right\|_2^2 \\
 &= \left\| \sum_{k: p_k < n < d_k, p_k < m} \sum_{j=0}^{p_k-2} p_k X_{k,j} + \sum_{k: p_k < n < d_k, m \leq p_k} \sum_{j=0}^{m-1} p_k X_{k,j} \right\|_2^2 \\
 &= \sum_{k: p_k < n < d_k, p_k < m} p_k^2 (p_k - 1) \alpha_k^2 + \sum_{k: p_k < n < d_k, m \leq p_k} p_k^2 m \alpha_k^2 \\
 &= \sum_{k: p_k < n < d_k, p_k < m} \frac{p_k - 1}{k} + \sum_{k: p_k < n < d_k, m \leq p_k} \frac{m}{k}.
 \end{aligned}$$

By Lemma 2 (with $a = 2$, $p = 1$, $c = 1$) and $m \leq n$ we have

$$\sum_{k: p_k < n < d_k, p_k < m} \frac{p_k - 1}{k} = O\left(\frac{n}{\log n}\right).$$

Because $\log_2 m \leq \log_3 n$ if and only if $m \leq n^{\log_3 2}$, the maximum

$$\max_{1 \leq m \leq n} \frac{1}{n} \sum_{k: p_k < n < d_k, m \leq p_k} \frac{m}{k} = \max_{1 \leq m \leq n} \frac{1}{n} \sum_{\log_3 n < k < \log_2 n, \log_2 m \leq k} \frac{m}{k}$$

is achieved for $m \geq n^{\log_3 2}$. The sum

$$\sum_{\log_3 n < k < \log_2 n, \log_2 m \leq k} 1/k = \sum_{\log_2 m < k < \log_2 n} 1/k$$

can be estimated from above by $\int_{\log_2 m}^{\log_2 n} (1/x) dx$; hence

$$\frac{m}{n} \sum_{\log_2 m < k < \log_2 n} \frac{1}{k} \leq \frac{m}{n} (\log \log n - \log \log m).$$

Using the mean value theorem, we get (for $m < n$)

$$\begin{aligned}
 \frac{m}{n} \sum_{\log_2 m < k < \log_2 n} \frac{m}{k} &\leq \frac{m}{n} \frac{\log \log n - \log \log m}{\log n - \log m} \frac{\log n - \log m}{n - m} (n - m) \\
 &\leq \frac{m}{n} \frac{1}{\log m} \frac{1}{m} (n - m) = \frac{n - m}{\log m}.
 \end{aligned}$$

Because the maximum of $(1/n) \sum_{\log_3 n < k < \log_2 n, \log_2 m \leq k} m/k$ is achieved for $m \geq n^{\log_3 2}$, it is estimated from above by $O(n/\log n)$. Therefore

$$(18) \quad \max_{1 \leq m \leq n} \frac{1}{n} \left\| \sum_{j=0}^{m-1} Y_{n,j} - \tilde{S}_{m,1} \right\|_2^2 = O\left(\frac{1}{\log n}\right).$$

We have

$$\tilde{S}_{m,1} = \sum_{j=0}^{m-1} V_j,$$

where

$$V_j = \sum_{k: p_k < n < d_k, m > p_k, j \geq p_k - 1} p_k (X_{k,j} - X_{k,d_k+j}).$$

Because the V_j are mutually independent and

$$\|V_j\|_2^2 \leq 2 \sum_{k: p_k < n < d_k} p_k^2 \alpha_k^2,$$

which has been estimated (cf. (4)) by $2(\log \log 3 - \log \log 2)$.

By the invariance principle for independent random variables (see [B]),

$$(19) \quad \lambda^2 \mu \left(\max_{0 \leq m \leq n-1} |\tilde{S}_{m,1}| > \lambda \sqrt{n} \right) \rightarrow 0 \text{ for } \lambda \rightarrow \infty, \text{ and } n \rightarrow \infty.$$

We have

$$(20) \quad \|\hat{S}_{m,1}\|_2 \leq 2 \sum_{k: p_k < n < d_k, m > p_k} p_k d_k \epsilon_k = O \left(\frac{1}{n^2} \right).$$

Hence by Lemma 3(i)

$$(21) \quad \lambda^2 \mu \left(\max_{0 \leq m \leq n-1} |\hat{S}_{m,1}| > \lambda \sqrt{n} \right) \rightarrow 0 \text{ for } n \rightarrow \infty.$$

From (10), (12), (14), (16), (18), (20) we get (i), and from (11), (13), (15), (17), (19), (21) we get (ii). \square

From Lemma 4(i), Lemma 5(i), and Lemma 6(i) we obtain

Lemma 7.

$$\max_{1 \leq \ell \leq n} \frac{1}{\sqrt{n}} \|S_\ell(f) - \sum_{j=0}^{\ell-1} Y_{n,j}\|_2 = O \left(\frac{1}{\log n} \right)^{1/2}.$$

Theorem 1 follows from Lemma 7 and Lemma 1(ii). \square

Theorem 3 follows from Lemma 4(ii), Lemma 5(ii), and Lemma 6(ii). \square

Recall the definitions

$$Y_{n,i} = \sum_{k: p_k < n < d_k} p_k (X_{k,i} - X_{k,d_k+i}), \quad Z_{n,j} = Y_{n,j} \frac{\sqrt{2(\log \log 3 - \log \log 2)}}{\|Y_{n,j}\|_2}.$$

We define

$$Y_i = \sum_{k: p_k < i < d_k} p_k (X_{k,i} - X_{k,d_k+i}),$$

$$Z_i = Y_i \frac{\sqrt{2(\log \log 3 - \log \log 2)}}{\|Y_i\|_2}, \quad i = 0, 1, 2, \dots$$

Lemma 8.

$$\max_{1 \leq \ell \leq n} \frac{1}{\sqrt{n}} \left\| \sum_{j=0}^{\ell-1} (Z_j - Y_{n,j}) \right\|_2 = O \left(\frac{\log \log \log \log n}{\log n} \right)^{1/2}.$$

Proof of Lemma 8. First, let us prove

$$(22) \quad \max_{1 \leq \ell \leq n} \frac{1}{\sqrt{n}} \left\| \sum_{j=0}^{\ell-1} (Y_j - Y_{n,j}) \right\|_2 = O \left(\frac{\log \log \log \log n}{\log n} \right)^{1/2}.$$

Let $1 \leq \ell \leq n$. We have

$$\sum_{j=0}^{\ell-1} (Y_j - Y_{j,n}) = \sum_{j=0}^{\ell-1} \sum_{k: p_k < j < d_k \leq n} p_k X_{k,j} - \sum_{j=0}^{\ell-1} \sum_{k: j \leq p_k < n < d_k} p_k X_{k,j};$$

hence

$$(23) \quad \left\| \sum_{j=0}^{\ell-1} (Y_j - Y_{j,n}) \right\|_2^2 \leq \sum_{j=0}^{n-1} \sum_{k: p_k < j < d_k \leq n} \frac{1}{k} + \sum_{j=0}^{n-1} \sum_{k: j \leq p_k < n < d_k} \frac{1}{k}.$$

Because $p_k < j < d_k \leq n$ iff $\log_3 j < k < \log_2 j$, $k \leq \log_3 n$, and $\log_2 j \leq \log_3 n$ iff $j \leq n^{\log_3 2}$, we have

$$(24) \quad \sum_{j=0}^{n-1} \sum_{k: p_k < j < d_k \leq n} \frac{1}{k} = \sum_{j=0}^{[n^{\log_3 2}]} \sum_{\log_3 j < k < \log_2 j} \frac{1}{k} + \sum_{j=[n^{\log_3 2}]+1}^{n-1} \sum_{\log_3 j < k < \log_3 n} \frac{1}{k}$$

and

$$(25) \quad \sum_{j=0}^{n-1} \sum_{k: j \leq p_k < n < d_k} \frac{1}{k} \leq \sum_{j=0}^{[n^{\log_3 2}]} \sum_{\log_3 j < k < \log_2 j} \frac{1}{k} + \sum_{j=n^{\log_3 2}+1}^{n-1} \sum_{\log_2 j \leq k < \log_2 n} \frac{1}{k}.$$

By (4), for all j we have

$$\sum_{\log_3 j < k < \log_2 j} \frac{1}{k} \leq (\log \log 3 - \log \log 2) + \frac{3 \log 3}{\log j};$$

hence for $c \geq (\log \log 3 - \log \log 2) + (3 \log 3 / \log n)$

$$(26) \quad \sum_{j=0}^{[n^{\log_3 2}]} \sum_{\log_3 j < k < \log_2 j} \frac{1}{k} \leq c(1 + n^{\log_3 2}).$$

In the same way as we derived (4) we get

$$\begin{aligned} \sum_{\log_2 j \leq k < \log_2 n} \frac{1}{k} &\leq \log \left(\frac{\log n}{\log j} \right) + \frac{3}{\log_2 j}, \\ \sum_{\log_3 j \leq k < \log_3 n} \frac{1}{k} &\leq \log \left(\frac{\log n}{\log j} \right) + \frac{3}{\log_3 j}, \end{aligned}$$

and hence

$$(27) \quad \begin{aligned} \sum_{j=n^{\log_3 2}+1}^{n-1} \sum_{\log_2 j < k < \log_2 n} \frac{1}{k} &\leq \sum_{j=n^{\log_3 2}+1}^{n-1} \log \left(\frac{\log n}{\log j} \right) + \frac{3n}{\log_2 n}, \\ \sum_{j=n^{\log_3 2}+1}^{n-1} \sum_{\log_3 j < k < \log_3 n} \frac{1}{k} &\leq \sum_{j=n^{\log_3 2}+1}^{n-1} \log \left(\frac{\log n}{\log j} \right) + \frac{3n}{\log_3 n}. \end{aligned}$$

For $j \geq n^{\log_3 2}$ there exists $c > 0$ such that

$$\log \left(\frac{\log n}{\log j} \right) < c \left(\frac{\log n}{\log j} - 1 \right).$$

Let $K = \log \log n$. For $\log n \geq 2 \log \log n$ we have

$$\begin{aligned}
 & \sum_{j=[n^{\log_3 2}]}^{n-1} \log \left(\frac{\log n}{\log j} \right) \\
 &= \sum_{j=[n^{\log_3 2}]}^{[n/\log n]} \log \left(\frac{\log n}{\log j} \right) + \sum_{j=[n/\log n]+1}^{n-1} \log \left(\frac{\log n}{\log j} \right) \\
 (28) \quad &\leq \frac{n}{\log n} (\log \log 3 - \log \log 2) + \sum_{j=[n/\log n]+1}^{n-1} c \left(\frac{\log n}{\log j} - 1 \right) \\
 &\leq \frac{n}{\log n} + \sum_{j=[n/\log n]+1}^{[n/K]} c \left(\frac{\log n}{\log j} - 1 \right) + \sum_{j=[n/K]+1}^{n-1} c \left(\frac{\log n}{\log j} - 1 \right) \\
 &\leq \frac{n}{\log n} + \frac{cn}{K} \left(\frac{\log n}{\log n - \log \log n} - 1 \right) + cn \left(\frac{\log n}{\log n - \log \log K} - 1 \right) \\
 &\leq \frac{n}{\log n} + \frac{2cn}{K} \frac{\log \log n}{\log n} + 2cn \frac{\log \log K}{\log n}.
 \end{aligned}$$

From (23)–(28) it follows that there exists a constant C such that for all n sufficiently big

$$(22) \quad \max_{1 \leq \ell \leq n} \left\| \sum_{j=0}^{\ell-1} (Y_j - Y_{j,n}) \right\|_2^2 \leq C \frac{n \log \log \log \log n}{\log n}.$$

It remains to show that

$$\max_{1 \leq \ell \leq n} \left\| \sum_{j=0}^{\ell-1} (Y_j - Z_j) \right\|_2^2 = O \left(\frac{n \log \log \log \log n}{\log n} \right).$$

From the mutual orthogonality of the Y_i we get

$$\begin{aligned}
 & \left\| \sum_{j=0}^{\ell-1} (Y_j - Z_j) \right\|_2 = \left\| \sum_{j=0}^{\ell-1} Y_j \left(1 - \frac{\sqrt{2(\log \log 3 - \log \log 2)}}{\|Y_j\|_2} \right) \right\|_2 \\
 &= \sum_{j=0}^{\ell-1} (\|Y_j\|_2 - \sqrt{2(\log \log 3 - \log \log 2)})^2 \\
 &\leq 2 \sum_{j=0}^{\ell-1} (Y_{n,j} - \sqrt{2(\log \log 3 - \log \log 2)})^2 + 2 \sum_{j=0}^{\ell-1} \|Y_j - Y_{n,j}\|_2^2.
 \end{aligned}$$

By (22) and Lemma 1(i) the last two summands on the right are $O(n/\log^2 n)$ and $O(n \log \log \log n / \log n)$, respectively. \square

Theorem 2 follows from Lemma 7 and Lemma 8. \square

Proof of Theorem 4. For every positive integer n

$$S_n(f) = \sum_{j=0}^{n-1} Z_j + \sum_{j=0}^{n-1} (Y_{n,j} - Z_j) + \left(S_n(f) - \sum_{j=0}^{n-1} Y_{n,j} \right).$$

By Lemma 4(i), Lemma 5(i), and Lemma 6(i) the difference $S_n(f) - \sum_{j=0}^{n-1} Y_{n,j}$ can be expressed as

$$\begin{aligned} S_n(f) - \sum_{j=0}^{n-1} Y_{n,j} &= \tilde{g}'_{(n)} - \tilde{g}'_{(n)} \circ T^n + \tilde{C}_{*,n} + \tilde{D}_{*,n} + \tilde{E}_{*,n} + S''_{n+} + \tilde{S}_{n,2} \\ &+ \left(\tilde{S}_{n,1} - \sum_{j=0}^{n-1} Y_{n,j} \right) + \hat{g}'_{(n)} - \hat{g}'_{(n)} \circ T^n + \hat{C}_{*,n} + \hat{D}_{*,n} + \hat{E}_{*,n} + \hat{S}''_{n+} + \hat{S}_{n,1} + \hat{S}_{n,1}. \end{aligned}$$

The random variables $\sum_{j=0}^{n-1} (Y_{n,j} - Z_j)$, $\tilde{g}'_{(n)}$, $\tilde{g}'_{(n)} \circ T^n$, $\tilde{C}_{*,n}$, $\tilde{D}_{*,n}$, $\tilde{E}_{*,n}$, S''_{n+} , $\tilde{S}_{n,1} - \sum_{j=0}^{n-1} Y_{n,j}$, and $\tilde{S}_{n,2}$ are Gaussian with zero means (as the $X_{n,j}$ have zero means), and (by Lemmas 4(i), 5(i), 6(i), and 8) their variances are bounded by $n \log \log \log n / \log n$.

For any $\epsilon > 0$ and Gaussian random variables A_n with variances bounded by $K \log \log \log n / (\log n \log \log n)$ for some K , we have

$$\mu(|A_n| > \epsilon) = \sqrt{\frac{2}{\pi}} \int_{\epsilon \sqrt{\log n \log \log n / (K \log \log \log n)}}^{\infty} e^{-x^2/2} dx < n^{-2}$$

for n sufficiently big. By the Borel-Cantelli lemma we thus get

$$(29) \quad \lim_{n \rightarrow \infty} \frac{A_n}{\sqrt{n \log \log n}} = 0 \quad \text{a.s.}$$

for $A_n = (\sum_{j=0}^{n-1} (Y_{n,j} - Z_j) + \tilde{g}'_{(n)} + \tilde{g}'_{(n)} \circ T^n + \tilde{C}_{*,n} + \tilde{D}_{*,n} + \tilde{E}_{*,n} + S''_{n+} + \tilde{S}_{n,1} - \sum_{j=0}^{n-1} Y_{n,j} + \tilde{S}_{n,2})$. Let us define

$$\hat{g} = \sum_{k=1}^{\infty} |\hat{g}'_k|.$$

Then

$$|\hat{g}'_{(n)}| \leq \hat{g}$$

for every n ; and because $\|\hat{g}'_k\|_2 \leq d_k p_k \epsilon_k = 6^{-2k}$ (cf. the proof of Lemma 4), $\hat{g} \in L^2$. From the integrability of \hat{g}^2 it follows that

$$\sum_{n=1}^{\infty} \mu(|\hat{g}'_{(n)} \circ T^n| > \epsilon \sqrt{n}) \leq \sum_{n=1}^{\infty} \mu(\hat{g}^2 > \epsilon^2 n) < \infty.$$

Hence by the Borel-Cantelli lemma

$$(30) \quad \lim_{n \rightarrow \infty} \frac{\hat{g}'_{(n)} \circ T^n}{\sqrt{n \log \log n}} = 0 \quad \text{a.s.}$$

and similarly

$$(31) \quad \lim_{n \rightarrow \infty} \frac{\hat{g}'_{(n)}}{\sqrt{n \log \log n}} = 0 \quad \text{a.s.}$$

By Lemma 5(i) and Lemma 6(i) there exists a K such that if A_n is any of the random variables $\hat{C}_{*,n}$, $\hat{D}_{*,n}$, $\hat{E}_{*,n}$, \hat{S}''_{n+} , $\hat{S}_{n,1}$, and $\hat{S}_{n,1}$, then $EA_n^2 \leq K/n^2$. Hence by the Chebyshev inequality

$$\mu(|A_n| > \epsilon \sqrt{n}) \leq \frac{K}{\epsilon^2 n^3},$$

and by the Borel-Cantelli lemma

$$(31) \quad \lim_{n \rightarrow \infty} \frac{1}{\sqrt{n \log \log n}} (\hat{C}_{*,n} + \hat{D}_{*,n} + \hat{E}_{*,n} + \hat{S}_{n+}'' + \hat{S}_{n,1} + \hat{S}_{n,1}) \text{ a.s.}$$

From (29)–(31) we get

$$\lim_{n \rightarrow \infty} \frac{S_n(f) - \sum_{j=0}^{n-1} Z_j}{\sqrt{n \log \log n}} = 0 \text{ a.s.}$$

□

3. CONCLUDING REMARKS

Using the previous constructions, E. Lesigne recently proved ([Le1]) that in an aperiodic and ergodic dynamical system there exists a function $f \in L^2$ such that for the process $(f \circ T^i)$ the almost sure central limit theorem holds. This means that for almost every ω the measures

$$\frac{1}{\log n} \sum_{k=1}^n \frac{1}{k} \delta_{S_n(\omega)/\sqrt{n}}$$

weakly converge to the standard normal distribution; here, δ_x denotes the Dirac measure at a point $x \in \mathbb{R}$ and $S_n(\omega)$ denotes $\sum_{i=0}^{n-1} f(T^i \omega)$.

According to [Le2] it seems that using the ideas of the previous proofs one can get the invariance principle (convergence of finite dimensional distributions) in the sense of the convergence to α -stable processes, $1 < \alpha < 2$. It would be very interesting to understand what happens for the Cauchy law ($\alpha = 1$).

Recently J.-P. Thouvenot and B. Weiss ([T-We]) showed that in every aperiodic and ergodic dynamical system and every probability law (i.e. probability measure on the real line) ν there exist a measurable function f and a sequence $c_n \rightarrow \infty$ such that the distributions of the normalized sums $S_n(f)/c_n$ converge weakly to ν . More generally, for any connected and closed set G of probability laws f and (c_n) can be found such that $\nu \in G$ if and only if there exists a subsequence $n_k \rightarrow \infty$ such that the distributions of the normalized sums $S_{n_k}(f)/c_{n_k}$ converge weakly to ν .

The set of functions f for which the central limit theorem holds is small: In the space of zero mean L^p functions, $1 \leq p \leq \infty$, for every sequence $a_n \rightarrow \infty$, $a_n/n \rightarrow 0$, there is a dense G_δ set of functions f such that for every probability law ν there exist $n_k \rightarrow \infty$ such the distributions of $(1/a_{n_k})S_{n_k}(f)$ converge weakly to ν (see [Vo1]; the case of $p = \infty$ is considered in [K-Vo]). A similar result holds in the space of continuous bounded functions on a metrizable compact where T is a homeomorphism (see [Li-Vo]).

In chaotic dynamical systems, sufficient smoothness can guarantee the central limit theorem (cf. e.g. [Ka], [Ch]), but for irrational rotations we get results similar to the L^p spaces (see [Li-Vo]).

ACKNOWLEDGEMENT

I would like to express my thanks to Professor Emmanuel Lesigne for pointing out an error in the proof of Proposition 2, many consultations, and much support. I also thank Professor Jean-Paul Thouvenot for consultations, and kind encouragement.

REFERENCES

- [B] Billingsley, P., *Convergence of Probability Measures*, J. Wiley, New York, 1968. MR **38**:1718
- [Bu-De] Burton, R. and Denker, M., *On the central limit theorem for dynamical systems*, Transactions Amer. Math. Soc. **302** (1987), 715-726. MR **88i**:60039
- [C-F-S] Cornfeld, I.P., Fomin, S.V., and Sinai Ya.G., *Ergodic Theory*, Springer-Verlag, New York, 1982. MR **87f**:28019
- [Ch] Chernov, N., *Limit theorems and Markov approximations for chaotic dynamical systems*, Probab. Theory Related Fields **101** (1995), 321-362. MR **96m**:28016
- [Cs-Re] Csörgö, M. and Révész, P., *Strong Approximations in Probability and Statistics*, Academic Press, New York, 1981. MR **84d**:60050
- [D-L-P-W] De La Rue, T., Ladouceur, S., Peškir, G., and Weber, M., *On the central limit theorem for aperiodic dynamical systems and applications*, preprint 1993.
- [De-K] Denker, M. and Keane, M., *Almost topological dynamical systems*, Israel J. Math. **34** (1979), 139-160. MR **82m**:54042
- [G] Gordin M.I., *The central limit theorem for stationary processes*, Soviet Math. Dokl. **10** (1969), 1174-1176. MR **40**:5012
- [Ha-He] Hall, P. and Heyde, C.C., *Martingale Limit Theory and its Application*, Academic Press, New York, 1980. MR **83a**:6001
- [Ka] Kac, M., *On the distribution of values of sums of the type $\sum f(2^k t)$* , Ann. Math. **47** (1946), 33-49. MR **7**:436f
- [K-Vo] Keane, M. and Volný, D., an unpublished result.
- [La1] Lacey, M., *On weak convergence in dynamical systems to self-similar processes with spectral representation*, Transactions Amer. Math. Soc. **328** (1991), 767-778. MR **92c**:60048
- [La2] Lacey, M., *On central limit theorems, modulus of continuity and Diophantine type for irrational rotations*, Journal d'Analyse Math. **61** (1993), 47-59. MR **95a**:60054
- [Le1] Lesigne, E., *Almost sure central limit theorem for strictly stationary processes*, preprint 1996.
- [Le2] Lesigne, E., personal communication, 1996.
- [Li-Vo] Liardet, P. and Volný, D., *Sums of continuous and differentiable functions in dynamical systems*, Israel J. of Mathematics **98** (1997), 29-60. CMP 97:15
- [Ru] Rudolph, D., *Fundamentals of Measurable Dynamics*, Clarendon Press, Oxford, 1990. MR **92e**:28006
- [T-We] Thouvenot, J.-P. and Weiss, B., personal communication.
- [Vo1] Volný, D., *On limit theorems and category for dynamical systems*, Yokohama Math. J. **38** (1990), 29-35. MR **92c**:28014
- [Vo2] Volný, D., *Approximating martingales and the central limit theorem for strictly stationary processes*, Stochastic Processes and their Appl. **44** (1993), 41-74. MR **93m**:28021

UNIVERSITÉ DE ROUEN, UPRES-A CNRS 60 85, SITE COLBERT, 76821 MONT-SAINT-AIGNAN
CEDEX, FRANCE

E-mail address: dalibor.volny@univ-rouen.fr