THE INVARIANCE PRINCIPLE FOR A CLASS OF DEPENDENT RANDOM FIELDS

UDC 519.21

D. V. PORYVAĬ

ABSTRACT. Sufficient conditions for the tightness of a family of distributions of partial sum set-indexed processes constructed from symmetric random fields are obtained in this paper. We require that the moments of order s , $s > 2$, exist. The dependence structure of the field is described by the β_1 -mixing coefficients decreasing with a power rate. Assuming that a field is stationary and applying a result of D. Chen (1991) on the convergence of finite-dimensional distributions of the processes we obtain the invariance principle.

1. INTRODUCTION

The asymptotic behavior of partial sum set-indexed processes is studied in a number of papers (see, for example, [2], [3], [6], [7], [11]). The results of the above papers are obtained for smoothed partial sums defined on a subclass $\mathfrak{A}$ of the Borel sets of the cube $[0, 1]^d$. More precisely, let $X = \{X_j, j \in \mathbf{Z}^d\}$ be a stationary field defined on some probability space $(\Omega, \mathfrak{D}, \mathbf{P})$. Consider the processes

$$Z_n(A) = n^{-d/2} \sum_{j \in \mathbf{Z}^d} b_{nj}(A) X_j, \quad A \in \mathfrak{A}, \quad n \in \mathbf{N},$$

where $j = (j_1, \dots, j_d)$, $C_j = (j_1 - 1, j_1] \times \dots \times (j_d - 1, j_d]$ is a unit cube, $|\cdot|$ is Lebesgue measure in $\mathbf{R}^d$, $b_{nj}(A) = |(nA) \cap C_j|$, and $nA = \{nx: x \in A\}$.

Let $\overline{\mathfrak{A}}$ be the closure of $\mathfrak{A}$ with respect to the pseudometric $d_L(A, B) = |A \Delta B|$ defined for $A, B \in \overline{\mathfrak{A}}$. Denote by $C(\overline{\mathfrak{A}})$ the space of real continuous functions on $\overline{\mathfrak{A}}$ equipped with the sup-norm.

We consider symmetric fields X (see, for example, §4.2 in [12]) constructed from identically distributed random variables X_j . Recall that a field X is called symmetric if the finite-dimensional distributions of the fields X and

$$\varepsilon X = \{\varepsilon_j X_j, j \in \mathbf{Z}^d\}$$

coincide where $\varepsilon = \{\varepsilon_j, j \in \mathbf{Z}^d\}$ is the Rademacher field that does not depend on X .

We are interested in obtaining sufficient conditions for the invariance principle, that is, conditions for the convergence in distribution in the space $C(\overline{\mathfrak{A}})$ of the processes Z_n to the process $\sqrt{\eta}Z$ as $n \rightarrow \infty$, where Z is a standard Brownian motion,

$$\eta = \sum_{k \in \mathbf{Z}^d} \mathbf{E}(X_0 X_k \mid \mathfrak{I}),$$

2000 *Mathematics Subject Classification.* Primary 60F17, 60G60.
Supported in part by the RFFI grant 03-01-00724.

and $\mathfrak{I}$ is the σ -algebra of events that are invariant under the shifts of the field X (Z does not depend on η). The standard Brownian motion Z is defined as a mean zero Gaussian process with sample paths in $C(\overline{\mathfrak{A}})$ and such that $E(Z(A)Z(B)) = |A \cap B|$ for $A, B \in \overline{\mathfrak{A}}$. The existence of such a process Z is proved in [9] under some entropy conditions posed on the class $\mathfrak{A}$. These conditions are given in terms of the so-called entropy with inclusion.

We need more notation to describe the dependence structure of the field X . Let $\beta(\sigma_1, \sigma_2)$ be the coefficient of absolute regularity of the σ -algebras $\sigma_1, \sigma_2 \subset \mathfrak{D}$ (see, for example, [1]). Put

$$\rho(G_1, G_2) = \inf\{\|x - y\| : x \in G_1, y \in G_2\}$$

where $G_1, G_2 \subset \mathbf{Z}^d$ and $\|x\| = \max_{1 \leq i \leq d} |x_i|$ for $x = (x_1, \dots, x_d) \in \mathbf{R}^d$. For $n \in \mathbf{N}$ and $k, m \in \mathbf{N} \cup \{\infty\}$ we introduce the mixing coefficients:

$$(1) \quad \beta_X(n, k, m) = \sup\{\beta(\sigma_X(G_1), \sigma_X(G_2)) : \sharp(G_1) \leq k, \sharp(G_2) \leq m, \rho(G_1, G_2) \geq n\}$$

where the sets G_1 and G_2 are separated by some hyperplane in $\mathbf{R}^d$, $\sigma_X(G)$ is the σ -algebra generated by the field X in the set $G \subset \mathbf{Z}^d$, and $\sharp(G)$ denotes the cardinality of G . For $x, y, z \geq 1$, we put $\beta_X(x, y, z) = \beta_X([x], [y], [z])$ where $[\cdot]$ is the integer part of a number.

The convergence of finite-dimensional distributions of the processes Z_n is proved in [7] under a condition on the dependence of $\sigma_X(\{j\})$ and $\sigma_X(G)$, such that

$$\rho(\{j\}, G) \geq n, \quad n \in \mathbf{N}.$$

At the same time, the condition on the tightness of the family of distributions of the processes Z_n in $C(\overline{\mathfrak{A}})$ relies on the dependence of $\sigma_X(\{i, j\})$ and $\sigma_X(G)$ such that

$$\rho(\{i, j\}, G) \geq n, \quad n \in \mathbf{N}.$$

The main aim of this paper is to obtain a sufficient condition for the tightness of the distributions of Z_n in terms of the coefficients $\beta_X(n, 1, m)$, so we avoid two-point subsets of $\mathbf{Z}^d$ in (1) (instead, the condition will involve those G_1 for which $\sharp(G_1) = 1$). There are, of course, some extra conditions on the moments of the field X and on the structure of the class $\mathfrak{A}$.

To solve the problem we generalize the method of the paper [3] to the case of weakly dependent fields. This generalization is due to the so-called reconstruction technique (see, for example, §1.2.2 in [8]) developed in the paper [5]. Our method also uses truncation of the original random variables, appropriate approximations of elements of the class $\mathfrak{A}$, and some maximal inequalities.

2. ENTROPY CONDITIONS

We introduce the following conditions posed on the entropy with inclusion for a family $\mathfrak{A}$. Let $g(\varepsilon)$, $\varepsilon \in [0, 1]$, be an increasing function such that $g(\varepsilon) \rightarrow \infty$ as $\varepsilon \rightarrow 0$ and

$$(2) \quad \int_0^1 (\varepsilon^{-1} H(\varepsilon))^{1/2} g(\varepsilon) d\varepsilon < \infty$$

where $H(\varepsilon) = \log N_I(\varepsilon, \mathfrak{A}, d_L)$ and $N_I(\varepsilon, \mathfrak{A}, d_L)$ denotes the minimal number $k \geq 1$ for which there are measurable sets $A_i^{(1)}$ and $A_i^{(2)}$ of $[0, 1]^d$, $1 \leq i \leq k$, such that for all $A \in \mathfrak{A}$ there exists i such that $A_i^{(1)} \subset A \subset A_i^{(2)}$ and $|A_i^{(2)} \setminus A_i^{(1)}| \leq \varepsilon$.

It follows from metric entropy condition (2) that $\overline{\mathfrak{A}}$ is a compact set, and therefore $C(\overline{\mathfrak{A}})$ is a separable space. We define the exponent r of the metric entropy of $\mathfrak{A}$ by $r = \inf\{s > 0 : \log N_I(\varepsilon, \mathfrak{A}, d_L) = O(\varepsilon^{-s}) \text{ as } \varepsilon \rightarrow 0\}$. It is easy to see that metric entropy condition (2) holds if $r < 1$. Therefore all the classes of sets studied in [3] satisfy condition (2).

3. TIGHTNESS

Here is the main result of the paper.

Theorem 1. *Let X be a symmetric field of identically distributed random variables X_j . Assume that $\mathfrak{A}$ is a family of subsets of $[0, 1]^d$ satisfying metric entropy condition (2). Moreover let*

- (i) $\mathbb{E} |X_0|^s < \infty$ for some $s > 2$;
- (ii) $\limsup_{\varepsilon \downarrow 0} f^{2\tau}(\varepsilon) \cdot \beta_X(g^{1/d}(\varepsilon) - 1, 1, f^{2\tau}(\varepsilon)/g(\varepsilon)) < \infty$ where $\tau = s/(s-2)$ and $f(\varepsilon) = (\varepsilon^{-1}H(\varepsilon))^{1/2}g(\varepsilon)$.

Then the family of distributions of the processes $Z_n = \{Z_n(A) : A \in \mathfrak{A}\}$ is tight in the space $C(\overline{\mathfrak{A}})$.

Proof. Given $\delta > 0$ consider the family of sets

$$\mathbb{S}_\delta = \{A \setminus B : A, B \in \mathfrak{A} \text{ such that } |A \setminus B| \leq \delta\}.$$

Since $N_I(\varepsilon, \mathbb{S}_\delta, d_L) \leq N_I(\varepsilon/2, \mathfrak{A}, d_L)^2$, condition (2) holds for $\mathbb{S}_\delta$. Thus the process Z is continuous in the space $(\mathbb{S}_\delta, d_L)$. Let

$$\|f\|_D = \sup_{x \in D} |f(x)|$$

for all real functions f defined on the set D . The functional

$$\mathbf{w}(Z_n, \delta) = \sup\{|Z_n(B) - Z_n(C)| : B, C \in \mathfrak{A}, |B \triangle C| < \delta\}, \quad \delta > 0,$$

is the modulus of continuity of the process Z_n in the space $C(\overline{\mathfrak{A}})$. It is clear that Z_n is a random element in the space $C(\overline{\mathfrak{A}})$. Since $\mathbf{w}(Z_n, \delta) \leq 2\|Z_n\|_{\mathbb{S}_\delta}$, the family of distributions is tight if for all $M > 0$

$$(3) \quad \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\|Z_n\|_{\mathbb{S}_\delta} > M) = 0.$$

We split the proof of (3) into several steps.

Step 1. First we truncate the random variables X_j . According to (i) there is a sequence $\{c_n\}_{n \geq 1}$ such that $c_n \rightarrow 0$ and $\lim_{n \rightarrow \infty} n^d \mathbb{P}(|X_0|^s > c_n^2 n^d) = 0$. If

$$b_n^2 = n^{2d/s-d} c_n^{4/s},$$

then $\lim_{n \rightarrow \infty} n^d \mathbb{P}(X_0^2 > b_n^2 n^d) = 0$. Fix constants M and $\delta > 0$ and put

$$(4) \quad \gamma_n = \inf \{\gamma > 0 : n^d \mathbb{P}(X_0^2 > \gamma^2 n^d) < M \delta b_n^{-1}\} \wedge b_n.$$

Let

$$Z'_n = n^{-d/2} \sum_j b_{nj} X_j \mathbb{1}_{\{|X_j| > b_n n^{d/2}\}}, \quad Z''_n = n^{-d/2} \sum_j b_{nj} X_j \mathbb{1}_{\{\gamma_n n^{d/2} < |X_j| \leq b_n n^{d/2}\}}.$$

If the sequence $\{b_n\}$ is defined as above, then

$$n^d \mathbb{P}(|X_0| > \gamma_n n^{d/2}) \leq M \delta b_n^{-1}$$

and $b_{nj}(\cdot) \leq 1$. Thus by the Chebyshev inequality

$$\mathbb{P}(\|Z'_n\|_{\mathbb{S}_\delta} > M) \leq \mathbb{P}\left(\bigcup_{0 \leq j \leq n-1} \{|X_j| > b_n n^{d/2}\}\right) = o(1),$$

$$\mathbb{P}(\|Z''_n\|_{\mathbb{S}_\delta} > M) \leq M^{-1} \mathbb{E} \|Z''_n\|_{\mathbb{S}_\delta} \leq M^{-1} b_n \sum_{0 \leq j \leq n-1} \mathbb{P}(|X_j| > \gamma_n n^{d/2}) \leq \delta.$$

Here $\mathbf{1} = (1, \dots, 1)$ and $\mathbf{0} \leq j \leq n\mathbf{1}$ is equivalent to the inequalities $0 \leq j_i \leq n$ for all $i = 1, \dots, d$. Therefore condition (3) holds in the general case if it holds for the truncated partial sum process

$$Z_n^T = Z_n - Z'_n - Z''_n = n^{-d/2} \sum_j b_{nj} X_j \mathbb{1}_{\{|X_j| \leq \gamma_n n^{d/2}\}}$$

instead of Z_n .

Step 2. We apply the so-called stratification procedure where the interval $(0, \gamma_n n^{d/2}]$ is partitioned into appropriate subintervals. Let $F(x) = \mathbb{P}(|X_0| > x)$, $x \in \mathbf{R}_+$. Put

$$Q_F(u) = \inf\{x \geq 0: F(x) \leq u\}, \quad u \in (0, 1].$$

For $0 < \beta < 1$ let $\mu_k = \beta^k$ and $a_k = Q_F(\mu_k)$, $k = 0, \dots, k_n$, where

$$k_n = \max\left\{k: a_k < \gamma_n n^{d/2}\right\}.$$

Then $\mathbb{P}(|X_0| > a_k) \leq \mu_k$ for $k = 0, \dots, k_n$. Consider the intervals $J_k = (a_k, a_{k+1}]$, $0 \leq k \leq k_n$, where $a_{k_n+1} = \gamma_n n^{d/2}$.

If $Y_j = |X_j|$, then

$$Z_n^T = \sum_{k \leq k_n} \theta_k \nu_{nk}$$

where $\theta_k = a_{k+1} \mu_k^{1/2}$ and

$$\nu_{nk}(A) = (n^d \mu_k)^{-1/2} \sum_j b_{nj}(A) a_{k+1}^{-1} X_j \mathbb{1}_{\{Y_j \in J_k\}}.$$

It is clear that $a_{k+1}^{-1} Y_j \leq 1$ if $Y_j \in J_k$. It is easy to see that

$$\sum_{k=0}^{k_n} \theta_k^2 = \sum_{k=0}^{k_n} a_{k+1}^2 \mu_k \leq \sum_{k=0}^{\infty} Q_F^2(\beta^{k+1}) \beta^k \leq 1/(\beta(1-\beta)) \mathbb{E} X_0^2 < \infty$$

for all $n \in \mathbf{N}$.

Below we use the functions f and H that are introduced above. The only properties of the functions H and f we use in the proof are that $H(\varepsilon)$ is an upper bound of $\log N_I(\varepsilon, \mathfrak{A}, d_L)$ and f is integrable. Thus without loss of generality we may assume that

$$(5) \quad H \text{ is continuous, decreases, and } H(\varepsilon) \geq 1 + \log(\varepsilon^{-1})$$

and f is a decreasing function. Hence its inverse function f^{inv} is well defined. Put $\delta_{nk_n} = f^{\text{inv}}((n^d \mu_{k_n})^{1/2}/4)$. For $0 \leq k < k_n$ we choose δ_{nk} such that

$$(6) \quad n^d \mu_k = 16H(\delta_{nk})g^2(\delta_{nk_n})\delta_{nk}^{-1}.$$

This can be done, since H has the inverse function. Taking into account (4) we see that $n^d \mu_{k_n} \geq M\delta_{nk_n}^{-1} \rightarrow \infty$, whence $\delta_{nk_n} \rightarrow 0$ by (6). Note that the aim of the second truncation above, that is, the subtraction of the process Z''_n , is to satisfy the latter relation.

Step 3. Now we consider a finite net of subsets that approximate ν_{nk} on $A \in \mathbb{S}_\delta$ by its values on the subsets of the net that are close to A . The family $\mathfrak{A}$ is totally bounded, thus there are finite nets $\mathbb{D}_{nk}^{(l)}$, $l = 1, 2$, whose cardinalities are less than or equal to $\exp(2H(\delta_{nk}))$, respectively, and such that for all $A \in \mathbb{S}_\delta$ there is $D_{nk}^{(l)}(A) \in \mathbb{D}_{nk}^{(l)}$ for which $D_{nk}^{(1)}(A) \subset A \subset D_{nk}^{(2)}(A)$ and $|D_{nk}^{(2)}(A) \setminus D_{nk}^{(1)}(A)| \leq 2\delta_{nk}$.

Since ν_{nk} is additive, we represent Z_n^T as follows:

$$(7) \quad Z_n^T(A) = \sum_{k \leq k_n} \theta_k \nu_{nk} \left(D_{nk}^{(1)}(A) \right) + \sum_{k \leq k_n} \theta_k \nu_{nk} \left(\left(A \setminus D_{nk}^{(1)}(A) \right) \right) = Z_n^{(1)}(A) + Z_n^{(2)}(A).$$

Step 4. Further we apply the reconstruction technique. Let

$$L_n = \{j/n : j \in \{1, 2, \dots, n\}^d\}$$

and $C_{n,j} = (\mathbf{j} - n^{-1}\mathbf{1}, \mathbf{j}]$. Let $[0, 1]^d$ be represented as follows: $[0, 1]^d = \bigcup_{l \in L_{p_n}} C_{p_n, l}$ where $p_n = \lfloor n/m_n \rfloor$, $m_n^d = N$, and $N = g(\delta_{nk_n})$. The intersections of an arbitrary cube $C_{p_n, l}$ with cubes of the family $\{C_{n,j}, j \in L_n\}$ form a family of smaller cubes. These small cubes are numbered in every cube of the family $\{C_{p_n, l}, l \in L_{p_n}\}$ and denoted by I_{nli} , $i = \{1, \dots, N\}$. Put

$$I_{ni} = \bigcup_{l \in L_{p_n}} I_{nli}.$$

Then every element I_{ni} is a union of cubes whose sides are of length $1/n$ and the distances between the cubes are at least $1/p_n - 1/n$.

The reasoning above implies that

$$\nu_{nk}(A) = \sum_{i=1}^N \nu_{nk}(A \cap I_{ni}) = \sum_{i=1}^N \sum_{l \in L_{p_n}} \nu_{nk}(A \cap I_{nli})$$

and

$$\nu_{nk}(A \cap I_{nli}) = (n^d \mu_k)^{-1/2} \sum_{j \in nS(n, l, i)} |n(A \cap I_{nli} \cap C_{n,j})| a_{k+1}^{-1} X_j \mathbb{1}_{\{Y_j \in J_k\}}$$

where

$$S(n, l, i) = \{j \in L_n : C_{n,j} \cap I_{nli} \neq \emptyset\}.$$

If n and i are fixed, then the set $S(n, l, i)$ contains only a single element and the distance between these sets is at least $1/p_n - 1/n$.

To every $j \in nS(n, l, i)$ there corresponds a triple (n, l, i) . Then

$$Y_{nli} = Y_j$$

for $j \in nS(n, l, i)$ and $\nu_{nk}(A \cap I_{nli}) = (n^d \mu_k)^{-1/2} |n(A \cap I_{nli})| a_{k+1}^{-1} X_{nli} \mathbb{1}_{\{Y_{nli} \in J_k\}}$.

The following result plays the key role in the reconstruction technique.

Lemma 1 ([5]). *Let X and Y be two random variables assuming values in Polish spaces S_1 and S_2 , respectively. Suppose that the probability space where X and Y are defined is essentially rich in the sense that there exists a random variable U that is uniform on the interval $[0, 1]$ and independent of both X and Y . Then there exists a random variable Y^* having the same distribution as Y , independent of X , and such that*

$$\mathbb{P}(Y \neq Y^*) = \beta(\sigma(X), \sigma(Y)).$$

Moreover the random variable Y^* can be represented as $Y^* = f(X, Y, U)$ where

$$f: S_1 \times S_2 \times [0, 1] \rightarrow S_2$$

is some measurable function.

Let ψ be a one-to-one mapping from $[1, \sharp(L_{p_n})] \cap \mathbb{N}$ to L_{p_n} such that $\psi(m) <_{\text{lex}} \psi(m')$ for all $1 \leq m < m' \leq \sharp(L_{p_n})$ where the symbol $<_{\text{lex}}$ stands for the lexicographic order. Then

$$\{X_{nli}, l \in L_{p_n}\} = \{X_{n, \psi(m), i}, m \in [1, \sharp(L_{p_n})] \cap \mathbb{N}\}.$$

Given numbers n and i we use an induction and construct the independent random variables

$$\{\tilde{X}_{nli}, l \in L_{p_n}\}$$

such that the distribution of $\tilde{X}_{nli}$ coincides with that of X_{nli} and

$$\mathbb{P}(\tilde{X}_{nli} \neq X_{nli}) \leq \beta_X(m_n - 1, 1, p_n^d).$$

At the first step of the induction we put $\tilde{X}_{n,\psi(1),i} = X_{n,\psi(1),i}$, while at the induction step r ($1 < r < \sharp(L_{p_n})$) we apply Lemma 1 with

$$X = \left(\tilde{X}_{n,\psi(m),i} \right)_{1 \leq m < r}, \quad Y = X_{n,\psi(r),i} \quad \tilde{X}_{n,\psi(r),i} = Y^*.$$

As a result we get the inequalities

$$\mathbf{P}(\tilde{X}_{nli} \neq X_{nli}) = \beta(\sigma\{X\}, \sigma\{Y\}) \leq \beta(\sigma\{X_{n,\psi(m),i}, 1 \leq m < r\}, \sigma\{X_{n,\psi(r),i}\}).$$

Put $F_{ni} = \bigcup_{l \in L_{p_n}} \{\tilde{X}_{nli} \neq X_{nli}\}$. Then

$$(8) \quad \mathbf{P}(F_{ni}) \leq p_n^d \cdot \beta_X(m_n - 1, 1, p_n^d).$$

This completes the process of reconstruction of the field X .

Step 5. First we prove relation (3) for the process $Z_n^{(2)}(A)$. We represent $Z_n^{(2)}$ as follows:

$$Z_n^{(2)}(A) = \sum_{i=1}^N \sum_{k \leq k_n} \theta_k \nu_{nk} \left((A \setminus D_{nk}^{(1)}(A)) \cap I_{ni} \right).$$

If $U_{nli}^k(A) = n^{-d/2} |n(A \cap I_{nli})| \mathbb{1}_{\{Y_{nli} \in J_k\}}$, then

$$|\nu_{nk}(A)| \leq \sum_{i=1}^N (\mu_k)^{-1/2} \sum_{l \in L_{p_n}} U_{nli}^k(A).$$

Since $\mathbf{P}(Y_j \in J_k) \leq \mu_k$, we obtain

$$(\mu_k)^{-1/2} \sum_{l \in L_{p_n}} \mathbf{E}(U_{nli}^k(A)) \leq (n^d \mu_k)^{1/2} \sum_{l \in L_{p_n}} |A \cap I_{nli}|$$

where A is an arbitrary set belonging to $\mathfrak{B}([0, 1]^d)$. This implies for $A \in \mathbb{S}_\delta$ the following estimate:

$$\begin{aligned} \left| \nu_{nk} \left((A \setminus D_{nk}^{(1)}(A)) \cap I_{ni} \right) \right| &\leq \sum_{i=1}^N (\mu_k)^{-1/2} \sum_{l \in L_{p_n}} U_{nli}^k \left(D_{nk}^{(2)}(A) \setminus D_{nk}^{(1)}(A) \right) \\ &\leq \sum_{i=1}^N (\mu_k)^{-1/2} \sum_{l \in L_{p_n}} V_{nli}^k \left(D_{nk}^{(2)}(A) \setminus D_{nk}^{(1)}(A) \right) + (n^d \mu_k)^{1/2} \left| D_{nk}^{(2)}(A) \setminus D_{nk}^{(1)}(A) \right| \end{aligned}$$

where $V_{nli}^k = U_{nli}^k - \mathbf{E} U_{nli}^k$. Let $\tilde{V}_{nli}^k = \tilde{U}_{nli}^k - \mathbf{E} \tilde{U}_{nli}^k$ where

$$\tilde{U}_{nli}^k(A) = n^{-d/2} |n(A \cap I_{nli})| \mathbb{1}_{\{\tilde{X}_{nli} \in J_k\}}$$

and $\lambda_{nk} = 16N(\delta_{nk} H(\delta_{nk}))^{1/2}$. Then we estimate $\mathbf{P}(\|Z_n^{(2)}\|_{\mathbb{S}_\delta} > M)$ by

$$\begin{aligned} (9) \quad &\mathbf{P} \left(\sum_{i=1}^N \sum_{k \leq k_n} \theta_k \sup_{A \in \mathbb{S}_\delta} \left| \nu_{nk} \left((A \setminus D_{nk}^{(1)}(A)) \cap I_{ni} \right) \right| > \sum_{k \leq k_n} \theta_k \lambda_{nk} \right) \\ &\leq \sum_{i=1}^N \mathbf{P}(F_{ni}) + \sum_{i=1}^N \sum_{k \leq k_n} R_{ni}^k \end{aligned}$$

where

$$\begin{aligned} R_{ni}^k &= \mathbf{P} \left(\sup_{A \in \mathbb{S}_\delta} \left| \sum_{l \in L_{p_n}} \tilde{V}_{nli}^k \left(D_{nk}^{(2)}(A) \setminus D_{nk}^{(1)}(A) \right) \right| > p_{nk} N \right), \\ p_{nk} N &= (\mu_k)^{1/2} \lambda_{nk} / N - n^{d/2} \mu_k 2\delta_{nk} / N. \end{aligned}$$

Step 6. Now we are going to apply the Bernstein inequality (see [4]) to estimate R_{ni}^k . This can be done, since $\tilde{V}_{nli}^k(E)$ for $E \in \mathbb{S}_\delta$ are independent random variables. We need the following estimates:

$$(10) \quad \left| \tilde{V}_{nli}^k(E) \right| \leq n^{-d/2} |n(E \cap I_{nli})| \leq n^{-d/2},$$

$$(11) \quad \text{Var} \left(\sum_{l \in L_{pn}} \tilde{V}_{nli}^k(E) \right) \leq \sum_{l \in L_{pn}} \text{Var}(\mathbb{1}_{\{Y_0 \in J_k\}}) |E \cap I_{nli}| \leq |E \cap I_{ni}| \mu_k.$$

Putting $E = D_{nk}^{(2)}(A) \setminus D_{nk}^{(1)}(A)$ in (10) and (11) and applying the Bernstein inequality we obtain

$$\begin{aligned} R_{ni}^k &\leq \left(\# \left(\mathbb{D}_{nk}^{(1)} \right) \right)^2 \max_{A \in \mathbb{S}_\delta} \mathbb{P} \left(\left| \sum_{l \in L_{pn}} \tilde{V}_{nli}^k \left(D_{nk}^{(2)}(A) \setminus D_{nk}^{(1)}(A) \right) \right| > p_{nkN} \right) \\ &\leq \exp \left(4H(\delta_{nk}) - \frac{p_{nkN}^2}{2(\mu_k 2\delta_{nk} + 1/3n^{-d/2} p_{nkN})} \right). \end{aligned}$$

In view of (6) we have $n^{d/2} \mu_k \delta_{nk} = 4N(H(\delta_{nk})\delta_{nk})^{1/2}(\mu_k)^{1/2}$. Thus

$$R_{ni}^k \leq \exp \left(4H(\delta_{nk}) - \frac{64\mu_k H(\delta_{nk})\delta_{nk}}{2(\mu_k 2\delta_{nk} + 2/3\delta_{nk}\mu_k)} \right) = \exp(-8H(\delta_{nk})).$$

Then it follows from (5) that

$$(12) \quad \sum_{i=1}^N R_{ni}^k \leq N \exp(-8H(\delta_{nk})) \leq N\delta_{nk} \leq N(\delta_{nk}H(\delta_{nk}))^{1/2}.$$

Condition (2) implies that $\int_0^{\delta_{nk_n}} f(\varepsilon) d\varepsilon \rightarrow 0$ as $n \rightarrow \infty$. In turn, the latter relation implies that

$$(13) \quad \sum_{k \leq k_n} N(\delta_{nk}H(\delta_{nk}))^{1/2} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Indeed, put

$$q_{nk} = (n^d \mu_k)^{1/2} / 4 = N(H(\delta_{nk})/\delta_{nk})^{1/2}$$

and note that $q_{nk} \leq f(\delta_{nk})$ and $q_{n,k+1} = \beta^{1/2} q_{nk}$. Hence

$$\begin{aligned} \sum_{k \leq k_n} N(\delta_{nk}H(\delta_{nk}))^{1/2} &\leq q_{nk_n} f^{-1}(q_{nk_n}) + \left(1 - \beta^{1/2}\right)^{-1} \sum_{k < k_n} (q_{nk} - q_{n,k+1}) f^{-1}(q_{nk}) \\ &\leq q_{nk_n} f^{-1}(q_{nk_n}) + \left(1 - \beta^{1/2}\right)^{-1} \int_{q_{nk_n}}^{\infty} f^{-1}(x) dx \\ &\leq \left(1 - \beta^{1/2}\right)^{-1} \int_0^{\delta_{nk_n}} f(\varepsilon) d\varepsilon \rightarrow 0, \quad n \rightarrow \infty, \end{aligned}$$

whence $\sum_{k \leq k_n} \theta_k g(\delta_{nk_n})(\delta_{nk}H(\delta_{nk}))^{1/2} \rightarrow 0$ as $n \rightarrow \infty$. Therefore

$$(14) \quad \sum_{k \leq k_n} \sum_{i=1}^N R_{ni}^k \rightarrow 0 \quad \text{and} \quad \sum_{k \leq k_n} \theta_k \lambda_{nk} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Step 7. Now we prove the relation

$$(15) \quad \sum_{i=1}^N \mathbb{P}(F_{ni}) \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

which, in view of (8), follows from $n^d \beta_X(m_n - 1, 1, p_n^d) \rightarrow 0$ as $n \rightarrow \infty$. The latter relation is equivalent to

$$(16) \quad n^d \beta_X \left(g^{1/d}(\delta_{nk_n}) - 1, 1, n^d/g(\delta_{nk_n}) \right) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

According to (6)

$$n^d = \left(\frac{n^d}{H(\delta_{nk_n})N^2} \right)^\tau \cdot \frac{H^\tau(\delta_{nk_n})N^{2\tau}}{n^{d(\tau-1)}} = \frac{16^\tau}{(\mu_{k_n})^\tau n^{d(\tau-1)}} \cdot f^{2\tau}(\delta_{nk_n})$$

where $\tau = s/(s-2)$ is defined in the statement of Theorem 1. The construction of μ_{k_n} and (4) imply that

$$\lim_{n \rightarrow \infty} (\mu_{k_n})^{-\tau} n^{-d(\tau-1)} \leq \lim_{n \rightarrow \infty} n^d b_n^\tau / (M\delta)^\tau = 0.$$

It is clear that the latter relation holds for b_n specified above. Then the relation

$$\lim_{n \rightarrow \infty} f^{2\tau}(\delta_{nk_n}) \cdot \beta_X \left(g^{1/d}(\delta_{nk_n}) - 1, 1, f^{2\tau}(\delta_{nk_n})/g(\delta_{nk_n}) \right) < \infty$$

implies (16).

Taking (9), (14), and (15) into account, we see that $\mathbf{P}(\|Z_n^{(2)}\|_{\mathbb{S}_\delta} > M) \rightarrow 0$ as $n \rightarrow \infty$.

Step 8. Now it remains to consider the processes $Z_n^{(1)}$ defined by (7). When considering these processes we may face a problem that some of the approximating sets $D_{nk}^{(1)}(A)$ are too close together, and this may not allow us to obtain a suitable Gaussian approximation. To avoid this problem we apply the following idea. Let $\mathbb{S}_{nk}$ be the maximal subset of $\mathbb{S}_\delta$ such that $|C_1 \triangle C_2| \geq 2\delta_{nk}$ for all $C_1 \neq C_2$. Then for any $A \in \mathbb{S}_\delta$ there is an element $C_{nk}(A)$ of $\mathbb{S}_{nk}$ such that $|C_{nk}(A) \triangle A| < 2\delta_{nk}$. Thus

$$(17) \quad \left| C_{nk}(A) \triangle D_{nk}^{(1)}(A) \right| < 4\delta_{nk}.$$

Put

$$Z_n^{(3)}(A) = \sum_{k \leq k_n} \theta_k \nu_{nk}(C_{nk}(A)), \quad Z_n^{(4)}(A) = \sum_{k \leq k_n} \theta_k \left\{ \nu_{nk} \left(D_{nk}^{(1)}(A) \right) - \nu_{nk}(C_{nk}(A)) \right\}.$$

Then $Z_n^{(1)} = Z_n^{(3)}(A) + Z_n^{(4)}(A)$.

Proceeding in the same way as in the case of the process $Z_n^{(2)}$ and substituting $D_{nk}^{(1)}(A) \triangle C_{nk}(A)$ instead of E in (10) and in (11) we obtain from (17) that

$$\mathbf{P} \left(\|Z_n^{(4)}\|_{\mathbb{S}_\delta} > M \right) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Step 9. Recall that if a field $(X_j)_j$ is symmetric, then its distribution coincides with that of the field $(\varepsilon_j X_j)_j$ where $(\varepsilon_j)_j$ is a Rademacher field that does not depend on X . Without loss of generality one can thus consider the field $\tilde{Z}_n^{(3)}(A) = \sum_j \varepsilon_j S_{nj}(A)$ instead of $Z_n^{(3)}(A)$ where $S_{nj}(A) = \sum_{k \leq k_n} n^{-d/2} b_{nk}(C_{nk}(A)) X_j \mathbb{1}_{\{Y_j \in J_k\}}$. We have

$$(18) \quad \mathbf{P} \left(\left\| Z_n^{(3)} \right\|_{\mathbb{S}_\delta} > M \right) = \mathbf{P} \left(\left\| \tilde{Z}_n^{(3)} \right\|_{\mathbb{S}_\delta} > M \right) \leq 1/M \mathbf{E} \left\| \tilde{Z}_n^{(3)} \right\|_{\mathbb{S}_\delta}.$$

Step 10. Now we need

Lemma 2 ([3]). *Let $\{f_j, j \in T\}$ be a finite set of real functions defined on the set D and let $\{v_j, j \in T\}$ be a family of nonnegative random variables such that $\mathbf{E}(v_j) = 1$. Further*

let $\{\varepsilon_j, j \in T\}$ be a family of random variables that do not depend on $\{v_j, j \in T\}$. Then

$$\mathbb{E} \left\| \sum_{j \in T} \varepsilon_j f_j \right\|_D \leq \mathbb{E} \left\| \sum_{j \in T} \varepsilon_j v_j f_j \right\|_D$$

provided the norms are measurable.

Let $\{g_j, j \in \mathbf{Z}^d\}$ be independent standard Gaussian random variables that do not depend on $\{\varepsilon_j, j \in \mathbf{Z}^d\}$. Since the fields X , ε , and $g = \{g_j, j \in \mathbf{Z}^d\}$ are independent, we may assume that these fields are defined on different probability spaces with measures $\mathbb{P}_X$, $\mathbb{P}_\varepsilon$, and $\mathbb{P}_g$, respectively. Denote by $\mathbb{E}_X$, $\mathbb{E}_\varepsilon$, and $\mathbb{E}_g$ the partial integration with respect to $\mathbb{P}_X$, $\mathbb{P}_\varepsilon$, and $\mathbb{P}_g$, respectively. In particular, $\mathbb{E}(\cdot) = \mathbb{E}_X \mathbb{E}_\varepsilon \mathbb{E}_g(\cdot)$. Put $\mu = 1/\mathbb{E}|g_0|$. Since $\varepsilon_j|g_j| = g_j$ in law, we set $v_j = \mu|g_j|$ in Lemma 2 and obtain

$$(19) \quad \mathbb{E}_\varepsilon \left\| \widetilde{Z}_n^{(3)} \right\|_{\mathbb{S}_\delta} \leq \mu \mathbb{E}_g \|Z_n^g\|_{\mathbb{S}_\delta}$$

where $Z_n^g(A) = \sum_j g_j S_{nj}(A)$. Note that all the norms are measurable, since the upper bounds are considered with respect to the finite set $\bigcup_{k \leq k_n} \mathbb{S}_{nk}$.

The process Z_n^g is Gaussian with respect to the measure $\mathbb{P}_g$. We compare the Gaussian process Z_n^g with another Gaussian process constructed from the Brownian motion Z . Namely let $G^{(k)}, k \geq 1$, be independent copies of the process Z . Let

$$(20) \quad G_n(A) = \sum_{k \leq k_n} 2\theta_k G^{(k)}(C_{nk}(A)),$$

$$Q_{nk}(A) = (n^d \mu_k)^{-1} \sum_j b_{nj}(A) \mathbb{1}_{\{Y_j \in J_k\}}, \quad \Delta_{nk} = \{E \triangle F : E \neq F \in \mathbb{S}_{nk}\}.$$

Then for $A, B \in \mathbb{S}_\delta$ we have $\mathbb{E}_g (Z_n^g(A) - Z_n^g(B))^2 = \sum_j (S_{nj}(A) - S_{nj}(B))^2$ and this does not exceed

$$\begin{aligned} & n^{-d} \sum_{k \leq k_n} \sum_j (b_{nj}(C_{nk}(A)) - b_{nj}(C_{nk}(B)))^2 a_{k+1}^2 \mathbb{1}_{\{Y_j \in J_k\}} \\ & \leq \sum_{k \leq k_n} \theta_k^2 Q_{nk}(C_{nk}(A) \triangle C_{nk}(B)). \end{aligned}$$

On the other hand

$$\mathbb{E}(G_n(A) - G_n(B))^2 = \sum_{k \leq k_n} 4\theta_k^2 |C_{nk}(A) \triangle C_{nk}(B)|.$$

Therefore

$$\mathbb{E}_g (Z_n^g(A) - Z_n^g(B))^2 \leq \mathbb{E}(G_n(A) - G_n(B))^2$$

on the event $D_n = \{Q_{nk}(A) \leq 4|A| \text{ for all } A \in \Delta_{nk} \text{ and } k \leq k_n\}$.

Step 11. Below we need the following result.

Lemma 3 ([10]). *Let $\{Y_i(t), t \in D\}$, $i = 1, 2$, be centered Gaussian processes indexed by a countable set D such that*

$$0 \in \{Y_1(t, \omega) : t \in D, \omega \in \Omega\} \quad \text{almost surely.}$$

Assume that

$$\mathbb{E}(Y_1(t) - Y_1(s))^2 \leq \mathbb{E}(Y_2(t) - Y_2(s))^2$$

for all $s, t \in D$. Then $\mathbb{E}\|Y_1\|_D \leq 2\mathbb{E}\|Y_2\|_D$.

The condition $0 \in \{Y_1(t, \omega) : t \in D, \omega \in \Omega\}$ almost surely holds if D is a class of sets containing the empty set. Note that this is the case for our consideration (recall that $\emptyset \in \mathbb{S}_{nk}$).

By Lemma 3 we have

$$(21) \quad \mathbf{E}_g \|Z_n^g\|_{\mathbb{S}_\delta} \leq 2 \mathbf{E} \|G_n\|_{\mathbb{S}_\delta}$$

on the event D_n . It is convenient to introduce the following notation: $Z_{nk}(A) = \sum_{l=k}^{k_n} 2\theta_l G^{(l)}(A)$ and

$$W_{nk}(A) = Z_{nk}(C_{nk}(A)) - Z_{nk}(C_{n,k-1}(A)), \quad v = 2 \left(\sum_{l \geq 0} \theta_l^2 \right)^{1/2}.$$

Changing the order of summation in (20) we get

$$(22) \quad \begin{aligned} G_n(A) &= \sum_{k \leq k_n} 2\theta_k G^{(k)}(C_{n0}(A)) + \sum_{k \leq k_n} \sum_{l=1}^k 2\theta_k \left\{ G^{(k)}(C_{nl}(A)) - G^{(k)}(C_{n,l-1}(A)) \right\} \\ &= Z_{n0}(C_{n0}(A)) + \sum_{k=1}^{k_n} W_{nk}(A). \end{aligned}$$

Lemma 4 ([9]). *There exists an universal constant K such that*

$$\mathbf{E} \|Z\|_{\mathbb{S}_\delta} \leq K \int_0^\delta (\varepsilon^{-1} \log N_I(\varepsilon, \mathbb{S}_\delta, d_L))^{1/2} d\varepsilon + K\delta.$$

Then $\mathbf{E} \|Z_{n0}\|_{\mathbb{S}_\delta} \leq v \mathbf{E} \|Z\|_{\mathbb{S}_\delta} \rightarrow 0$ as $\delta \rightarrow 0$. We see that $W_{nk}(A)$ is a Gaussian random variable, and (17) implies that $\mathbf{E} W_{nk}^2(A) \leq 4v^2\delta_{nk}$. Thus

$$\begin{aligned} \mathbf{P} \left(\|W_{nk}\|_{\mathbb{S}_\delta} > t (4v^2\delta_{nk})^{1/2} \right) &\leq \sharp(\mathbb{S}_{nk}) \sharp(\mathbb{S}_{n,k-1}) (2/\sqrt{2\pi}) t^{-1} \exp(-t^2/2) \\ &\leq \exp(-(t^2 - 8H(\delta_{n,k-1}))/2) \\ &\leq \exp(-(t^2 - 8\beta^{-1}H(\delta_{nk}))/2) \leq \exp(-t^2/4) \end{aligned}$$

for all $t \geq (16\beta^{-1}H(\delta_{nk}))^{1/2}$ if n is sufficiently large. Hence

$$\mathbf{E} \|W_{nk}\|_{\mathbb{S}_\delta} \leq (4v^2\delta_{nk})^{1/2} \left((16\beta^{-1}H(\delta_{nk}))^{1/2} + 4 \right).$$

Metric entropy condition (2) implies that $\sum_{1 \leq k \leq k_n} \mathbf{E} \|W_{nk}\|_{\mathbb{S}_\delta} \rightarrow 0$ as $n \rightarrow \infty$. Combining (18)–(22) we see that

$$\mathbf{P} \left(\|Z_n^{(3)}\|_{\mathbb{S}_\delta} > M \right) \leq \mathbf{P}(D_n^c) + o(1).$$

Step 12. To complete the proof of Theorem 1 it is sufficient to show that

$$(23) \quad \mathbf{P}(D_n^c) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

As before

$$Q_{nk}(A) = n^{-d/2} \mu_k^{-1} \sum_{i=1}^N \sum_{l \in L_{p_n}} V_{nli}^k(A) + |A|.$$

Taking the above reasoning into account we obtain from the definition of the event D_n^c that

$$(24) \quad \begin{aligned} \mathbf{P}(D_n^c) &= \mathbf{P}(\text{there exist } k \leq k_n \text{ and } A \in \Delta_{nk} : Q_{nk}(A) > 4|A|) \\ &\leq \mathbf{P} \left(\bigcup_{i=1}^N F_{ni} \right) + \sum_{i=1}^N \sum_{k \leq k_n} \hat{\mathbf{R}}_{ni}^k \end{aligned}$$

where

$$\hat{R}_{ni}^k = \sharp(\Delta_{nk}) \sup_{A \in \Delta_{nk}} \mathbb{P} \left(n^{-d/2} (\mu_k)^{-1} \left| \sum_{l \in L_{pn}} \tilde{V}_{nli}^k(A) \right| > 3|A|/N \right).$$

As we have already shown before, conditions (i) and (ii) imply that the first term in (24) approaches 0 as $n \rightarrow \infty$. We estimate $\hat{R}_{ni}^k$ in the second term by

$$\hat{R}_{ni}^k \leq \left(\sharp(\mathbb{D}_{nk}^{(1)}) \right)^2 \sup_{A \in \Delta_{nk}} \mathbb{P} \left(\left| \sum_{l \in L_{pn}} \tilde{V}_{nli}^k(A) \right| > 3|A|/N n^{d/2} \mu_k \right).$$

Applying the Bernstein inequality once more and the estimate $|A| \geq 2\delta_{nk}$ for all $A \in \Delta_{nk}$ we get

$$\begin{aligned} \hat{R}_{ni}^k &\leq \sup_{A \in \Delta_{nk}} \exp \left(4H(\delta_{nk}) - \frac{9|A|^2/N^2 n^d \mu_k^2}{2(\mu_k|A| + 1/3 n^{-d/2} 3|A|/N n^{d/2} \mu_k)} \right) \\ &\leq \exp(-68H(\delta_{nk})). \end{aligned}$$

The latter inequality holds, since (6) implies that $\mu_k n^d |A|/N \geq 32H(\delta_{nk})N$. Taking (12) and (13) into account, we get

$$(25) \quad \sum_{i=1}^N \sum_{k \leq k_n} \hat{R}_{ni}^k \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Therefore (24) and (25) imply (23).

Theorem 1 is proved. $\square$

4. WEAK CONVERGENCE

We use a result of the paper [7] to prove the convergence of finite-dimensional distributions of the processes Z_n .

We call a Borel set A regular if the Lebesgue measure of its boundary is zero.

Theorem 2. *Let $\mathfrak{A}$ be a family of regular sets of $[0, 1]^d$ satisfying metric entropy condition (2). Let X be a symmetric strictly stationary field and let assumption (i) of Theorem 1 be satisfied. If*

$$(ii') \quad \limsup_{\varepsilon \downarrow 0} f^{2\tau}(\varepsilon) \cdot \beta_X(g^{1/d}(\varepsilon), 1, \infty) < \infty,$$

then the invariance principle holds.

Proof. The tightness of the family of distributions of $\{Z_n\}_{n \in \mathbf{N}}$ follows from Theorem 1. To prove the convergence of finite-dimensional distributions of the fields $\{Z_n(B), B \in \mathfrak{A}\}$ to those of the field $\{Z(B), B \in \mathfrak{A}\}$ it is sufficient (see [7]) to check that

$$(26) \quad \sum_{n \in \mathbf{N}} n^{d-1} \alpha_X^{(s-2)/s}(n, 1, \infty) < \infty.$$

The latter condition follows from (ii'). $\square$

Acknowledgement. The author is grateful to Professor A. V. Bulinskii for the setting of the problem and help in writing this paper.

BIBLIOGRAPHY

1. A. V. Bulinskiĭ, *Limit Theorems under Weak Dependence Conditions*, Moscow University, Moscow, 1989. (Russian)
2. A. V. Bulinskiĭ and E. Shabanovich, *Asymptotic behavior of some functionals of positively or negatively dependent random fields*, Fundam. Prikl. Mat. **4** (1998), no. 2, 479–492. (Russian) MR1801168 (2001h:60033)
3. K. Alexander and R. Pyke, *A uniform central limit theorem for set-indexed partial-sum processes with finite variance*, Ann. Probab. **14** (1986), no. 2, 582–597. MR0832025 (88b:60084)
4. G. Bennett, *Probability inequalities for the sum of independent random variables*, J. Amer. Statist. Assoc. **57** (1962), 33–45.
5. H. C. P. Berbee, *Random Walks with Stationary Increments and Renewal Theory*, Mathematical Centre Tracts, vol. 112, Amsterdam, 1979. MR0547109 (81e:60093)
6. D. Chen, *A uniform central limit theorem for nonuniform φ -mixing random fields*, Ann. Probab. **19** (1991), no. 2, 636–649. MR1106280 (93a:60050)
7. J. Dedecker, *Exponential inequalities and functional central limit theorems for random fields*, ESAIM, Probabilités et Statistique **5** (2001), 77–104. MR1875665 (2003a:60054)
8. P. Doukhan, *Mixing. Properties and Examples*, Lecture Notes in Statistics, vol. 85, Springer, New York, 1994. MR1312160 (96b:60090)
9. R. M. Dudley, *Sample functions of the Gaussian process*, Ann. Probab. **1** (1973), 66–103. MR0346884 (49:11605)
10. R. V. Fernique, *Régularité des trajectoires des fonctions aléatoires gaussiennes*, Lecture Notes in Math. **480**, Springer, Berlin, 1975, pp. 1–96. MR0413238 (54:1355)
11. C. M. Goldie and P. Greenwood, *Variance of set-indexed sums of mixing random variables and weak convergence of set-indexed processes*, Ann. Probab. **14** (1985), 815–839. MR0841586 (88e:60038b)
12. M. Ledoux and M. Talagrand, *Probability in Banach Spaces*, Springer, Berlin, 1991. MR1102015 (93c:60001)

DEPARTMENT OF PROBABILITY THEORY, MECHANICS AND MATHEMATICS FACULTY, MOSCOW STATE UNIVERSITY, MOSCOW, RUSSIA

E-mail address: denis@orc.ru

Received 27/FEB/2003

Translated by V. SEMENOV