Elect. Comm. in Probab. **12** (2007), 184–199

ELECTRONIC COMMUNICATIONS in PROBABILITY

ASYMPTOTIC RESULTS FOR EMPIRICAL MEASURES OF WEIGHTED SUMS OF INDEPENDENT RANDOM VARIABLES

BERNARD BERCU
Institut de Mathématiques de Bordeaux, Université Bordeaux 1, 351 cours de la libération 33405 Talence cedex, France
email: Bernard.Bercu@math.u-bordeaux1.fr

WŁODZIMIERZ BRYC[1]
Department of Mathematical Sciences, University of Cincinnati, 2855 Campus Way, PO Box 210025, Cincinnati, OH 45221-0025, USA
email: Wlodzimierz.Bryc@UC.edu

Submitted January 24, 2007, *accepted in final form* May 8 2007

AMS 2000 Subject classification: Primary: 60F15; Secondary: 60F05, 60F10
Keywords: almost sure central limit theorem, large deviations,normal approximation, periodogram

Abstract
We investigate the asymptotic behavior of weighted sums of independent standardized random variables with uniformly bounded third moments. The sequence of weights is given by a family of rectangular matrices with uniformly small entries and approximately orthogonal rows. We prove that the empirical CDF of the resulting partial sums converges to the normal CDF with probability one. This result implies almost sure convergence of empirical periodograms, almost sure convergence of spectral distribution of circulant and reverse circulant matrices, and almost sure convergence of the CDF generated from independent random variables by independent random orthogonal matrices. In the special case of trigonometric weights, the speed of the almost sure convergence is described by a normal approximation as well as a large deviation principle.

1 Introduction

Let (X_n) be a sequence of independent and identically distributed random variables with $\mathbb{E}[X_n] = 0$ and $\mathbb{E}[X_n^2] = 1$ and denote

$$S_n = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} X_t.$$

[1]RESEARCH PARTIALLY SUPPORTED BY NSF GRANT #DMS-0504198.

The almost sure central limit theorem (ASCLT) states that, for any $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{t=1}^{n} \frac{1}{t} 1_{\{S_t \le x\}} = \Phi(x)$$

with probability one, where Φ stands for the standard normal distribution [3], [7], [14], [21]. One can observe that the ASCLT does not hold for the Césaro averaging. Assume now that (X_n) has all moments finite and consider the weighted sum

$$S_{n,k} = \sqrt{\frac{2}{n}} \sum_{t=1}^{n} X_t \cos\left(\frac{\pi k t}{n}\right). \tag{1.1}$$

It was recently established by Massey, Miller and Sinsheimer [17] that, for any $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} 1_{\{S_{n,k} \le x\}} = \Phi(x) \tag{1.2}$$

with probability one together with the uniform convergence. This result is closely related to the limiting spectral distribution of random symmetric circulant matrices [6] since the eigenvalues of theses matrices are exactly given by (1.1). The goal of this paper is to answer to several natural questions.

(a) Is it possible to remove the assumption of identical distribution?

(b) Can the moment condition be relaxed?

(c) Can the trigonometric coefficients be replaced by other numbers?

(d) Can the multivariate version of convergence (1.2) be established?

(e) Is it possible to prove a CLT or a LDP associated with (1.2)?

We shall propose positive answers to these questions extending [17, Theorem 5.1] in various directions. First of all, we shall show that it is possible to deal with a sequence (X_n) of independent random variables defined on a common probability space $(\Omega, \mathcal{A}, \mathbb{P})$ without assuming that they share the same distribution. In addition, we shall only require that the third moments of (X_n) are uniformly bounded. Next, we shall allow more general weights and we will prove the multivariate version of convergence (1.2). Finally, we shall provide a CLT as well as a LDP.

2 Main results

Let $(\mathbf{U}^{(n)})$ be a family of real rectangular $r_n \times n$ matrices where (r_n) is an increasing sequence of integers which goes to infinity with $1 \le r_n \le n$. We shall assume that there exist some constants $C, \delta > 0$ which do not depend on n such that

$$\text{(A}_1\text{)} \qquad \max_{1\le k\le r_n,\, 1\le t\le n} |u_{k,t}^{(n)}| \le \frac{C}{(\log(1+r_n))^{1+\delta}},$$

$$\text{(A}_2\text{)} \qquad \max_{1\le k,\, l\le r_n} \left| \sum_{t=1}^{n} u_{k,t}^{(n)} u_{l,t}^{(n)} - \delta_{k,l} \right| \le \frac{C}{(\log(1+r_n))^{1+\delta}}.$$

For application to periodograms, we will also need to consider pairs $(\mathbf{U}^{(n)}, \mathbf{V}^{(n)})$ of such matrices and we shall assume that

$$(\mathrm{A}_3) \qquad \max_{1 \le k,\, l \le r_n} \left| \sum_{t=1}^{n} u_{k,t}^{(n)} v_{l,t}^{(n)} \right| \le \frac{C}{(\log(1+r_n))^{1+\delta}}.$$

These assumptions are not really restrictive and they are fulfilled in many situations. For example, consider the sequence of real rectangular $r_n \times n$ matrices $(\mathbf{U}^{(n)}, \mathbf{V}^{(n)})$ with $r_n \le \lfloor \frac{n-1}{2} \rfloor$ given, for all $1 \le k \le r_n$ and $1 \le t \le n$, by

$$u_{k,t}^{(n)} = \sqrt{\frac{2}{n}} \cos\left(\frac{2\pi kt}{n}\right), \qquad v_{k,t}^{(n)} = \sqrt{\frac{2}{n}} \sin\left(\frac{2\pi kt}{n}\right). \tag{2.1}$$

Then (A_1) holds trivially while (A_2) and (A_3) follow from $2r_n < n$ and the following well known trigonometric identities where $1 \le k < l \le n$.

$$\sum_{t=1}^{n} \cos\left(\frac{2\pi kt}{n}\right) \cos\left(\frac{2\pi lt}{n}\right) = \begin{cases} 0 & \text{if} \quad k+l \neq n, \\ n/2 & \text{if} \quad k+l = n, \end{cases} \tag{2.2}$$

$$\sum_{t=1}^{n} \sin\left(\frac{2\pi kt}{n}\right) \sin\left(\frac{2\pi lt}{n}\right) = \begin{cases} 0 & \text{if} \quad k+l \neq n, \\ -n/2 & \text{if} \quad k+l = n. \end{cases} \tag{2.3}$$

In addition,

$$\sum_{t=1}^{n} \cos\left(\frac{2\pi kt}{n}\right) \sin\left(\frac{2\pi lt}{n}\right) = 0, \tag{2.4}$$

$$\sum_{t=1}^{n} \cos^2\left(\frac{2\pi kt}{n}\right) = n - \sum_{t=1}^{n} \sin^2\left(\frac{2\pi kt}{n}\right) = \begin{cases} n/2 & \text{if} \quad 2k \neq n, \\ n & \text{if} \quad 2k = n. \end{cases} \tag{2.5}$$

In order to avoid cumbersome notation, we only state our first result in the univariate and bivariate cases. The d-variate extension requires introducing d sequences of matrices that satisfy assumptions (A_1) and (A_2) with each pair from the n-th level satisfying condition (A_3) and the proof follows essentially the same lines.

Theorem 1 (ASCLT)**.** *Assume that (X_n) is a sequence of independent random variables such that $\mathbb{E}[X_n] = 0$, $\mathbb{E}[X_n^2] = 1$ and $\sup \mathbb{E}[|X_n|^3] < \infty$. Let $(S_{n,k})$ be the sequence of weighted sums*

$$S_{n,k} = \sum_{t=1}^{n} u_{k,t}^{(n)} X_t \tag{2.6}$$

with $k = 1, 2, \ldots, r_n$ where $(\mathbf{U}^{(n)})$ is a family of real rectangular $r_n \times n$ matrices satisfying (A_1) and (A_2). Then, we have the almost sure uniform convergence

$$\lim_{n \to \infty} \sup_{x \in \mathbb{R}} \left| \frac{1}{r_n} \sum_{k=1}^{r_n} 1_{\{S_{n,k} \le x\}} - \Phi(x) \right| = 0. \tag{2.7}$$

In addition, let $(T_{n,k})$ be the sequence of weighted sums

$$T_{n,k} = \sum_{t=1}^{n} v_{k,t}^{(n)} X_t \tag{2.8}$$

with $k = 1, 2, \ldots, r_n$ where $(\mathbf{V}^{(n)})$ is a family of real rectangular $r_n \times n$ matrices such that (A_1) and (A_2) hold with $v_{k,t}^{(n)}$ in place of $u_{k,t}^{(n)}$. Assume that the sequence of pairs $(\mathbf{U}^{(n)}, \mathbf{V}^{(n)})$ satisfies (A_3). Then there is a measurable set $\Delta \subset \Omega$ of probability one such that for all $x, y \in \mathbb{R}$, we have the almost sure convergence on Δ

$$\lim_{n\to\infty} \frac{1}{r_n} \sum_{k=1}^{r_n} 1_{\{S_{n,k}\le x, T_{n,k}\le y\}} = \Phi(x)\Phi(y). \tag{2.9}$$

For the univariate case with trigonometric coefficients given by (2.1), we also have the companion weak limit theorem and the large deviation principle under the restrictions on the rate of growth of the sequence (r_n). The most attractive case

$$r_n = [(n-1)/2]$$

which corresponds to the spectral measures of random circulant matrices, is unfortunately not covered by our result. Some LDP for spectra of other random matrices can be found in [11, Chapter 5].

Theorem 2 (CLT)**.** *Assume that (X_n) is a sequence of independent random variables such that $\mathbb{E}[X_n] = 0$, $\mathbb{E}[X_n^2] = 1$ and satisfying for some constant $\tau > 0$,*

$$\sup_{n\ge 1} \mathbb{E}[|X_n|^3 \exp(|X_n|/\tau)] \le \tau. \tag{2.10}$$

Consider the sequence of weighted sums $(S_{n,k})$ given by (2.6), where $(\mathbf{U}^{(n)})$ corresponds to the trigonometric weights given by (2.1). If (r_n) is such that

$$\frac{(\log n)^2}{n} r_n^3 \to 0, \tag{2.11}$$

then for all $x \in \mathbb{R}$,

$$\frac{1}{\sqrt{r_n}} \sum_{k=1}^{r_n} \left(1_{\{S_{n,k}\le x\}} - \Phi(x)\right) \xrightarrow[n\to+\infty]{\mathcal{D}} \mathcal{N}\left(0, \Phi(x)(1-\Phi(x))\right). \tag{2.12}$$

In addition, we also have

$$\sqrt{r_n} \sup_{x\in\mathbb{R}} \left| \frac{1}{r_n} \sum_{k=1}^{r_n} 1_{\{S_{n,k}\le x\}} - \Phi(x) \right| \xrightarrow[n\to+\infty]{\mathcal{D}} \mathcal{L} \tag{2.13}$$

where $\mathcal{L}$ stands for the Kolmogorov-Smirnov distribution.

Remark 1. *The conclusions (2.12), (2.13) also hold if conditions (2.10) and (2.11) are replaced by the assumption that there is $p > 0$ such that*

$$\sup_{n\ge 1} \mathbb{E}[|X_n|^{2+p}] < \infty \qquad \text{as soon as} \qquad r_n^3 n^{-p/(2+p)} \to 0.$$

This follows from our proof, using [20, Section 5, Corollary 5] instead of Lemma 3. For example, if $p = 1$, it is necessary to assume that $r_n^9 = o(n)$.

Remark 2. *The Kolmogorov-Smirnov distribution is the distribution of the supremum of the absolute value of the Brownian bridge.*

The large deviation principle was motivated by the LDP from the Brosamler-Schatte almost sure CLT, see [16, Theorem 1]. To formulate the result, we need to introduce additional notation. Let $\mathcal{M}_1(\mathbb{R})$ denote the Polish space of probability measures on the Borel sets of $\mathbb{R}$ with the topology of weak convergence. For the sequence of weighted sums $(S_{n,k})$ given by (2.6), consider the empirical measures

$$\mu_n = \frac{1}{r_n}\sum_{k=1}^{r_n}\delta_{S_{n,k}} \tag{2.14}$$

The rate function $I : \mathcal{M}_1(\mathbb{R}) \to [0,\infty]$ in our LDP is the relative entropy with respect to the standard normal law. More precisely, if $\phi(x)$ denotes the standard normal density and $\nu \in \mathcal{M}_1(\mathbb{R})$, we have

$$I(\nu) = \int_{\mathbb{R}} \log\frac{f(x)}{\phi(x)} f(x)\,dx$$

if ν is absolutely continuous with respect to the Lebesgue measure of $\mathbb{R}$ with density f and the integral exists and $I(\nu) = +\infty$ otherwise. It is well known that the level sets $I^{-1}[0,a]$ are compact for $a < \infty$. The conclusion of our next result is the LDP for the empirical measures μ_n with speed r_n and good rate function I.

Theorem 3 (LDP)**.** *Assume that (X_n) shares the same assumptions as in Theorem 2. Consider the sequence of weighted sums $(S_{n,k})$ given by* (2.6)*, where $(\mathbf{U}^{(n)})$ corresponds to the trigonometric weights given by* (2.1)*. If (r_n) is such that*

$$\frac{r_n^4}{n} \to 0 \qquad \text{and} \qquad \frac{\log n}{r_n} \to 0,$$

then for all closed sets F and open sets G in $\mathcal{M}_1(\mathbb{R})$,

$$\limsup_{n\to\infty} \frac{1}{r_n}\log\mathbb{P}(\mu_n \in F) \leq -\inf_{\nu\in F} I(\nu)$$

and

$$\liminf_{n\to\infty} \frac{1}{r_n}\log\mathbb{P}(\mu_n \in G) \geq -\inf_{\nu\in G} I(\nu).$$

All technical proofs are postponed to section 4. We shall now provide several applications of our results.

3 Applications

3.1 Application to periodograms

The empirical periodogram associated with the sequence (X_n) is defined, for all λ in the torus $\mathbb{T} = [-\pi,\pi[$, by

$$I_n(\lambda) = \frac{1}{n}\left|\sum_{t=1}^{n} e^{-it\lambda}X_t\right|^2.$$

The empirical distribution of the periodogram is the random CDF given, for all $x \geq 0$, by

$$F_n(x) = \frac{1}{r_n} \sum_{k=1}^{r_n} 1_{\{I_n(2\pi k/n) \leq x\}}$$

where $r_n = \lfloor \frac{n-1}{2} \rfloor$. Theorem 1 strengthens the conclusion of [13, Proposition 4.1] to almost sure convergence at the expense of the assumption that third moments are finite.

Corollary 4. *Assume that* (X_n) *is a sequence of independent random variables such that* $\mathbb{E}[X_n] = 0$, $\mathbb{E}[X_n^2] = 1$ *and* $\sup \mathbb{E}[|X_n|^3] < \infty$. *Then, we have the almost sure uniform convergence*

$$\lim_{n \to \infty} \sup_{x \geq 0} |F_n(x) - (1 - \exp(-x))| = 0.$$

Proof. Following [13, (2.1)], we can rewrite F_n as the CDF of the empirical measure

$$\mu_n = \frac{1}{r_n} \sum_{k=1}^{r_n} \delta_{\frac{1}{2}(S_{n,k}^2 + T_{n,k}^2)}$$

where $(S_{n,k}, T_{n,k})$ are defined by (2.6) and (2.8) and $(\mathbf{U}^{(n)}, \mathbf{V}^{(n)})$ are the trigonometric weights given by (2.1). Consequently, Corollary 4 immediately follows from Theorem 1. As a matter of fact, let $h : E \to F$ be a continuous mapping of Polish spaces. If a sequence of discrete measure (ν_n) converges weakly to some probability measure ν on the Borel sigma-field of E, then $(\nu_n \circ h^{-1})$ converge weakly to the probability measure $\nu \circ h^{-1}$, see e.g. [4, Theorem 29.2]. We apply it to

$$\nu_n = \frac{1}{r_n} \sum_{k=1}^{r_n} \delta_{(S_{n,k}, T_{n,k})}$$

and to the continuous mapping $h : \mathbb{R}^2 \to \mathbb{R}$ given by

$$h(x, y) = \frac{1}{2}(x^2 + y^2).$$

On the one hand, we clearly have $\mu_n = \nu_n \circ h^{-1}$. On the other hand, as ν is the product of two independent standard normal distributions, the limiting distribution $\mu = \nu \circ h^{-1}$ is simply the standard exponential distribution. Hence, for all $x \geq 0$, $F(x) = 1 - \exp(-x)$. Finally, as this limit is a continuous CDF, it is well known, see [4, Exercise 14.8], that the convergence is uniform. □

3.2 Application to symmetric circulant and reverse circulant matrices

Corollary 5. *The weak convergence in [5, Theorem 5] and in [6] holds with probability one.*

Proof. One can find in [5, Theorem 5] and [6] the analysis of the limiting spectral distribution of the $n \times n$ symmetric random matrices with typical eigenvalues of the form

$$\pm \sqrt{(S_{n,k}^2 + T_{n,k}^2)/2}$$

where $(\mathbf{U}^{(n)}, \mathbf{V}^{(n)})$ are the trigonometric weights given by (2.1) and $r_n = \lfloor \frac{n-1}{2} \rfloor$ see [6, Lemma 1]. After omitting at most two eigenvalues which do not modify the convergence of the spectral measure, Theorem 1 implies that the convergence holds with probability one by the same arguments as in the proof of Corollary 4. □

Assume now that $(\mathbf{A}_n)$ is a family of symmetric random circulant matrix formed from the sequence of independent random variables (X_n) by taking as the first row

$$[\mathbf{A}_n]_{1,t} = X_t$$

with $t = 1, 2, \ldots, [(n+1)/2]$ and $[\mathbf{A}_n]_{1,t} = [\mathbf{A}_n]_{1,n-t}$ for the other indices. The next corollary strengthens [6, Remark 2] to almost sure convergence and removes the assumption of integrability and identical distribution in [17, Theorem 1.5]. To justify the later claim, we note that a "palindromic matrix" analyzed in [17] differs from $\mathbf{A}_n$ by the last row and column only. Thus their ranks differ by at most one and asymptotically "palindromic matrices" and random circulant matrices have the same spectrum, see [1, Lemma 2.2].

Corollary 6. *Assume that (X_n) is a sequence of independent random variables with common mean $m = \mathbb{E}[X_n] = 0$, common variance $\sigma^2 = Var(X_n) > 0$ and uniformly bounded third moments $\sup \mathbb{E}[|X_n|^3] < \infty$. Then, the spectral distribution of the random matrix*

$$\frac{1}{\sigma\sqrt{n}}\mathbf{A}_n$$

converges weakly with probability one to the standard normal distribution.

Proof. Subtracting the rank 1 matrix does not change the asymptotic of the spectral distribution. Consequently, without loss of generality, we may assume that $m = 0$. Rescaling the random variables by $\sigma > 0$ we can also assume that $\sigma^2 = 1$. With the exception of at most two eigenvalues, the remaining eigenvalues of $\mathbf{A}_n/\sqrt{n}$ are of multiplicity two and are given by (2.6) with the trigonometric weights given by (2.1), see [6, Remark 2]. Finally, the weak convergence with probability one of the spectral distribution of $\mathbf{A}_n/\sqrt{n}$ to $\mathcal{N}(0,1)$ follows from Theorem 1. ☐

3.3 Application to random orthogonal matrices

A well known result of Poincaré says that if $\mathbf{U}^{(n)}$ is a random orthogonal matrix uniformly distributed on $O(n)$ and $x_n \in \mathbb{R}^n$ is a sequence of vectors of norm $\sqrt{n}$ then the first k coordinates of $\mathbf{U}^{(n)}x_n$ are asymptotically normal and independent, see e.g. [4, Exercise 29.9].

Corollary 7. *Assume that (X_n) is a sequence of independent random variables such that $\mathbb{E}[X_n] = 0$, $\mathbb{E}[X_n^2] = 1$ and $\sup \mathbb{E}[|X_n|^3] < \infty$. Consider the sequence of weighted sums $(S_{n,k})$ given by (2.6) where $(\mathbf{U}^{(n)})$ is a family of random orthogonal matrices uniformly distributed on $O(n)$ and independent of (X_n). Then, we have the almost sure uniform convergence*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \left| \frac{1}{n} \sum_{k=1}^{n} 1_{\{S_{n,k} \le x\}} - \Phi(x) \right| = 0. \tag{3.1}$$

Remark 3. *This result has a direct elementary proof, which we learned from Jack Silverstein. His proof shows that the result holds also for i.i.d. random variables with finite second moments. Here we derive it as a corollary to Theorem 1.*

Proof. Orthogonal matrices satisfy (A_2) with $r_n = n$. By [12, Theorem 1], (A_1) holds with probability 1. Therefore, redefining $\mathbf{U}^{(n)}$ and (X_t) on the product probability space $\Omega_U \times \Omega_X$, by [12, Theorem 1], there is a subset Δ_U of probability 1 such that for each $\omega_1 \in \Delta_U$, by Theorem 1, one can find a measurable subset $\Omega_{X,\omega_1} \subset \Omega_X$ of probability one such that (3.1) holds. By Fubini's Theorem, the set of all pairs (ω_1, ω_2) for which (3.1) holds has probability one. ☐

4 Proofs

4.1 Proof of Theorem 1

In order to use the well known method of the characteristic function, let $s, t \in \mathbb{R}$ and consider the random variable

$$\Phi_n(s,t) = \frac{1}{r_n} \sum_{k=1}^{r_n} \exp(isS_{n,k} + itT_{n,k}).$$

Lemma 1. *For all $s, t \in \mathbb{R}$, one can find some constant $C(s,t) > 0$ which does not depend on n such that for n large enough*

$$\mathbb{E}\left[|\Phi_n(s,t) - \Phi(s,t)|^2\right] \leq \frac{C(s,t)}{(\log(1+r_n))^{1+\delta}}. \tag{4.1}$$

Proof. For all $s, t \in \mathbb{R}$, denote by $L_n(s,t)$ the left hand side of (4.1). We have the decomposition

$$L_n(s,t) = \frac{1}{r_n^2} \sum_{k=1}^{r_n} \sum_{l=1}^{r_n} l_n(s,t,k,l)$$

where

$$l_n(s,t,k,l) = \mathbb{E}\left[(\exp(isS_{n,k} + itT_{n,k}) - \Phi(s,t))\,(\exp(-isS_{n,l} - itT_{n,l}) - \Phi(s,t))\right].$$

In addition, if

$$\varphi_n(t) = \mathbb{E}[\exp(itX_n)],$$

we clearly have

$$l_n(s,t,k,l) = a_n(s,t,k,l) - b_n(s,t,k,l) \tag{4.2}$$

where

$$\begin{aligned} a_n(s,t,k,l) &= \mathbb{E}[\exp(isS_{n,k} + itT_{n,k} - isS_{n,l} - itT_{n,l})] - \Phi^2(s,t) \\ &= \prod_{j=1}^{n} \varphi_j\left(s(u_{k,j}^{(n)} - u_{l,j}^{(n)}) + t(v_{k,j}^{(n)} - v_{l,j}^{(n)})\right) - \Phi^2(s,t), \end{aligned}$$

$$\begin{aligned} b_n(s,t,k,l) &= \Phi(s,t)\left(\mathbb{E}[\exp(isS_{n,k} + itT_{n,k})] + \mathbb{E}[\exp(-isS_{n,l} - itT_{n,l})] - 2\Phi(s,t)\right) \\ &= \Phi(s,t)\left(\prod_{j=1}^{n} \varphi_j(su_{k,j}^{(n)} + tv_{k,j}^{(n)}) + \prod_{j=1}^{n} \varphi_j(-su_{l,j}^{(n)} - tv_{l,j}^{(n)}) - 2\Phi(s,t)\right). \end{aligned}$$

We shall now proceed to bound $l_n(s,t,k,l)$ for all $1 \leq k, l \leq r_n$. First of all, we clearly have $l_n(s,t,k,k) \leq 2$. Moreover, we will show how to bound $a_n(s,t,k,l)$ inasmuch as the bound for $b_n(s,t,k,l)$ can be handled similarly. We obviously have

$$\max_{k \neq l} |a_n(s,t,k,l)| \leq A_n(s,t) + B_n(s,t), \tag{4.3}$$

where

$$A_n(s,t) = \max_{k\neq l}\Big|\prod_{j=1}^n \varphi_j(s(u_{k,j}^{(n)}-u_{l,j}^{(n)})+t(v_{k,j}^{(n)}-v_{l,j}^{(n)})) - \prod_{j=1}^n \exp(-\frac{1}{2}(s(u_{k,j}^{(n)}-u_{l,j}^{(n)})+t(v_{k,j}^{(n)}-v_{l,j}^{(n)}))^2)\Big|$$

and

$$B_n(s,t) = \max_{k\neq l}\left|\prod_{j=1}^n \exp(-\frac{1}{2}(s(u_{k,j}^{(n)}-u_{l,j}^{(n)})+t(v_{k,j}^{(n)}-v_{l,j}^{(n)}))^2) - \Phi^2(s,t)\right|.$$

It follows from (A_1) that for n large enough

$$0 \le \max_{j,k,l}\left|s(u_{k,j}^{(n)}-u_{l,j}^{(n)})+t(v_{k,j}^{(n)}-v_{l,j}^{(n)})\right| \le 1. \tag{4.4}$$

Moreover, by use of the well known inequality (27.13) of [4], we deduce from (A_2) and (4.4) that

$$\begin{aligned} A_n(s,t) &\le \alpha \max_{k\neq l}\sum_{j=1}^n \left|s(u_{k,j}^{(n)}-u_{l,j}^{(n)})+t(v_{k,j}^{(n)}-v_{l,j}^{(n)})\right|^3 \\ &\le \frac{a(s,t)}{(\log(1+r_n))^{1+\delta}}\Big(\max_k \sum_{j=1}^n (u_{k,j}^{(n)})^2 + \max_l \sum_{j=1}^n (v_{l,j}^{(n)})^2\Big) \\ &\le \frac{A(s,t)}{(\log(1+r_n))^{1+\delta}}. \end{aligned} \tag{4.5}$$

In order to bound $B_n(s,t)$, set

$$d_n(k,l,s,t) = \sum_{j=1}^n (s(u_{k,j}^{(n)}-u_{l,j}^{(n)})+t(v_{k,j}^{(n)}-v_{l,j}^{(n)}))^2 - 2(s^2+t^2).$$

Assumptions (A_2) and (A_3) imply that

$$\begin{aligned} |d_n(k,l,s,t)| \le s^2 &\left|\sum_{j=1}^n ((u_{k,j}^{(n)})^2+(u_{l,j}^{(n)})^2)-2\right| + t^2\left|\sum_{j=1}^n ((v_{k,j}^{(n)})^2+(v_{l,j}^{(n)})^2)-2\right| \\ &+ 2s^2\left|\sum_{j=1}^n u_{k,j}^{(n)}u_{l,j}^{(n)}\right| + 2t^2\left|\sum_{j=1}^n v_{k,j}^{(n)}v_{l,j}^{(n)}\right| \\ &+ 2\left|st\sum_{j=1}^n u_{k,j}^{(n)}v_{k,j}^{(n)} - u_{l,j}^{(n)}v_{k,j}^{(n)} + u_{l,j}^{(n)}v_{l,j}^{(n)} - u_{k,j}^{(n)}v_{l,j}^{(n)}\right| \le \frac{D(s,t)}{(\log(1+r_n))^{1+\delta}}. \end{aligned}$$

By the elementary fact that for all $a,b>0$, $|\exp(-a)-\exp(-b)| \le |a-b|$, we obtain that

$$B_n(s,t) \le \frac{B(s,t)}{(\log(1+r_n))^{1+\delta}}. \tag{4.6}$$

Consequently, we deduce from (4.3), (4.5) and (4.6) that

$$\max_{k\neq l} |a_n(s,t,k,l)| \leq \frac{C(s,t)}{3(\log(1+r_n))^{1+\delta}}.$$

One can verify that the same inequality holds for $b_n(s,t,k,l)$. Finally, we obtain from (4.2) that

$$L_n(s,t) \leq \frac{2}{r_n} + \frac{2C(s,t)}{3(\log(1+r_n))^{1+\delta}} ,$$

which completes the proof of Lemma 1. □

To prove almost sure convergence, we will use the following lemma.

Lemma 2 ([15, Theorem 1])**.** *Assume that $(Y_{n,k})$ is a sequence of uniformly bounded $\mathbb{C}$-valued and possibly dependent random variables. Let (r_n) be an increasing sequence of integers which goes to infinity. If*

$$Z_n = \frac{1}{r_n}\sum_{k=1}^{r_n} Y_{n,k}$$

and for some constant $C > 0$,

$$\mathbb{E}[|Z_n|^2] \leq \frac{C}{(\log(1+r_n))^{1+\delta}} , \tag{4.7}$$

then (Z_n) converges to zero almost surely.

Proof of Theorem 1. In order to prove the second part of Theorem 1, we apply Lemma 2 to uniformly bounded random variables

$$Y_{n,k} = \exp(isS_{n,k} + itT_{n,k}) - \Phi(s,t)$$

with $\Phi(s,t) = \exp(-(s^2+t^2)/2)$. It follows from Lemma 1 that the condition (4.7) of Lemma 2 is satisfied. Therefore, (Z_n) converges to zero almost surely. Since $Z_n = \Phi_n(s,t) - \Phi(s,t)$, this shows that $\Phi_n(s,t) \to \Phi(s,t)$ almost surely as $n \to \infty$, and we immediately deduce (2.9) from [18, Theorem 2.6]. The proof of the first part of Theorem 1 is similar, and essentially consists of taking $t = 0$ in the above calculations. Once we establish the weak convergence on a set Δ of probability 1, due to continuity of $\Phi(x)$, the convergence is uniform in x for every $\omega \in \Delta$, see [4, Exercise 14.8]. □

4.2 Proof of Theorems 2 and 3

The proofs rely on strong approximation of the partial sum processes indexed by the Lipschitz functions $f_k(x) = \cos(2\pi kx)$, compare [10, Theorems 2.1, 2.2]. We derive suitable approximation directly from the following result.

Lemma 3 (Sakhanenko [19, Theorem 1])**.** *Consider a sequence (X_n) which satisfies the assumptions of Theorem 2. Then, it exists some constant $c > 0$ such that for every $n \geq 1$ one can realize $X_1, X_2, \ldots, X_n$ on a probability space on which there are i.i.d. $\mathcal{N}(0,1)$ random variables $\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_n$ such that the partial sums*

$$S_t = \sum_{i=1}^{t} X_i \qquad \textit{and} \qquad \widetilde{S}_t = \sum_{i=1}^{t} \widetilde{X}_i$$

satisfy

$$\mathbb{E}\left[\exp\left(\frac{c}{\tau}\max_{1\leq t\leq n}|S_t-\widetilde{S}_t|\right)\right]\leq 1+\frac{n}{\tau}\,, \tag{4.8}$$

where τ is given by (2.10).

We use this keystone Lemma 3 as follows. For every $n\geq 1$, we redefine $X_1, X_2, \ldots, X_n$ onto a new probability space $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$ on which we have the i.i.d. standard normal random variables $\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_n$ which satisfy (4.8). Then, we define the sequence of weighted sums $(S_{n,k})$ by (2.6) and we also denote

$$\widetilde{S}_{n,k}=\sqrt{\frac{2}{n}}\sum_{t=1}^{n}\widetilde{X}_t\cos\left(\frac{2\pi kt}{n}\right) \tag{4.9}$$

with $k=1,2,\ldots,r_n$. The assumptions and the conclusions of Theorems 2 and 3 are not affected by such a change. One can observe from the trigonometric identities (2.2 to 2.5) that for every fixed n, the random variables $\widetilde{S}_{n,1}, \widetilde{S}_{n,2}, \ldots, \widetilde{S}_{n,r_n}$ are i.i.d with standard $\mathcal{N}(0,1)$ distribution. Therefore, for all $x, y\in\mathbb{R}$, if $x_n = x + y/\sqrt{r_n}$, we have the CLT

$$\frac{1}{\sqrt{r_n}}\sum_{k=1}^{r_n}\left(1_{\{\widetilde{S}_{n,k}\leq x_n\}}-\Phi(x_n)\right)\xrightarrow[n\to+\infty]{\mathcal{D}}\mathcal{N}\left(0,\Phi(x)(1-\Phi(x))\right). \tag{4.10}$$

This is just the normal approximation for the binomial random variable $\mathcal{B}(r_n, p_n)$ where the probability of success $p_n=\Phi(x_n)$ converges to $p=\Phi(x)$. In addition, we also have from the Kolmogorov-Smirnov Theorem

$$\sqrt{r_n}\sup_{x\in\mathbb{R}}\left|\frac{1}{r_n}\sum_{k=1}^{r_n}1_{\{\widetilde{S}_{n,k}\leq x\}}-\Phi(x)\right|\xrightarrow[n\to+\infty]{\mathcal{D}}\mathcal{L}, \tag{4.11}$$

where $\mathcal{L}$ stands for the Kolmogorov-Smirnov distribution. Furthermore, consider the corresponding empirical measure

$$\widetilde{\mu}_n=\frac{1}{r_n}\sum_{k=1}^{r_n}\delta_{\widetilde{S}_{n,k}}. \tag{4.12}$$

By Sanov's Theorem, see e.g. [8], we have for all closed sets F and for all open sets G in $\mathcal{M}_1(\mathbb{R})$,

$$\limsup_{n\to\infty}\frac{1}{r_n}\log\mathbb{P}(\widetilde{\mu}_n\in F)\leq-\inf_{\nu\in F}I(\nu)$$

and

$$\liminf_{n\to\infty}\frac{1}{r_n}\log\mathbb{P}(\widetilde{\mu}_n\in G)\geq-\inf_{\nu\in G}I(\nu).$$

Our goal is to deduce Theorems 2 and 3 from these results.

Proof of Theorem 2. For all $x\in\mathbb{R}$, denote

$$Z_n(x)=\frac{1}{\sqrt{r_n}}\sum_{k=1}^{r_n}1_{\{S_{n,k}\leq x\}}\qquad\text{and}\qquad F_n(x)=\frac{1}{r_n}\sum_{k=1}^{r_n}1_{\{S_{n,k}\leq x\}}.$$

Let $\widetilde{Z}_n(x)$ and $\widetilde{F}_n(x)$ be the corresponding sums associated with $\widetilde{S}_{n,k}$ given by (4.9) from Lemma 3. Fix $\varepsilon > 0$, and let $\varepsilon_n = \varepsilon\sqrt{2\pi/r_n}$. If

$$A_n = \Big\{ \max_{1\le k\le r_n} |S_{n,k} - \widetilde{S}_{n,k}| > \varepsilon_n \Big\}$$

and $R_n = \sqrt{r_n}\, 1_{A_n}$, we clearly have

$$\widetilde{Z}_n(x-\varepsilon_n) - R_n \le Z_n(x) \le \widetilde{Z}_n(x+\varepsilon_n) + R_n.$$

Hence, it follows from the trivial bound $|\Phi(x\pm\varepsilon_n) - \Phi(x)| \le \varepsilon_n/\sqrt{2\pi}$ that

$$\begin{cases} Z_n(x) - \sqrt{r_n}\Phi(x) \ \ge\ \widetilde{Z}_n(x-\varepsilon_n) - \sqrt{r_n}\Phi(x-\varepsilon_n) - \varepsilon - R_n, \\ Z_n(x) - \sqrt{r_n}\Phi(x) \ \le\ \widetilde{Z}_n(x+\varepsilon_n) - \sqrt{r_n}\Phi(x+\varepsilon_n) + \varepsilon + R_n. \end{cases} \tag{4.13}$$

In addition, we also have

$$\begin{cases} \sqrt{r_n}(F_n(x) - \Phi(x)) \ \ge\ \sqrt{r_n}(\widetilde{F}_n(x-\varepsilon_n) - \Phi(x-\varepsilon_n)) - \varepsilon - R_n, \\ \sqrt{r_n}(F_n(x) - \Phi(x)) \ \le\ \sqrt{r_n}(\widetilde{F}_n(x+\varepsilon_n) - \Phi(x+\varepsilon_n)) + \varepsilon + R_n, \end{cases}$$

which gives

$$\left| \sqrt{r_n}\sup_x |F_n(x) - \Phi(x)| - \sqrt{r_n}\sup_x |\widetilde{F}_n(x) - \Phi(x)| \right| \le \varepsilon + R_n \tag{4.14}$$

We now claim that (R_n) goes to zero in probability. As a matter of fact, we have

$$\max_{1\le k\le r_n} |S_{n,k} - \widetilde{S}_{n,k}| = \sqrt{\frac{2}{n}}\max_k \left| \sum_{t=1}^{n} (X_t - \widetilde{X}_t)\cos\left(\frac{2\pi kt}{n}\right)\right|$$

$$\le \sqrt{\frac{2}{n}}\max_k \left| \sum_{t=1}^{n-1} (S_t - \widetilde{S}_t)\left(\cos\left(\frac{2\pi kt}{n}\right) - \cos\left(\frac{2\pi k(t+1)}{n}\right)\right)\right| + \sqrt{\frac{2}{n}}\left|S_n - \widetilde{S}_n\right|.$$

Since the cosine is a Lipschitz function, we obtain that

$$\begin{aligned} \max_{1\le k\le r_n} |S_{n,k} - \widetilde{S}_{n,k}| \ &\le\ \sqrt{\frac{2}{n}}\max_k \sum_{t=1}^{n-1} |S_t - \widetilde{S}_t|\frac{2\pi k}{n} + \sqrt{\frac{2}{n}}|S_n - \widetilde{S}_n| \\ &\le\ \sqrt{\frac{2}{n}}(1+2\pi r_n)\max_{1\le t\le n} |S_t - \widetilde{S}_t|. \end{aligned} \tag{4.15}$$

From Lemma 3 together with Markov inequality, we obtain for large enough n so that $r_n/\sqrt{n} \le c/\tau$,

$$\begin{aligned} \mathbb{P}(A_n) &\le \mathbb{P}\left(\max_{1\le t\le n} |S_t - \widetilde{S}_t| \ge \frac{\varepsilon\sqrt{2\pi n}}{\sqrt{r_n}(1+2\pi r_n)}\right) \\ &\le \exp\left(-\frac{c\varepsilon\sqrt{2\pi n}}{\tau\sqrt{r_n}(1+2\pi r_n)}\right)\mathbb{E}\left[\exp\left(c/\tau \max_{1\le t\le n} |S_t - \widetilde{S}_t|\right)\right] \\ &\le \exp\left(\log(1+\frac{n}{\tau}) - \frac{c\varepsilon\sqrt{2\pi n}}{\tau\sqrt{r_n}(1+2\pi r_n)}\right) \to 0. \end{aligned}$$

Consequently $R_n \to 0$ in probability. Finally, as $\varepsilon > 0$ is arbitrary, we deduce (2.12) from (4.10) and (4.13) while (2.13) follows from (4.11) and (4.14), which achieves the proof of Theorem 2. $\square$

Our proof of Theorem 3 is based on the following approximation lemma.

Lemma 4 ([2, Theorem 4.9])**.** *Suppose that the sequence of random variables $(S_{n,k})$ and $(\widetilde{S}_{n,k})$ are such that, for every $\theta > 0$,*

$$\limsup_{n\to\infty} \frac{1}{r_n} \log \mathbb{E}\Big[\exp(\theta \sum_{k=1}^{r_n} |S_{n,k} - \widetilde{S}_{n,k}|)\Big] \leq 1. \tag{4.16}$$

If the sequence of empirical measures $(\widetilde{\mu}_n)$ given by (4.12) satisfies a LDP in $\mathcal{M}_1(\mathbb{R})$ with speed r_n and good rate function I, then the sequence of empirical measures (μ_n) given by (2.14) satisfies a LDP in $\mathcal{M}_1(\mathbb{R})$ with the same speed r_n and the same rate function I.

Proof of Theorem 3. The large deviation principle for the sequence $(\widetilde{\mu}_n)$ follows from Sanov's Theorem. In order to complete the proof, we only need to verify assumption (4.16) of Lemma 4. Inequality (4.15) implies that for n large enough, there is some constant $C > 0$ such that, for every $\theta > 0$,

$$\begin{aligned}
\mathbb{E}\Big[\exp(\theta \sum_{k=1}^{r_n} |S_{n,k} - \widetilde{S}_{n,k}|)\Big] &\leq \mathbb{E}\Big[\exp(\theta r_n \max_{1\leq k\leq r_n} |S_{n,k} - \widetilde{S}_{n,k}|)\Big] \\
&\leq \mathbb{E}\Big[\exp(C r_n^2 n^{-1/2} \max_{1\leq t\leq n} |S_t - \widetilde{S}_t|)\Big].
\end{aligned}$$

Since $r_n^4 = o(n)$, we infer from (4.8) that for n large enough,

$$\frac{1}{r_n} \log \mathbb{E}\Big[\exp(\theta \sum_{k=1}^{r_n} |S_{n,k} - \widetilde{S}_{n,k}|)\Big] \leq \frac{\log(1 + n/\tau)}{r_n},$$

which completes the proof of Theorem 3 as $\log(n) = o(r_n)$. $\square$

Acknowledgement The second-named author (WB) thanks S. Miller for the early version of [17], and M. Peligrad for a helpful discussion. The authors thank J. Silverstein for showing us the elementary direct proof of Corollary 7.

References

[1] Bai, Z. D. Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica 9*, 3 (1999), 611–677. With comments by G. J. Rodgers and Jack W. Silverstein; and a rejoinder by the author. MR1711663

[2] Baxter, J. R., and Jain, N. C. An approximation condition for large deviations and some applications. In *Convergence in ergodic theory and probability (Columbus, OH, 1993)*, vol. 5 of *Ohio State Univ. Math. Res. Inst. Publ.* de Gruyter, Berlin, 1996, pp. 63–90. MR1412597

[3] Berkes, I., and Csáki, E. A universal result in almost sure central limit theory. *Stochastic Process. Appl. 94* (2001), 105–134. MR1835848

[4] Billingsley, P. *Probability and measure*, third ed. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1995. A Wiley-Interscience Publication. MR1324786

[5] Bose, A., Chatterjee, S., and Gangopadhyay, S. Limiting spectral distributions of large dimensional random matrices. *J. Indian Statist. Assoc. 41*, 2 (2003), 221–259. MR2101995

[6] Bose, A., and Mitra, J. Limiting spectral distribution of a special circulant. *Statist. Probab. Lett. 60*, 1 (2002), 111–120. MR1945684

[7] Brosamler, G. An almost sure everywhere central limit theorem. *Math. Proc. Cambridge Philos. Soc. 104* (1988), 561–574. MR0957261

[8] Dembo, A., and Zeitouni, O. *Large deviations techniques and applications*, second ed., vol. 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1998. MR1619036

[9] Fay, G., and Soulier, P. The periodogram of an i.i.d. sequence. *Stochastic Process. Appl. 92*, 2 (2001), 315–343. MR1817591

[10] Grama, I., and Nussbaum, M. A functional Hungarian construction for sums of independent random variables. *Ann. Inst. H. Poincaré Probab. Statist. 38*, 6 (2002), 923–957. En l'honneur de J. Bretagnolle, D. Dacunha-Castelle, I. Ibragimov. MR1955345

[11] Hiai, F., and Petz, D. *The semicircle law, free random variables and entropy*, vol. 77 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000. MR1746976

[12] Jiang, T. Maxima of entries of Haar distributed matrices. *Probab. Theory Related Fields 131*, 1 (2005), 121–144. MR2105046

[13] Kokoszka, P., and Mikosch, T. The periodogram at the Fourier frequencies. *Stochastic Process. Appl. 86*, 1 (2000), 49–79. MR1741196

[14] Lacey, M. T., and Phillip, W. A note on the almost sure central limit theorem. *Statist. Probab. Lett. 9* (1990), 201–205. MR1045184

[15] Lyons, R. Strong laws of large numbers for weakly correlated random variables. *Michigan Math. J. 35*, 3 (1988), 353–359. MR0978305

[16] March, P., and Seppäläinen, T. Large deviations from the almost everywhere central limit theorem. *J. Theoret. Probab. 10*, 4 (1997), 935–965. MR1481655

[17] Massey, A., Miller, S. J., and Sinsheimer, J. Distribution of Eigenvalues of Real Symmetric Palindromic Toeplitz Matrices and Circulant Matrices. *Journ. Theoret. Probab.* (2005). (to appear), arXiv:math.PR/0512146.

[18] Patrizia Berti, L. P., and Rigo, P. Almost sure weak convergence of random probability measures. *Stochastics: An International Journal of Probability and Stochastic Processes 78* (2006), 91–97. MR2236634

[19] Sakhanenko, A. I. Rate of convergence in the invariance principle for variables with exponential moments that are not identically distributed. In *Limit theorems for sums of random variables*, vol. 3 of *Trudy Inst. Mat. (English translation in: Limit theorems for sums of random variables, eds. Balakrishna and A.A.Borovkov. Optimization Software, New York, 1985.).* "Nauka" Sibirsk. Otdel., Novosibirsk, 1984, pp. 4–49. MR0749757

[20] Sakhanenko, A. I. On the accuracy of normal approximation in the invariance principle [translation of Trudy Inst. Mat. (Novosibirsk) **13** (1989), Asimptot. Analiz Raspred. Sluch. Protsess., 40–66; MR 91d:60082]. *Siberian Adv. Math. 1*, 4 (1991), 58–91. Siberian Advances in Mathematics. MR1138005

[21] Schatte, P. On strong versions of the almost sure central limit theorem. *Math. Nachr. 137* (1988), 249–256. MR0968997