

STRONG APPROXIMATION FOR SET-INDEXED PARTIAL SUM PROCESSES VIA KMT CONSTRUCTIONS III

EMMANUEL RIO

ABSTRACT. We generalize the results of Komlós, Major and Tusnády concerning the strong approximation of partial sums of independent and identically distributed random variables with a finite r -th moment to the case when the parameter set is two-dimensional. The most striking result is that the rates of convergence are exactly the same as in the one-dimensional case.

1. INTRODUCTION AND RESULTS

In this paper, we continue the research started in Rio (1993a, 1993b, 1990). So, the purpose of this paper is to establish strong invariance principles for partial-sum processes indexed by a class \mathcal{S} of subsets of the unit cube $]0, 1]^d$. These processes are determined by an array $(X_i)_{i \in \mathbb{Z}^d}$ of random variables. If \mathcal{S} is any collection of subsets of the unit cube, we define the partial-sum process $\{X(\nu S) : S \in \mathcal{S}\}$ by

$$X(\nu S) = \sum_{i \in \nu S} X_i.$$

When $d = 1$, \mathcal{S} is the class of subintervals of $[0, 1]$ and $(X_i)_{i > 0}$ is a sequence of independent and identically distributed real-valued random variables with a finite r -th moment for some $r > 2$, Komlós, Major and Tusnády (1975, 1976) proved that a sequence $(Y_i)_{i > 0}$ of independent and identically distributed Gaussian variables may be constructed in such a way that

$$\sup_{S \in \mathcal{S}} |X(\nu S) - Y(\nu S)| = o(\nu^{d/r}) \text{ a.s.} \quad (1.1)$$

By the second Borel-Cantelli lemma, the rates of strong approximation appearing above are optimal [Breiman (1967)].

Our aim is to obtain optimal rates of convergence in strong invariance principles for partial-sum processes when the time parameter is multi-dimensional. In that case the optimal rates depend mostly on the geometric properties of the class \mathcal{S} [see Beck (1985) and Rio (1993a)]. When \mathcal{S} is a Vapnik-Chervonenkis class of subsets of the unit cube with uniformly

ESAIM: Probability and Statistics is an electronic journal with URL address <http://www.emath.fr/ps/>

Received by the Journal May 21, 1996. Revised April 24, 1997. Accepted for publication June 6, 1997.

© Société de Mathématiques Appliquées et Industrielles.

bounded perimeter (such as the class of Euclidean balls or the class of boxes...) Rio (1993a) obtained

$$\sup_{S \in \mathcal{S}} |X(\nu S) - Y(\nu S)| = o(\nu^{(d-1)/2} \sqrt{\log \nu}) \text{ a.s.} \quad (1.2)$$

for any $r > 2d/(d-1)$ and (1.1) for $2 < r < 2d/(d-1)$. It comes from the lower bounds of Beck (1985) that (1.2) is optimal, up to an eventual factor $\sqrt{\log \nu}$, when \mathcal{S} is the class of Euclidean balls [Rio (1993a)].

However, Beck (1985) proved that for the class of Euclidean balls and the class of lower-left orthants the lower bounds to the distances between the normalized empirical measure and the nearest Brownian measure have entirely different order of magnitude. For example, when \mathcal{S} is the class of lower-left orthants, the lower bounds to the distance between the normalized empirical measure $Z_n(\cdot)$ and the Brownian measure are of the order of $n^{-1/2}(\log n)^{(d-1)/2}$. Furthermore, in the two-dimensional case, the strong invariance principle for the normalized uniform empirical process holds with rate $n^{-1/2}(\log n)^2$ [Tusnády (1977)]. By contrast, when \mathcal{S} is the class of Euclidean balls, the lower bounds to the distance between the normalized empirical measure $Z_n(\cdot)$ and the Brownian measure are of the order of $n^{-1/(2d)}$.

So one can think that (1.2) may be improved. In this paper, we prove that (1.1) holds for any $r > 2$, when $d = 2$ and \mathcal{S} is the class of lower-left orthants. Let us now state our main result.

NOTATION 1.1. Throughout the paper, \mathbb{Z}^2 is equipped with the product order. We set $\mathbf{1} = (1, 1)$.

THEOREM 1.2. *Let $(X_i)_{i \in \mathbb{Z}^2}$ be an array of real-valued independent random variables with common distribution function F such that*

$$\int_{\mathbb{R}} x^2 dF(x) = 1 \quad \text{and} \quad \int_{\mathbb{R}} x dF(x) = 0.$$

Assume furthermore that

$$\int_{\mathbb{R}} |x|^r dF(x) < \infty \quad \text{for some } r \geq 4.$$

Then there exists an array $(Y_i)_{i \in \mathbb{Z}^2}$ of independent standard normal random variables such that

$$\sup_{m \in [1, \nu]^2} \left| \sum_{\mathbf{1} \leq i \leq m} (X_i - Y_i) \right| = o(\nu^{2/r}) \text{ a.s.}$$

Now we prove Theorem 1.2. The proof of this theorem is mainly based on a two-dimensional construction of the two arrays of independent random variables, which combines the matching method of Tusnády (1977) [this matching method can be modified to obtain a strong invariance principle for the two-dimensional Poisson process] together with a powerful extension of the results of Komlós, Major and Tusnády (1976) to non identically distributed random variables, due to Sakhanenko (1985a, 1985b). This technique was already used in Rio (1990) to obtain an error term of the order of $\nu^{2/r}(\log \nu)^2$. In this paper, we remove the extra factor $(\log \nu)^2$ using adaptive truncations as in Bass (1985) and the new levels of truncations introduced in Laurent-Bonvalot (1991). Unfortunately the matching techniques fail when $d \geq 3$ and we are unable to generalize this result to higher dimensions. Now we describe our method of construction of the two arrays.

2. CONSTRUCTION OF THE ARRAYS

Contrary to Rio (1993a, 1993b), we construct the two arrays in \mathbb{Z}^2 . The main tool for the construction is a matching argument, which is similar to the matching argument previously used by Tusnády (1977). In order to construct the arrays, we need to introduce some more notation.

DEFINITION 2.1. Let $(X_i)_{i \in \mathbb{Z}^2}$ be a sequence of real-valued random variables with mean zero. For any nonincreasing cadlag function $H : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, let H^{-1} denote the cadlag inverse function of H , which is defined by

$$H^{-1}(u) = \sup\{t \in \mathbb{R}^+ : H(t) > u\},$$

with the convention that $\sup \emptyset = 0$. For any real-valued random variable X with distribution function F , we denote by Q_X or Q_F the inverse function of $H_X : t \rightarrow \mathbb{P}(|X| > t)$.

Throughout the sequel, we need the following representation of a random variable Z with distribution function F .

NOTATION 2.2. Let Z be a random variable with distribution function F and δ be a random variable with uniform distribution, independent of Z . Let

$$H_Z(t - 0) = \mathbb{P}(|Z| \geq t) = \lim_{s \nearrow t} H_Z(s)$$

and u be the uniformly distributed random variable which is defined by

$$u = H_Z(|Z|) + \delta(H_Z(|Z| - 0) - H_Z(|Z|)).$$

Then $|Z| = Q_F(u)$ almost surely. We denote by G the bivariate distribution function of (u, Z) and by F_L the distribution function of $Z \mathbb{1}_{u \geq 2^{-L-1}}$.

In order to construct the arrays, we need the following lemma, due to Skorohod (1976).

LEMMA 2.3. *If r_1 and r_2 are random variables with values in two Polish spaces R_1 and R_2 and δ is a random variable with uniform distribution over $[0, 1]$ independent of (r_1, r_2) , then there exists a measurable function $\zeta : R_1 \times R_2 \times [0, 1] \rightarrow [0, 1]$ and a measurable function $g : R_2 \times [0, 1] \rightarrow R_1$ for which there hold the following relations: $\zeta(r_1, r_2, \delta)$ is independent of r_2 , has uniform distribution over $[0, 1]$ and $r_1 = g(r_2, \zeta(r_1, r_2, \delta))$ a.s.*

From Lemma 2.3, it follows that it suffices to construct the two arrays in another probability space. Let $(Z_i)_{i \in \mathbb{Z}^2}$ be an array of independent random variables with common distribution function F . In order to construct the arrays, we will divide \mathbb{Z}^2 into dyadic squares. Clearly, for any natural integer N , $[1, 2^{N+1}]^2 \setminus [1, 2^N]^2$ is a disjoint union of three dyadic squares $(C_{q,N})_{q=1,2,3}$ of the length 2^N . So we will define the arrays $(X_i)_{i \in C_{q,N}}$ and $(Y_i)_{i \in C_{q,N}}$ as a measurable function of $(Z_i)_{i \in C_{q,N}}$ and some auxiliary random variable. The main step of the proof is the following theorem.

THEOREM 2.4. *Let F be a distribution function satisfying the assumptions of Theorem 1 and Q be any nonincreasing function from $]0, 1]$ to \mathbb{R}^+ such that $Q \geq Q_F$ and $u \rightarrow u^{1/(2^r)}Q(u)$ is nonincreasing.*

An array $(X_i)_{i \in [1, 2^N]^2}$ of independent random variables with common distribution function F and an array $(\bar{Y}_i)_{i \in [1, 2^N]^2}$ of independent and Gaussian random variables with

$$\mathbb{E}(\bar{Y}_i) = \int_{\mathbb{R}} x dF_{2N}(x) \quad \text{and} \quad \mathbb{E}(\bar{Y}_i^2) = \int_{\mathbb{R}} x^2 dF_{2N}(x) \tag{a}$$

may be constructed in such a way that, for any u in $]0, 1]$,

$$\mathbb{P}\left(\sup_{m \in [1, 2^N]^2} \left| \sum_{1 \leq i \leq m} (X_i - \bar{Y}_i) \right| > c_r Q(u)\right) \leq 2^{2N+1} u \tag{b}$$

for some constant c_r depending only on r and on $\int_{\mathbb{R}} x^4 dF(x)$.

Proof. Clearly we may assume that the underlying probability space contains an array $(u_i, Z_i)_i$ of independent random variables with common distribution G .

In order to prove Theorem 2.4, we shall use repeatedly the following extension of the results of Komlós, Major and Tusnády (1976), due to Sakhanenko (1985a, 1985b).

THEOREM 2.5. *Let $\alpha \geq 2$ and $(\xi_i)_{i \in [1, n]}$ be a sequence of independent real-valued random variables having a finite moment of order α . Then there exists a sequence $(\eta_i)_{i \in [1, n]}$ of independent Gaussian random variables satisfying the condition*

$$\mathbb{E}(\xi_i) = \mathbb{E}(\eta_i), \quad \mathbb{E}(\xi_i^2) = \mathbb{E}(\eta_i^2) \quad \text{for any } i \in [1, n]$$

and such that

$$\mathbb{E}\left(\sup_{m \leq n} \left| \sum_{i=1}^m (\xi_i - \eta_i) \right|^\alpha\right) \leq C \alpha^{2\alpha} \sum_{i=1}^n \mathbb{E}(|\xi_i|^\alpha),$$

where C is some positive universal constant.

Let us now introduce the two-dimensional Haar basis on the space $\ell^2([1, 2^N]^2 \cap \mathbb{Z}^2)$ equipped with the usual inner product.

DEFINITION 2.6. Let $I_{j,k} =](k-1)2^j, k2^j] \cap \mathbb{Z}$. and $e_{j,k}$ be the indicator function of $I_{j,k}$. For any positive integer j we set $\tilde{e}_{j,k} = e_{j,k} - 2e_{j-1,2k}$. Then

$$\mathcal{B} = \{e_{N,1}\} \cup \{\tilde{e}_{j,k} : j \in [1, N], k \in [1, 2^{N-j}]\}$$

is an orthogonal basis of the space $\ell^2([1, 2^N] \cap \mathbb{Z})$ equipped with the usual inner product. Therefore $\mathcal{B} \otimes \mathcal{B}$ is an orthogonal basis of $\ell^2([1, 2^N]^2 \cap \mathbb{Z}^2)$ equipped with the usual inner product.

We denote by $\mathcal{B}' = \{e_{0,l} : l \in [1, 2^N]\}$ the usual orthonormal basis of $\ell^2([1, 2^N] \cap \mathbb{Z})$. Then

$$\mathcal{B}' \otimes \{e_{N,1}\} \cup \mathcal{B} \otimes \{\tilde{e}_{j,m} : j \in [1, N], m \in [1, 2^{N-j}]\}$$

is an orthogonal basis of $\ell^2([1, 2^N]^2 \cap \mathbb{Z}^2)$ equipped with the usual inner product.

Consequently, to define the Gaussian array $(\bar{Y}_i)_i$ it suffices to construct a Gaussian random vector with independent components

$$(V_{N,l} : l \in [1, 2^N], V_{L,k,N,1}, \tilde{V}_{L,k,j,m} : L, j, k, m)$$

fulfilling

$$\mathbb{E}(V_{N,l}) = 2^N \mathbb{E}(Z \mathbb{I}_{u \geq 2^{-2N-1}}), \text{Var } V_{N,l} = 2^N \text{Var}(Z \mathbb{I}_{u \geq 2^{-2N-1}}) \quad (2.1)$$

and

$$\mathbb{E}(V_{L,k,N,1}) = \mathbb{E}(\tilde{V}_{L,k,j,m}) = 0, \quad (2.2a)$$

$$2^{-L-N} \text{Var } V_{L,k,N,1} = 2^{-L-j} \text{Var } \tilde{V}_{L,k,j,m} = \text{Var}(Z \mathbb{I}_{u \geq 2^{-2N-1}}). \quad (2.2b)$$

Then the array $(\bar{Y}_i)_i$ will be uniquely defined by the equations

$$V_{N,l} = (\bar{Y} \mid e_{0,l} \otimes e_{N,1}), V_{L,k,N,1} = (\bar{Y} \mid e_{N,1} \otimes \tilde{e}_{L,k})$$

and

$$\tilde{V}_{L,k,j,m} = (\bar{Y} \mid \tilde{e}_{j,m} \otimes \tilde{e}_{L,k}).$$

We want to construct the Gaussian random array from the random variables $(u_i, Z_i)_i$. Throughout the construction, we assume that the underlying probability space is rich enough in the following sense: there exists a sequence $(\delta_l)_{l>0}$ of independent random variables having uniform distribution on $[0, 1]$ and being independent of the sequence $(u_i, Z_i)_i$.

We start by the construction of the partial sums in each row of the random variables. By Theorem 2.5, a sequence $(x_i)_{i \in [1, 2^{2N}]}$ of independent random variables with common distribution function F_{2N} and a sequence $(\eta_i)_{i \in [1, 2^{2N}]}$ of independent Gaussian random variables satisfying

$$\mathbb{E}(\eta_i) = \mathbb{E}(x_i) \quad \text{and} \quad \mathbb{E}(\eta_i^2) = \mathbb{E}(x_i^2)$$

may be constructed in such a way that

$$\mathbb{E} \left(\sup_{m \leq 2^{2N}} \left| \sum_{i=1}^m (x_i - \eta_i) \right|^{3r} \right) \leq C(3r)^{6r} \sum_{i=1}^{2^{2N}} \mathbb{E}(|x_i|^{3r}).$$

By Lemma 2.3, there exists a uniformly distributed random variable ζ independent of $(x_i)_i$ and a measurable function $\Psi_N : \mathbb{R}^{2^N} \times [0, 1] \rightarrow \mathbb{R}^{2^N}$ such that

$$\Psi_N(x_1 + \dots + x_{2^N}, \dots, x_{l2^N+1} + \dots + x_{(l+1)2^N}, \dots, \zeta) = (\eta_1 + \dots + \eta_{2^N}, \dots, \eta_{l2^N+1} + \dots + \eta_{(l+1)2^N}, \dots).$$

Let $Z_i^{2N} = Z_i \mathbb{I}_{u_i \geq 2^{-2N-1}}$. Clearly the random variables $(Z_i^{2N})_i$ are independent and identically distributed with distribution F_{2N} . Now we set $U_{N,l} = Z_{l,1}^{2N} + \dots + Z_{l,2^N}^{2N}$ and

$$(V_{N,1}, \dots, V_{N,l}, \dots) = \Psi_N(U_{N,1}, \dots, U_{N,l}, \dots, \delta_1). \quad (2.3)$$

Then $(V_{N,l})_{l \in [1,2^N]}$ is a Gaussian random vector with independent components fulfilling (2.1), and

$$\mathbb{E} \left(\sup_{m \leq 2^N} \left| \sum_{l=1}^m (U_{N,l} - V_{N,l}) \right|^{3r} \right) \leq C(3r)^{6r} 2^{2N} \mathbb{E}(|Z_1^{2N}|^{3r}). \tag{2.4}$$

In order to define the random fields $(X_i)_i$ and $(\bar{Y}_i)_i$ we will use a matching argument. Let Π be a random permutation of the space of sites, independent of the array $(u_i, Z_i)_i$. Assume furthermore that each row is invariant under the action of Π . Then $(Z_{\Pi(i)})_i$ is a sequence of independent random variables with common distribution function F and

$$U_{N,l} = Z_{\Pi(l,1)}^{2N} + \dots + Z_{\Pi(l,2^N)}^{2N}. \tag{2.5}$$

Hence (2.5) still holds if we replace Z_i by $X_i = Z_{\Pi(i)}$ in each site. So we will define the array $(X_i)_i$ and the Gaussian array $(\bar{Y}_i)_i$ from a random permutation Π , conditionally to the σ -field generated by $(u_i, Z_i)_i$.

The permutation Π will be defined as a product of random permutations. Let

$$\mathcal{F}_N = \sigma\{\delta_1, (u_i, Z_i) : i \in [1, 2^N]^2\}. \tag{2.6}$$

Clearly we may assume that the underlying probability space contains a sequence $(\Pi_L)_{L \in [1,N]}$ of independent random permutations of $[1, 2^N]^2$ for which the following properties hold: the sequence $(\Pi_L)_{L \in [1,N]}$ is independent of \mathcal{F}_N and Π_L has the structure below:

DEFINITION 2.7. For $i \leq 2^N - 2^{L-1}$, let $\tau_{L,l,i}$ denote the transposition between (l, i) and $(l, i + 2^{L-1})$. Let $(b_{L,l,i})_{(l,i) \in [1,2^N]^2}$ be an array of independent random variables with common distribution the Bernoulli law with parameter $(1/2)$. We set

$$\Pi_L = \prod_{l=1}^{2^N} \prod_{k=1}^{2^{N-L}} \prod_{i=1}^{2^{L-1}} \tau_{L,l,i}^{b_{N,l,i+(k-1)2^L}}.$$

Note that the rectangles $[1, m] \times I_{L,k}$ are invariant under the action of Π_L .

In order to define the Gaussian array, we need to introduce a nonincreasing filtration $(\mathcal{F}_L)_{L \in [0,N]}$. Throughout we assume that our auxiliary sequence $(\delta_l)_{l>0}$ of uniformly distributed random variables is independent of the σ -field generated by the array $(u_i, Z_i)_i$ and the random permutations $(\Pi_L)_{L \in [1,N]}$.

DEFINITION 2.8. For any L in $[1, N]$, we set

$$\mathcal{G}_L = \sigma\{\Pi_L, \delta_i : i \in [2^{N-L}, 2^{N-L+1}]\}$$

and we define the nonincreasing filtration $(\mathcal{F}_L)_{L \in [0,N]}$ by $\mathcal{F}_{L-1} = \mathcal{F}_L \vee \mathcal{G}_L$, where \mathcal{F}_N is defined by (2.6).

We now describe the $(N - L + 1)$ -th step of the construction, which is the construction of the random variables $V_{L,k,l}$. These random variables will

be \mathcal{F}_{L-1} -measurable and independent of \mathcal{F}_L . Here we need some additional notation.

DEFINITION 2.9. For L in $[0, N[$, let the array $(Z_i^{(L)})_i$ of independent random variables with distribution F be defined by $Z_i^{(L)} = Z_{\Pi_{L+1} \circ \dots \circ \Pi_N(i)}$. We set $Z_i^{(N)} = Z_i$. Let us define the array $(X_i)_i$ by $X_i = Z_i^{(0)}$. We also define the arrays of truncated random variables $(Z_i^{(L,K)})_i$ by

$$Z_i^{(L,K)} = Z_{\Pi_{L+1} \circ \dots \circ \Pi_N(i)} \mathbb{1}_{u_{\Pi_{L+1} \circ \dots \circ \Pi_N(i)} \geq 2^{-K-1}}.$$

We set $Z_i^{(0,K)} = X_i^K$

NOTATION 2.10. For $k \in [1, 2^{N-L}]$, we define the truncated innovations $U_{L,k,l}$ by $U_{L,k,l} = (X^{N+L-1} \mid e_{0,l} \otimes \tilde{e}_{L,k}) = (Z^{(L-1, N+L-1)} \mid e_{0,l} \otimes \tilde{e}_{L,k})$.

Let L be in $[1, N]$. Using Lemma 2.3, we will define the random vector $(V_{L,k,l})_{l \in [1, 2^N]}$ as measurable function of $(Z_{l,i}^{(L)})_{l,i}$ and $(b_{L,l,i})_{l,i}$ in such a way that this random vector will be independent of \mathcal{F}_L and Gaussian with the prescribed covariance structure.

For the sake of simplicity of notations, we fix L and take $k = 1$ for the construction. We first study the conditional law of the random vector $(U_{L,k,l})_{l \in [1, 2^N]}$ given \mathcal{F}_L . By definition of Π_L ,

$$U_{L,k,l} = \sum_{i=1}^{2^{L-1}} (1 - 2b_{L,l,i}) (Z_{l,i}^{(L, N+L-1)} - Z_{l, i+2^{L-1}}^{(L, N+L-1)}). \tag{2.7}$$

Let the random sequences $(\varepsilon_p)_p$ and $(z_p)_p$ be defined by

$$z_p = Z_{l_p, i_p}^{(L, N+L-1)} - Z_{l_p, i_p+2^{L-1}}^{(L, N+L-1)} \quad \text{and} \quad \varepsilon_p = 1 - 2b_{L, l_p, i_p} \tag{2.8}$$

where (l_p, i_p) is the unique pair of integers in $[1, 2^N] \times [1, 2^{L-1}]$ such that $p = (l_p - 1)2^{L-1} + i_p$.

Clearly the so defined random variables z_p are \mathcal{F}_L -measurable, independent and identically distributed. Hence the \mathcal{G}_{L-1} -measurable sequence $(\varepsilon_p)_{p \in [1, 2^{N+L-1}]}$ of independent Rademacher random variables is independent of the sequence $(z_p)_{p \in [1, 2^{N+L-1}]}$.

So, by Theorem 2.5 and Lemma 2.3, there exists a measurable function $\Psi : \mathbb{R}^{2^{N+L-1}} \times \mathbb{R}^{2^{N+L-1}} \times [0, 1] \rightarrow \mathbb{R}^{2^{N+L-1}}$ for which there hold the following properties:

PROPERTY 2.11. If δ is a random variable with uniform law over $[0, 1]$, independent of the sequence $(z_p, \varepsilon_p)_p$, then the random vector

$$\Psi(z_1, z_2, \dots, \varepsilon_1, \varepsilon_2, \dots, \delta) = (y_1, y_1, \dots, y_{2^{N+L-1}})$$

is a standard normal random vector conditional on the random variables $(z_p)_p$ and

$$\mathbb{E} \left(\sup_{m \leq 2^{N+L-1}} \left| \sum_{p=1}^m z_p (\varepsilon_p - y_p) \right|^{3r} \mid z_1, z_2, \dots \right) \leq C(3r)^{6r} \sum_{p=1}^{2^{N+L-1}} |z_p|^{3r}.$$

Hence the random variables $(z_p y_p)_p$ are Gaussian and independent conditional on the sequence $(z_p)_p$.

In order to define the Gaussian random variables $V_{L,k,N,1}$ and $\tilde{V}_{L,k,j,m}$, we need further notation.

NOTATION 2.12. Let $(\delta'_p)_p$ be a sequence of independent and uniformly distributed random variables, independent of $(z_1, z_2, \dots, \varepsilon_1, \varepsilon_2, \dots, \delta)$. We set

$$u'_p = H_{z_p}(|z_p|) + \delta'_p(h_{z_p}(|z_p| - 0) - H_{z_p}(|z_p|)).$$

Let $z_{l,i} = z_{i+(l-1)2^{L-1}}$, $y_{l,i} = y_{i+(l-1)2^{L-1}}$ and $u'_{l,i} = u'_{i+(l-1)2^{L-1}}$. We set

$$W_{N,1} = \sum_{l=1}^{2^N} \sum_{i=1}^{2^{L-1}} z_{l,i} y_{l,i}, \quad \tilde{W}_{N,1} = \sum_{l=1}^{2^N} \sum_{i=1}^{2^{L-1}} \tilde{e}_{N,1}(l) z_{l,i} y_{l,i}$$

and, for $j < N$ and m in $[1, 2^{N-j}]$,

$$W_{j,m} = \sum_{l=1}^{2^N} \sum_{i=1}^{2^{L-1}} e_{j,m}(l) z_{l,i} y_{l,i} \mathbb{1}_{u'_{l,i} \geq 2^{-j-L}}$$

and

$$\tilde{W}_{j,m} = \sum_{l=1}^{2^N} \sum_{i=1}^{2^{L-1}} \tilde{e}_{j,m}(l) z_{l,i} y_{l,i} \mathbb{1}_{u'_{l,i} \geq 2^{-j-L}}.$$

We also set

$$D_{N,1} = \sum_{l=1}^{2^N} \sum_{i=1}^{2^{L-1}} z_{l,i}^2, \quad \tilde{D}_{N,1} = \sum_{l=1}^{2^N} \sum_{i=1}^{2^{L-1}} \tilde{e}_{N,1}(l) z_{l,i}^2$$

and, for $j < N$ and m in $[1, 2^{N-j}]$,

$$D_{j,m} = \sum_{l=1}^{2^N} \sum_{i=1}^{2^{L-1}} e_{j,m}(l) z_{l,i}^2 \mathbb{1}_{u'_{l,i} \geq 2^{-j-L}},$$

$$\tilde{D}_{j,m} = \sum_{l=1}^{2^N} \sum_{i=1}^{2^{L-1}} \tilde{e}_{j,m}(l) z_{l,i}^2 \mathbb{1}_{u'_{l,i} \geq 2^{-j-L}}.$$

Throughout the end of this subsection, we study the distributions of the random variables conditional on the random variables $(z_p, \delta'_p)_p$. Clearly $\{W_{N,1}, \tilde{W}_{j,m} : (j, m) \in [1, N] \times [1, 2^{N-j}]\}$ is a conditionally Gaussian array. Let us study the covariance structure of this conditionally Gaussian array. We first note that

$$\text{Var } W_{j,m} = \text{Var } \tilde{W}_{j,m} = D_{j,m}, \quad \text{Cov}(W_{j,m}, \tilde{W}_{j,m}) = \tilde{D}_{j,m}. \tag{2.9}$$

Let (j, m) and (j', m') be two distinct indices. We may assume that $j \leq j'$. Then

$$\begin{aligned} \text{Cov}(\tilde{W}_{j,m}, \tilde{W}_{j',m'}) &= \sum_{l=1}^{2^N} \sum_{i=1}^{2^{L-1}} \tilde{e}_{j',m'}(l) \tilde{e}_{j,m}(l) z_{i,i}^2 \mathbb{1}_{u'_{i,i} \geq 2^{-j-L}} \\ &= \tilde{e}_{j',m'}(m2^j) \tilde{D}_{j,m}. \end{aligned} \tag{2.10}$$

In the same way one can prove that

$$\text{Cov}(W_{j,m}, \tilde{W}_{j',m'}) = \tilde{e}_{j',m'}(m2^j) D_{j,m}. \tag{2.11}$$

Note also that

$$\text{Cov}(\tilde{W}_{j,m}, W_{N,1}) = \tilde{D}_{j,m} \quad \text{and} \quad \text{Cov}(W_{j,m}, W_{N,1}) = D_{j,m}. \tag{2.12}$$

From (2.10) and (2.11) the random variables $T_{j,m} = D_{j,m} \tilde{W}_{j,m} - \tilde{D}_{j,m} W_{j,m}$ are independent, centered and Gaussian. Moreover, by (2.12), these random variables are independent of $W_{N,1}$.

It remains to normalize these new independent random variables. Let $\{\xi'_{N,1}, \tilde{\xi}'_{j,m} : (j, m) \in [1, N] \times [1, 2^{N-j}]\}$ be an array of independent standard normal random variables, independent of the σ -field generated by the previously defined random variables. Clearly

$$\text{Var } T_{j,m} = D_{j,m}(D_{j,m}^2 - \tilde{D}_{j,m}^2) = \Gamma_{j,m}.$$

Therefore we define the random variables $\tilde{V}_{j,m}$ by

$$\tilde{V}_{j,m} = \sqrt{\frac{2^{j+L} \text{Var } X_1^{2N}}{\Gamma_{j,m}}} (D_{j,m} \tilde{W}_{j,m} - \tilde{D}_{j,m} W_{j,m}) \quad \text{if } \Gamma_{j,m} \neq 0 \tag{2.13a}$$

and

$$\tilde{V}_{j,m} = \sqrt{2^{j+L} \text{Var } X_1^{2N}} \tilde{\xi}'_{j,m} \quad \text{if } \Gamma_{j,m} = 0. \tag{2.13b}$$

In the same way we set

$$V_{L,N,1} = D_{N,1}^{-1/2} \sqrt{2^{N+L} \text{Var } X_1^{2N}} W_{N,1} \quad \text{if } D_{N,1} \neq 0 \tag{2.14a}$$

and

$$V_{L,N,1} = \sqrt{2^{N+L} \text{Var } X_1^{2N}} \xi'_{j,m} \quad \text{otherwise.} \tag{2.14b}$$

By definition, the variables $\{V_{L,N,1}, \tilde{V}_{j,m} : j \in [1, N], m \in [1, 2^{N-j}]\}$ are independent, centered and Gaussian with

$$\text{Var } V_{L,N,1} = 2^{N+L} \text{Var } X_1^{2N} \quad \text{and} \quad \text{Var } \tilde{V}_{j,m} = 2^{j+L} \text{Var } X_1^{2N}$$

conditional on the random variables $(z_p, \delta'_p)_p$. Now, by Lemma 2.3 again there exists some uniformly distributed random variable u independent of the random sequence $(z_p, \varepsilon_p)_p$ and such that

$$(V_{L,N,1}, \tilde{V}_{j,m} : j \in [1.N], m \in [1, 2^{N-j}]) = \Phi_L(z_1, z_2, \dots, \varepsilon_1, \varepsilon_2, \dots, u) \tag{2.15}$$

for some measurable function Φ_L .

Let the random sequences $(\varepsilon_p^k)_p$ and $(z_p^k)_p$ be defined by

$$z_p = Z_{l_p, i_p}^{(L, N+L-1)} - Z_{l_p, i_p+2^{L-1}}^{(L, N+L-1)} \quad \text{and} \quad \varepsilon_p = 1 - 2b_{L, l_p, i_p}, \tag{2.16}$$

where (l_p, i_p) is the unique pair of integers in $[1, 2^N] \times [(k-1)2^L, (k-1/2)2^L]$ such that $p = (l_p - 1)2^{L-1} + i_p - (k-1)2^L$.

Clearly the so defined random variables z_p^k are \mathcal{F}_L -measurable, independent and identically distributed. Hence the \mathcal{G}_{L-1} -measurable sequence $(\varepsilon_p^k)_{p \in [1, 2^{N+L-1}]}$ of independent Rademacher random variables is independent of the sequence $(z_p)_{p \in [1, 2^{N+L-1}]}$.

DEFINITION 2.13. Let the random variables $V_{L,k,N,1} = (\bar{Y} \mid e_{N,1} \otimes \tilde{e}_{L,k})$ and $\tilde{V}_{L,k,j,m} = (\bar{Y} \mid \tilde{e}_{j,m} \otimes \tilde{e}_{L,k})$ be defined from $(z_1^k, z_2^k, \dots, \varepsilon_1^k, \varepsilon_2^k, \dots, \delta_{k+2^{N-L}})$ via Φ_L .

Since the sequences $(z_p^k, \varepsilon_p^k)_p$ are independent together and identically distributed when k varies, the so defined variables are \mathcal{F}_{L-1} -measurable. Furthermore they are independent, Gaussian and satisfy (2.2) conditionally to \mathcal{F}_L . Hence the array $(\bar{Y}_i)_i$ which is uniquely defined by Definition 2.13 and (2.3) is a Gaussian array fulfilling the assumptions of Theorem 2.4.

3. UPPER BOUNDS FOR THE CONSTRUCTION

NOTATION 3.1. We set $\Delta = \sup_{m \in [1, 2^N]^2} \left| \sum_{1 \leq i \leq m} (X_i - \bar{Y}_i) \right|$.

In order to control the error term, it will be convenient to use the Haar expansion of the indicator functions of the interval [see Bretagnolle and Massart (1989)].

DEFINITION 3.2. For any integer p in $[1, 2^N]$, let $k(j, p)$ be the integer such that p belongs to $I_{j, k(j, p)}$ and $\tilde{e}_j^p = \tilde{e}_{j, k(j, p)}$. Let the coefficients $\gamma_j(p)$ be defined by

$$\mathbb{1}_{[1, p]} = p2^{-N} e_{N,1} + \sum_{j=1}^N \gamma_j(p) \tilde{e}_j^p.$$

Clearly, for any array $(\xi_i)_i$ of real-valued random variables,

$$\sum_{1 \leq i \leq (m, p)} \xi_i = p2^{-N} (\xi \mid \mathbb{1}_{[1, m]} \otimes e_{N,1}) + \sum_{j=1}^N \gamma_j(p) (\xi \mid \mathbb{1}_{[1, m]} \otimes \tilde{e}_j^p). \tag{3.1}$$

Let

$$\Delta_1 = \sum_{i \in [1, 2^{2N}]^2} |X_i - X_i^{2N}| \quad \text{and} \quad \Delta_i = \Delta_{i,N} + \sum_{L=1}^N \sup_{k \in [1, 2^{N-L}]} \Delta_{i,L-1,k}$$

for $i = 2, 3$, where

$$\Delta_{2,N} = 0, \quad \Delta_{2,L-1,k} = \sum_{m=1}^{2^N} \sum_{p \in I_{L,k}} |X_{m,p}^{N+L} - X_{m,p}^{N+L-1}|,$$

$$\Delta_{3,N} = \sup_{m \leq 2^N} \left| \sum_{l=1}^m (U_{N,l} - V_{N,l}) \right|, \quad \Delta_{3,L-1,k} = \sup_{m \leq 2^N} \left| \sum_{l=1}^m (U_{L,k,l} - V_{L,k,l}) \right|$$

and $V_{L,k,l} = (\bar{Y} \mid e_{1,l} \otimes \tilde{e}_{L,k})$.

The following inequality derives straightforwardly from (3.1), using the same method as in Rio (1993a, pp. 772-773):

$$\Delta \leq \Delta_1 + \Delta_2 + \Delta_3.$$

In order to prove Theorem 2.4, we have to prove that, for $i = 1, 2, 3$,

$$\mathbb{P}(\Delta_i \geq C_r Q(u)) \leq 2^{2N+1} u \tag{3.2}$$

for some constant C_r depending only on r and on $\int_{\mathbb{R}} x^4 dF(x)$.

3.1. CONTROL OF Δ_1

Throughout the sequel, c will denote a positive constant depending only on r and on $\mathbb{E}(X_1^4)$, which may vary from line to line.

By definition of Δ_1 , $\Delta_1 = \Delta_{1,0} + \Delta_{1,1}$ where

$$\Delta_{1,0} = \sum_{i \in [1, 2^{2N}]^2} Q_F(u_i) \mathbb{1}_{u_i \leq u} \quad \text{and} \quad \Delta_{1,1} = \sum_{i \in [1, 2^{2N}]^2} Q_F(u_i) \mathbb{1}_{u_i \in]u, 2^{-2N-1}]}$$

Since $\mathbb{P}(\Delta_{1,0} \neq 0) \leq 2^{2N} u$, it suffices to prove that

$$\mathbb{P}(\Delta_{1,1} \geq cQ(u)) \leq 2^{2N} u. \tag{3.3}$$

Now

$$\mathbb{P}(Q_F(u_i) \mathbb{1}_{u_i \in]u, 2^{-2N-1}] \neq 0) \leq 2^{-2N-1}$$

and consequently we get from Theorem 2.13 in Petrov (1995, p. 63) that

$$\mathbb{E}(\Delta_{1,1}^{3r}) \leq 2c2^{2N} \int_u^1 Q^{3r}(v) dv \leq 4c2^{2N} u Q^{3r}(u) \tag{3.4}$$

because of the monotonicity of $v \rightarrow v^{3/2} Q^{3r}(v)$. Hence (3.4) follows, therefore completing the control of Δ_1 . □

In order to prove (3.2) for $i = 2, 3$, we will bound up the moments of order $3r$ of the random variables Δ_i . First of all we note that (3.2) follows from the upper bound

$$\mathbb{E}(\Delta_i^{3r}) \leq cQ^{3r}(2^{-2N-1}) \tag{3.5}$$

via Markov's inequality, seeing that, under (3.5),

$$\mathbb{P}(\Delta_i \geq c^{1/3r}Q(u)) \leq \left(\frac{Q(2^{-2N-1})}{Q(u)}\right)^{3r} \leq 2^{2N+1}u$$

for any $u \leq 2^{-2N-1}$, because of the monotonicity of $u \rightarrow u^{3/2}Q^{3r}(u)$.

Now, by the Hölder inequality,

$$\begin{aligned} \mathbb{E}(\Delta_i^{3r}) &\leq \left(\sum_{L=0}^N (1 \vee L)^{-2}\right)^{3r-1} \left(\mathbb{E}(\Delta_{i,N}^{3r}) + \sum_{L=1}^N L^{6r-2} \mathbb{E}\left(\sup_{k \leq 2^{L-1}} \Delta_{i,N-L,k}^{3r}\right)\right) \\ &\leq 3^{3r} \left(\mathbb{E}(\Delta_{i,N}^{3r}) + \sum_{L=1}^N \sum_{k=1}^{2^{L-1}} L^{6r-2} \mathbb{E}(\Delta_{i,N-L,k}^{3r})\right). \end{aligned} \tag{3.6}$$

Hence (3.5) follows from (3.6) and the estimates below via the monotonicity of $u \rightarrow u^{3/2}Q^{3r}(u)$.

PROPOSITION 3.3. *Under the assumptions of Theorem 2.4, for $i = 2, 3$ there exists some constant c such that*

$$\mathbb{E}(\Delta_{i,N}^{3r}) \leq c2^{2N} \int_{2^{-2N-1}}^1 Q^{3r}(u)du \tag{a}$$

and, for any L in $[1, N]$ and any k in $[1, 2^{N-L}]$,

$$\mathbb{E}(\Delta_{i,L-1,k}^{3r}) \leq c2^{N+L} \int_{2^{-N-L-1}}^1 Q^{3r}(u)du. \tag{b}$$

Proof. We start by the proof in the case $i = 2$.

3.2. CONTROL OF Δ_2

To prove (3.2), we only need to prove Proposition 3.3(b). Since

$$\sum_{m=1}^{2^N} \sum_{p \in I_{L,k}} \mathbb{P}(|X_{m,p}^{N+L} - X_{m,p}^{N+L-1}| \neq 0) \leq 1,$$

we get from Theorem 2.13 in Petrov (1995, p. 63) that

$$\begin{aligned} \mathbb{E}(\Delta_{2,L-1,k}^{3r}) &\leq c2^{N+L} \mathbb{E}(|X_1^{N+L} - X_1^{N+L-1}|^{3r}) \\ &\leq c2^{N+L} \int_{2^{-N-L-1}}^1 Q^{3r}(u)du. \end{aligned}$$

Hence Proposition 3.3(b) holds.

3.3. CONTROL OF Δ_3

When $i = 3$, Proposition 3.3(a) follows immediately from (2.4). To prove Proposition 3.3(b), we first note that the random variables $(\Delta_{3,L-1,k})_k$ are identically distributed. So we take $k = 1$ throughout.

NOTATION 3.4. Let $W_{L,1,l} = z_{l,1}y_{l,1} + \dots + z_{l,2^{L-1}}y_{l,2^{L-1}}$, with the same notations as in section 2. We set

$$\Delta_4 = \sup_{m \leq 2^N} \left| \sum_{l=1}^m (U_{L,1,l} - W_{L,1,l}) \right| \quad \text{and} \quad \Delta_5 = \sup_{m \leq 2^N} \left| \sum_{l=1}^m (W_{L,1,l} - V_{L,1,l}) \right|.$$

Clearly

$$\Delta_{3,L-1,1} \leq \Delta_4 + \Delta_5. \tag{3.7}$$

Now, by Property 2.11,

$$\begin{aligned} \mathbb{E}(\Delta_4^{3r}) &\leq C(6r)^{6r} 2^{N+L} \mathbb{E}(|X_1^{N+L-1}|^{3r}) \\ &\leq C(6r)^{6r} 2^{N+L} \int_{2^{-N-L}}^1 Q^{3r}(u) du. \end{aligned} \tag{3.8}$$

It remains to bound up the $3r$ -th moment of Δ_5 .

NOTATION 3.5. Let $\Delta_{6,N} = |V_{L,N,1} - W_{N,1}|$ and $\Delta_{6,j-1,m} = |\tilde{V}_{j,m} - \tilde{W}_{j,m}|$ for j in $[1, N]$ and m in $[1, 2^{N-j}]$. Set $\Delta_{7,N} = 0$. For $m = 1, 2$, let

$$\Delta_{7,N-1,m} = \sup_{p \in [(m-1)2^{N+L-2}, m2^{N+L-2}]} \left| \sum_{i=1+(m-1)2^{N+L-2}}^p z_p y_p \mathbb{I}_{2^{N+L}} u'_p < 2 \right|$$

and, for j in $[1, N - 1]$ and m in $[1, 2^{N-j}]$,

$$\Delta_{7,j-1,m} = \sup_{p \in [(m-1)2^{j+L-2}, m2^{j+L-2}]} \left| \sum_{i=1+(m-1)2^{j+L-2}}^p z_p y_p \mathbb{I}_{2^{j+L}} u'_p \in [1, 2] \right|.$$

Using the Haar expansion of $\mathbb{I}_{[1,m]}$ introduced in Definition 3.2 and arguing as in Rio (1993a, pp. 772-773), we get that

$$\Delta_5 \leq \Delta_6 + 2\Delta_7,$$

where

$$\Delta_i = \Delta_{i,N} + \sum_{j=1}^N \sup_{m \in [1, 2^{N-j}]} \Delta_{i,j-1,m} \quad \text{for } i = 6, 7.$$

Hence Proposition 3.3(b) follows from the estimates below thanks to the Hölder inequality again and the monotonicity of $u \rightarrow u^{3/2}Q^{3r}(u)$.

PROPOSITION 3.6. *Under the assumptions of Theorem 2.4, there exists some constant c such that*

$$\mathbb{E}(\Delta_{6,N}^{3r}) \leq c2^{N+L} \int_{2^{-N-L-1}}^1 Q^{3r}(u) du \tag{a}$$

and, for $i = 6, 7$, any j in $[1, N]$ and any m in $[1, 2^{N-j}]$,

$$\mathbb{E}(\Delta_{i,j-1,m}^{3r}) \leq c2^{j+L} \int_{2^{-j-L-1}}^1 Q^{3r}(u)du. \tag{b}$$

Proof. To prove Proposition 3.6, we start by the case $i = 7$. Since the random variables $(\Delta_{i,j,m})_m$ are identically distributed, we can take $m = 1$ throughout.

Conditional on the random variables (z_p, u'_p) , the random variables $\Delta_{7,j,m}$ are suprema of partial sums of independent Gaussian random variables. Hence the Paul Lévy inequality yields

$$\mathbb{E}(\Delta_{7,N-1,1}^{3r} \mid z_1, z_2, \dots, u'_1, u'_2, \dots) \leq c \left(\sum_{p=1}^{2^{N+L-2}} z_p^2 \mathbb{1}_{u'_p < 2^{1-N-L}} \right)^{3r/2} \tag{3.9}$$

and

$$\mathbb{E}(\Delta_{7,j-1,1}^{3r} \mid z_1, z_2, \dots, u'_1, u'_2, \dots) \leq c \left(\sum_{p=1}^{2^{j+L-2}} z_p^2 \mathbb{1}_{2^{j+L} u'_p \in [1,2[} \right)^{3r/2}. \tag{3.10}$$

By Theorem 2.13 in Petrov (1995, p. 63) again it follows that

$$\mathbb{E}(\Delta_{7,N-1,1}^{3r}) \leq c2^{N+L-2} \mathbb{E}(z_1^{3r}) \tag{3.11}$$

and

$$\mathbb{E}(\Delta_{7,j-1,1}^{3r}) \leq c2^{j+L-2} \int_{2^{-j-L}}^1 Q_{z_1}^{3r}(u)du. \tag{3.12}$$

Noting that

$$Q_{z_1}(2u) \leq 2Q_F(u + 2^{-N-L}) \leq 2Q_F(u), \tag{3.13}$$

we then derive Proposition 3.6 in the case $i = 7$ from (3.11) and (3.12).

It remains to prove Proposition 3.6 for $i = 6$. We first note that, by the monotonicity of $u \rightarrow u^{3/2}Q^{3r}(u)$,

$$2^{j+L} \int_{2^{-j-L-1}}^1 Q^{3r}(u)du \leq Q(2^{-j-L-1}). \tag{3.14}$$

Hence Proposition 3.6(b) in the case $i = 6$ is an immediate consequence of the exponential estimates below.

CLAIM 3.7. *Under the assumptions of Theorem 2.4, there exists some positive constant c such that*

$$\mathbb{E}\left(\exp(c\Delta_{6,N}/Q(2^{-N-L}))\right) \leq 3, \tag{a}$$

$$\mathbb{E}\left(\exp(c\Delta_{6,j-1,1}/Q(2^{-1-j-L}))\right) \leq 3. \tag{b}$$

Proof of Claim 3.7. We start by the proof of Claim 3.7(a). By Lemma 2.3, there exists some standard normal random variable ξ independent of $D_{N,1}$ and such that

$$W_{N,1} - V_{L,N,1} = \xi(\sqrt{D_{N,1}} - \sqrt{\mathbb{E}(D_{N,1})}).$$

Hence

$$2\Delta_{6,N} \leq Q(2^{-N-L}) \xi^2 + (Q(2^{-N-L}))^{-1} \frac{(D_{N,1} - \mathbb{E}(D_{N,1}))^2}{\mathbb{E}(D_{N,1}) + D_{N,1}}. \tag{3.15}$$

Since ξ is a standard normal, $\mathbb{E}(\exp(t\xi^2)) = (1 - 2t)^{-1/2}$ and consequently $\mathbb{E}(\exp(\xi^2/3)) \leq 2$. To prove Claim 3.7(a), it remains to prove that

$$\mathbb{P}((D_{N,1} - \mathbb{E}(D_{N,1}))^2 \geq Mx(\mathbb{E}(D_{N,1}) + D_{N,1})) \leq 2 \exp(-x/3), \tag{3.16}$$

where $M = 4Q^2(2^{-N-L})$. Now $D_{N,1}$ is a sum of independent and nonnegative random variables each bounded by M . It follows that

$$\text{Var } D_{N,1} \leq M\mathbb{E}(D_{N,1}).$$

So, by Bernstein's inequality [see Pollard (1984, pp. 191-194)], on the one hand

$$\begin{aligned} \mathbb{P}((D_{N,1} - \mathbb{E}(D_{N,1}))^2 \geq Mx(\mathbb{E}(D_{N,1}) + D_{N,1})) \\ \leq \mathbb{P}((D_{N,1} - \mathbb{E}(D_{N,1}))^2 \geq \mathbb{E}(D_{N,1})Mx) \\ \leq 2 \exp\left(-x/(2 + (Mx/\mathbb{E}(D_{N,1}))^{1/2})\right) \leq 2 \exp(-x/3) \end{aligned} \tag{3.17}$$

for any $x \leq M^{-1}\mathbb{E}(D_{N,1})$. On the other hand, if $Mx > \mathbb{E}(D_{N,1})$, then

$$\begin{aligned} \mathbb{P}((D_{N,1} - \mathbb{E}(D_{N,1}))^2 \geq Mx(\mathbb{E}(D_{N,1}) + D_{N,1})) \\ \leq \mathbb{P}(|D_{N,1} - \mathbb{E}(D_{N,1})| \geq Mx) \\ \leq 2 \exp\left(-x/(1 + 2(Mx)^{-1}\mathbb{E}(D_{N,1}))\right) \leq 2 \exp(-x/3) \end{aligned} \tag{3.18}$$

by Bernstein's inequality again. Hence (3.16) holds and as a byproduct, we have:

$$\mathbb{P}(\Delta_{6,N} \geq Mx) \leq 4 \exp(-x/3). \tag{3.19}$$

Claim 3.7(a) follows.

We now prove Claim 3.7(b). For the sake of simplicity of notations, we omit the index 1 throughout. Clearly the decomposition below holds:

$$\tilde{V}_j - \tilde{W}_j = \tilde{V}_j - (\tilde{W}_j - \frac{\tilde{D}_j}{D_j}W_j) - \frac{\tilde{D}_j}{D_j}W_j$$

with the convention that $(\tilde{D}_j/D_j) = 0$ if $D_j = 0$. By Lemma 2.3, there exists some standard normal random vector (ξ, ξ') independent of (D_j, \tilde{D}_j) and such that

$$W_j = \xi^t \sqrt{D_j}, \tilde{V}_j = \xi \sqrt{2^{j+L} \text{Var } X_1^{2N}} \text{ and } \tilde{W}_j - \frac{\tilde{D}_j}{D_j}W_j = \xi \left(\frac{D_j^2 - \tilde{D}_j^2}{D_j} \right)^{1/2}.$$

Therefrom

$$\Delta_{6,j-1,1} \leq |\Delta_{8,j-1}| + |\Delta_{9,j-1}| + |\Delta_{10,j-1}|,$$

where

$$\Delta_{8,j-1} = -\xi' \sqrt{\tilde{D}_j^2/D_j} + \xi(\sqrt{D_j} - \sqrt{(D_j^2 - \tilde{D}_j^2)/D_j})$$

and

$$\Delta_{9,j-1} = \xi(\sqrt{\mathbb{E}(D_j)} - \sqrt{D_j}), \quad \Delta_{10,j-1} = \xi(\sqrt{2^{j+L} \text{Var } X_1^{2N}} - \sqrt{\mathbb{E}(D_j)}).$$

We first bound up $|\Delta_{8,j-1}|$. Since (ξ, ξ') is independent of (D_j, \tilde{D}_j) and has the distribution $N(0, I_2)$, there exists some standard normal variable ξ'' such that

$$\Delta_{8,j-1}^2 = 2(D_j - \sqrt{D_j^2 - \tilde{D}_j^2})\xi''^2.$$

Hence

$$|\Delta_{8,j-1}| \leq Q(2^{-j-L-1}) \xi''^2 + \tilde{D}_j^2/(D_j Q(2^{-j-L-1})),$$

so that it only remains to obtain the inequality below:

$$\mathbb{P}(\tilde{D}_j^2 \geq M'x D_j) \leq 2 \exp(-x/8) \tag{3.20}$$

where $M' = 4Q^2(2^{-j-L-1})$.

Let $(\varepsilon'_p)_p$ be a sequence of independent symmetric signs, independent of the sequence $(z_p, u'_p)_p$ and

$$\tilde{D}_j = \sum_{p=1}^{2^{j+L-2}} \varepsilon'_p (z_p^2 \mathbb{1}_{u'_p \geq 2^{-j-L}} - z_{p+2^{j+L-2}}^2 \mathbb{1}_{u'_{p+2^{j+L-2}} \geq 2^{-j-L}}).$$

Arguing as in Section 2, it is easily seen that the random vector (\tilde{D}'_j, D_j) has the same distribution as (\tilde{D}_j, D_j) . Now, by (3.13) and the Hoeffding inequality [see Pollard (1984, pp. 191-194)],

$$\mathbb{P}(|\tilde{D}'_j|^2 \geq t \mid z_1, z_2 \dots) \leq 2 \exp(-t(8M'D_j)^{-1}),$$

therefore completing the proof of (3.20).

To bound up the random variable $|\Delta_{9,j-1}|$, we proceed exactly as in the proof of Claim 3.7(a), In that way, we get that

$$\mathbb{P}(|\Delta_{9,j-1}| \geq M'x) \leq 4 \exp(-x/3). \tag{3.21}$$

Since $\Delta_{10,j-1}$ is a centered Gaussian random variable, it suffices to bound up the variance of $\Delta_{10,j-1}$. Now

$$\mathbb{E}(D_j) = 2^{j+L-1} \int_{2^{-j-L}}^1 Q_{z_1}^2(u) du.$$

Moreover

$$0 \leq 1 - \text{Var } X_1^{2N} \leq 2 \int_0^{2^{-2N-1}} Q_F^2(u) du \leq 2^{(1-2N)/2} \|X_1\|_4^2$$

by the Cauchy-Schwarz inequality. Hence there exists some positive integer N_0 depending only on $\int_{\mathbb{R}} x^4 dF(x)$ such that $\text{Var } X_1^{2N} \geq 1/2$ for $N \geq N_0$. Then

$$\begin{aligned} 2^{-j-L-2} \text{Var} \Delta_{10,j-1} &\leq \frac{1}{2} \left(\text{Var } X_1^{2N} - \frac{1}{2} \int_{2^{-j-L}}^1 Q_{z_1}^2(u) du \right)^2 \tag{3.22} \\ &\leq \left(\text{Var } X_1^{2N} - \int_{2^{-N-L}}^1 Q_F^2(u) du \right)^2 + \frac{1}{4} \left(\int_0^{2^{-j-L}} Q_{z_1}^2(u) du \right)^2 \\ &\leq \left(\text{Var } X_1^{2N} - \int_{2^{-N-L}}^1 Q_F^2(u) du \right)^2 + \left(4 \int_0^{2^{-j-L-1}} Q_F^2(u) du \right)^2 \end{aligned}$$

by (3.13). Now

$$\text{Var } X_1^{2N} - \int_{2^{-N-L}}^1 Q_F^2(u) du = \int_{2^{-2N-1}}^{2^{-N-L}} Q_F^2(u) du - (\mathbb{E} X_1^{2N})^2,$$

and therefore

$$\begin{aligned} \left| \text{Var } X_1^{2N} - \int_{2^{-N-L}}^1 Q_F^2(u) du \right| &\leq \int_{2^{-2N-1}}^{2^{-N-L}} Q_F^2(u) du + \mathbb{E}((X_1 - X_1^{2N})^2) \\ &\leq \int_0^{2^{-N-L}} Q_F^2(u) du. \tag{3.23} \end{aligned}$$

Hence, by (3.22) and (3.23),

$$\text{Var } \Delta_{10,j-1} \leq 68 \cdot 2^{j+L} \left(\int_0^{2^{-j-L}} Q_F^2(u) du \right)^2 \leq 68 \int_0^1 Q_F^4(u) du$$

by the Cauchy-Schwarz inequality. Claim 3.7(b) holds, therefore completing the proof of Proposition 3.3(b) for $i = 3$. Hence Theorem 2.4 holds.

4. STRONG APPROXIMATION

In this section, we derive Theorem 1.2 from Theorem 2.4. First of all, we prove that there exists some quantile function Q satisfying the assumptions of Theorem 2.4 and such that Q^r is integrable.

LEMMA 4.1. *Let Q be defined by $\sqrt{u}Q^r(u) = \sup_{t \in [u,1]} \sqrt{t}Q_F^r(t)$ for any u in $]0, 1[$. Then Q satisfies the assumptions of Theorem 2.4 and*

$$\int_0^1 Q^r(u) du \leq 2 \int_{\mathbb{R}} |x|^r dF(x).$$

Proof. Since Q_F^r is a nonincreasing function from $]0, 1[$ into \mathbb{R}^+ , there exists a Radon measure ν such that

$$Q_F^r(t) = \int_{\mathbb{R}} \mathbb{1}_{t < x} d\nu(x). \tag{4.1}$$

Hence

$$Q^r(u) \leq \int_{\mathbb{R}} (x/u)^{1/2} d\nu(x). \tag{4.2}$$

Then, integrating (4.2) and applying Fubini's theorem, we get Lemma 4.1. \square

Let Q be the function defined in Lemma 4.1. Applying Theorem 2.4 repeatedly, we may construct an array $(X_i)_{i \in \mathbb{Z}}$ of independent random variables with common distribution function F and an array $(\bar{Y}_i)_{i \in \mathbb{Z}}$ of independent Gaussian random variables with

$$\mathbb{E}(\bar{Y}_i) = \mathbb{E}(X_i^{2^N}), \text{ Var } \bar{Y}_i = \text{Var } X_i^{2^N} \text{ for any } i \in [1, 2^{N+1}]^2 \setminus [1, 2^N]^2$$

and $\mathbb{E}(Y_1) = \mathbb{E}(X_1^0)$, $\text{Var } Y_1 = \text{Var } X_1^0$ in such a way that, for any nonnegative integer L and any u in $]0, 1]$,

$$\mathbb{P}\left(\sup_{m \in [1, 2^{L+1}]^2} \left| \sum_{\substack{1 \leq i \leq m \\ i \notin [1, 2^L]^2}} (X_i - \bar{Y}_i) \right| > 3c_r Q(u)\right) \leq 3 \cdot 2^{2L+1} u \tag{4.3}$$

and $\mathbb{P}(|X_1 - Y_1| > c_r Q(u)) \leq 2u$. Now, applying (4.3) with $u = u_L = 2^{N-L}u$, summing on $L \in [0, N - 1]$ and noting that $\sum_{K \geq 0} Q(2^K u) \leq 4rQ(u)$,

we get that

$$\mathbb{P}\left(\sup_{m \in [1, 2^N]^2} \left| \sum_{1 \leq i \leq m} (X_i - \bar{Y}_i) \right| > 16rc_r Q(u)\right) \leq 2^{2N+4} u. \tag{4.4}$$

Let $(Y_i)_i$ be the array of independent standard normal random variables which is defined from $(\bar{Y}_i)_i$ by the normalization equations

$$\bar{Y}_i = \mathbb{E}(\bar{Y}_i) + \sqrt{\text{Var } \bar{Y}_i} Y_i. \tag{4.5}$$

We set

$$D_N = \sup_{m \in [1, 2^N]^2} \left| \sum_{1 \leq i \leq m} (X_i - \bar{Y}_i) \right| \text{ and } R_N = \sup_{m \in [1, 2^N]^2} \left| \sum_{1 \leq i \leq m} (Y_i - \bar{Y}_i) \right|.$$

Clearly

$$\sup_{m \in [1, 2^N]^2} \left| \sum_{1 \leq i \leq m} (X_i - Y_i) \right| \leq D_N + R_N.$$

To bound up R_N we note that, for any i in $[1, 2^{L+1}]^2 \setminus [1, 2^L]^2$,

$$|\mathbb{E}(\bar{Y}_i)| \leq \int_0^{2^{-2L-1}} Q_F(u) du \leq 2^{(2L+1)(-1+1/r)} \left(\int_0^{2^{-2L-1}} Q_F^r(u) du \right)^{1/r} \tag{4.6}$$

and

$$\sqrt{\text{Var}(Y_i - \bar{Y}_i)} \leq 2 \int_0^{2^{-2L-1}} Q_F^2(u) du \leq 2^{1-L} \|X_1\|_4^2 \tag{4.7}$$

by the Cauchy-Schwarz inequality [see Einmahl (1987) for a detailed proof of (4.7)]. Hence

$$\sum_{i \in [1, 2^N]^2} |\mathbb{E}(\bar{Y}_i)| = o(2^{2N/r}) \text{ as } N \rightarrow \infty \tag{4.8}$$

and

$$\sum_{i \in [1, 2^N]^2} \text{Var}(Y_i - \bar{Y}_i) \leq 2N \|X_1\|_4^2. \tag{4.9}$$

Since the class of lower-left orthants is a Vapnik-Chervonenkis class of subsets of \mathbb{R}^2 [see Assouad (1983)], it follows from classical estimates for the tail distribution of the supremum of a Gaussian process [see Dudley (1973)] that

$$\mathbb{E} \left(\sup_{m \in [1, 2^N]^2} \left| \sum_{1 \leq i \leq m} (\bar{Y}_i - \mathbb{E}(\bar{Y}_i) - Y_i) \right|^r \right) \leq cN^{r/2}. \tag{4.10}$$

Hence, by (4.8), (4.10) and Markov's inequality,

$$\mathbb{P}(R_N > \varepsilon 2^{2N/r}) \leq c\varepsilon^{-r} N^{r/2} 2^{-2N}$$

for N large enough and consequently $R_N = o(2^{2N/r})$ a.s.

It remains to prove that $D_N = o(2^{2N/r})$ a.s. Let H denote the cadlag inverse of Q . Clearly H is the tail function of $Q(U)$ for any uniformly distributed random variable U . Let us apply (4.4) with $u = H(\varepsilon 2^{2N/r})$. Since $Q(H(u)) \leq u$, we get that

$$\mathbb{P}(D_N > 16rc_r \varepsilon 2^{2N/r}) \leq 2^{2N+4} H(\varepsilon 2^{2N/r}) \tag{4.11}$$

and from now on the end of the proof is straightforward using the integrability of $x \rightarrow H(x^{1/r})$ and the Borel-Cantelli lemma. \square

ACKNOWLEDGEMENTS

I would like to thanks the referees for a careful reading of the paper and for pointing out errors in a previous version of the manuscript.

REFERENCES

ASSOUAD, P., (1983), Densité et dimension. *Ann. Inst. Fourier* **33** 233-282.
 BASS, R. F., (1985), Law of the iterated logarithm for partial-sum processes with finite variance. *Z. Wahrsch. verw. Gebiete* **70** 591-608.
 BECK, J., (1985), Lower bounds on the approximation of the multivariate empirical process. *Z. Wahrsch. verw. Gebiete* **70** 289-306.
 BREIMAN, L., (1967), On the tail behavior of sums of independent random variables. *Z. Wahrsch. verw. Gebiete* **9** 20-25.
 BRETAGNOLLE, J. and MASSART, P., (1989), Hungarian constructions from the non-asymptotic viewpoint. *Ann. Probab.* **17** 239-256.
 DUDLEY, R. M., (1973), Sample functions of the Gaussian process. *Ann. Probab.* **1** 66-103.
 EINMAHL, U., (1987), Strong invariance principle for partial sums of independent random vectors. *Ann. Probab.* **15** 1419-1440.

- KOMLÓS, J., MAJOR, P. and TUSNÁDY, G., (1975), An approximation of partial sums of independent rv's and the sample df. I. *Z. Wahrsch. verw. Gebiete* **32** 111-131.
- KOMLÓS, J., MAJOR, P. and TUSNÁDY, G., (1976), An approximation of partial sums of independent rv's and the sample df. II. *Z. Wahrsch. verw. Gebiete* **34** 35-58.
- LAURENT-BONVALOT, F., (1991) Approximation forte de processus empiriques et applications. *Thèse de doctorat, chapitre 2. Université Paris Sud.*
- PETROV, V. V., (1995), *Limit theorems of probability theory: sequences of independent random variables.* Oxford university press. Oxford.
- POLLARD, D., (1984), *Convergence of stochastic processes.* Springer Series in Statistics. Springer, Berlin.
- RIO, E., (1990), Approximation forte pour des processus de sommes partielles indexés par les quadrants. *Thèse de doctorat, chapitre 3. Université Paris Sud.*
- RIO, E., (1993a), Strong approximation for set-indexed partial sum Processes, via KMT constructions I. *Ann. Probab.* **21** 759-790.
- RIO, E., (1993b), Strong approximation for set-indexed partial sum Processes, via KMT constructions II. *Ann. Probab.* **21** 1706-1727.
- SAKHANENKO, A.I., (1985a), Rate of convergence in the invariance principle for non identically distributed variables with exponential moments. In *Advances in probability theory. Limit theorems for sums of random variables* (ed. A. A. Borovkov). 2-73. Optimization Software Inc., Publication Division. New York.
- SAKHANENKO, A. I., (1985b), Estimates in the invariance principle. *Trudy Inst. Mat. Sibirsk. Otdel, Akad. Nauka SSSR, Novosibirsk.* **5** 27-44.
- SKOROHOD, A. V., (1976), On a representation of random variables. *Theory Probab. Appl.* **21** 628-632.
- TUSNÁDY, G., (1977), A remark on the approximation of the sample DF in the multidimensional case. *Period. Math. Hung.* **8** 53-55.

EMMANUEL RIO, URA 0743 CNRS, UNIVERSITÉ PARIS SUD, BÂT. 425, MATHÉMATIQUE, F-91405 ORSAY CEDEX. E-MAIL: Emmanuel.Rio@math.u-psud.fr