Bernoulli **18**(1), 2012, 100–118
DOI: 10.3150/10-BEJ339

Central limit theorems for the excursion set volumes of weakly dependent random fields

ALEXANDER BULINSKI[1], EVGENY SPODAREV[2,*] and FLORIAN TIMMERMANN[2,**]

[1]*Moscow State University, Department of Mathematics and Mechanics, 119991 Moscow, Russia. E-mail: bulinski@mech.math.msu.su*
[2]*Ulm University, Institute of Stochastics, 89069 Ulm, Germany. E-mail:* [*]*evgeny.spodarev@uni-ulm.de;* [**]*florian.timmermann@uni-ulm.de*

The multivariate central limit theorems (CLT) for the volumes of excursion sets of stationary quasi-associated random fields on $\mathbb{R}^d$ are proved. Special attention is paid to Gaussian and shot noise fields. Formulae for the covariance matrix of the limiting distribution are provided. A statistical version of the CLT is considered as well. Some numerical results are also discussed.

Keywords: central limit theorem; dependence conditions; excursion sets; random fields

1. Introduction

An important research domain of modern probability theory is the investigation of geometric characteristics of random surfaces (see, e.g., [1–3]). The origin of interest often roots not only in pure mathematical challenges but also in various applications, including those in industry. We mention one motivating example for our study.

The contemporary method of papermaking goes back to the Han Dynasty period. Nowadays, the method is essentially the same, but machines in modern pulp and paper mills operate much faster. The surface structure of the paper during the forming process determines the quality of the production.

To model the paper surface, stationary random fields, say, shot noise (cf. [4]) or Gaussian, can be a reasonable first choice. Comparing by eye real paper image data and simulated realizations of such fields, one easily concludes that the similarities are striking. But it is hard to quantify how different these two images really are. To test whether the available image data originate from a realization of a specified stationary random field, the excursion sets can be considered.

We prove the central limit theorem (CLT) for volumes of excursion sets of a stationary field $X = \{X(t), t \in T\}$, $T \subset \mathbb{R}^d$, to characterize the surface generated by X. It is reasonable to assume that the field X could possess a dependence structure more general than positive or negative association used in a number of stochastic models; see, for example, [7]. Our main results yield uni- and multivariate CLT for quasi-associated random fields. The CLT is generalized in [12], page 80, having been obtained by other methods for volumes of excursion sets of stationary and isotropic Gaussian random fields. We also discuss the consistent estimators for the asymptotic covariance matrix that arises in the limiting distribution.

Note that we do not tackle here the interesting problems concerning the study of moving levels for excursion sets, the estimate of the convergence rate to the limit law and the analysis of the functionals in Gaussian random fields based on the Dobrushin–Major techniques. In this regard, we refer, to [12,15–17].

As to the problem of characterizing the paper quality taking into account the "hills" and "valleys" of its surface discernible with the help of microscope, it is by no means simple. In fact, we have to specify the admissible (average) number of such hills along with their size. Moreover, the thickness of the paper should be controlled as well (no holes or high peaks). Thus the study of the excursion sets for random fields is the first natural step to investigate such random surfaces. The application to paper surface image data will appear in a separate paper.

The present paper is organized as follows. Section 2 provides preliminaries on dependence concepts related to association and excursion sets for random fields. The CLT for the volumes of excursions of quasi-associated stationary random fields over one or finitely many levels are formulated and proved in Section 3. The special cases of stationary shot noise and Gaussian random fields are treated in more detail. Section 4 contains a statistical version of the limit theorems mentioned above where the (unknown) limiting covariance matrix is consistently estimated. Numerical results illustrating the limit theorem of Section 3 are given in Section 5. Finally, we conclude with the discussion of some open problems.

2. Preliminaries

In this section, we recall some dependence concepts for systems of random variables. Various examples can be found in [7]. After that, we introduce the excursion sets that are the main objects of this study. Then we consider the sequences of regular growing sets forming observation windows.

2.1. Dependence concepts for random fields

Consider a family, $X = \{X(t), t \in T\}$, of real-valued random variables, $X(t)$, defined on a probability space, $(\Omega, \mathcal{F}, \mathsf{P})$. A set, T, will be a subset of $\mathbb{R}^d$ or $\mathbb{Z}^d$. For $I \subset T$ let $X_I = \{X(t), t \in I\}$. Introduce the class $\mathcal{M}(n)$ consisting of real-valued, bounded, coordinate-wise non-decreasing Borel functions on $\mathbb{R}^n$, $n \in \mathbb{N}$. The cardinality of a finite set, U, will be denoted by $\operatorname{card} U$.

Definition 1. *A real-valued random field $X = \{X(t), t \in T\}$ is called positively associated* (*we write $X \in \mathsf{PA}$*) *if, for every disjoint finite set $I, J \subset T$ and any functions $f \in \mathcal{M}(\operatorname{card} I)$ and $g \in \mathcal{M}(\operatorname{card} J)$, one has*

$$\operatorname{cov}(f(X_I), g(X_J)) \geq 0. \tag{1}$$

Here, we use any permutation of (coordinates of) the column vector $(X(t_1), \dots, X(t_n))^\top$ for X_I, $I = \{t_1, \dots, t_n\} \subset T$ (and the analogous notation is employed for X_J); $\top$ stands for transposition. Definition 1, given for any (not necessarily disjoint) subsets I and $J \subset T$, introduces the family of *associated* random variables ($X \in \mathsf{A}$). The change of the sign of inequality (1) leads to the definition of negative association (one writes $X \in \mathsf{NA}$). Clearly, $X \in \mathsf{A}$ implies $X \in \mathsf{PA}$. Any collection of independent random variables is automatically PA and NA. Due to Pitt [19], a Gaussian family $X = \{X(t), t \in T\}$ of random variables is associated if and only if $\mathsf{cov}(X(s), X(t)) \geq 0$ for all $s, t \in T$. For such families, the concepts of A and PA coincide. A theorem by Joag-Dev and Proschan [13] states that a Gaussian family $X = \{X(t), t \in T\} \in \mathsf{NA}$ if and only if $\mathsf{cov}(X(s), X(t)) \leq 0$ for $s, t \in T$, $s \neq t$.

Let $BL(n)$ denote the class of *bounded Lipschitz* functions $f : \mathbb{R}^n \to \mathbb{R}$ ($n \in \mathbb{N}$) and

$$\mathsf{Lip}(f) = \sup_{x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|_1} < \infty, \qquad \|x\|_1 = \sum_{k=1}^{n} |x_k|, \ x = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n.$$

Since all norms are equivalent in $\mathbb{R}^n$, we sometimes use the Euclidean norm $\|x\|_2 = (\sum_{k=1}^n x_k^2)^{1/2}$ and the supremum norm $\|x\|_\infty = \max_{k=1,\ldots,n} |x_k|$ of $x \in \mathbb{R}^n$ for the sake of convenience.

Definition 2. *A random field $X = \{X(t), t \in T\}$ consisting of random variables $X(t)$ with $\mathsf{E}X(t)^2 < \infty$ is called quasi-associated ($X \in \mathsf{QA}$) if*

$$|\mathsf{cov}(f(X_I), g(X_J))| \leq \mathsf{Lip}(f)\,\mathsf{Lip}(g) \sum_{s \in I} \sum_{t \in J} |\mathsf{cov}(X(s), X(t))| \tag{2}$$

for all disjoint finite sets I, $J \subset T$ and any Lipschitz functions

$$f : \mathbb{R}^{\mathsf{card}\, I} \to \mathbb{R} \quad \text{and} \quad g : \mathbb{R}^{\mathsf{card}\, J} \to \mathbb{R}.$$

If $X \in \mathsf{PA}$ or $X \in \mathsf{NA}$ and $\mathsf{E}X(t)^2 < \infty$ for all $t \in T$, then (2) holds as was proved in [9]. Every Gaussian random field X (with covariance function taking both positive and negative values) is quasi-associated; see [20] and references therein.

Definition 3. *A real-valued random field $X = \{X(t), t \in \mathbb{Z}^d\}$ is called (BL, θ)-dependent ($X \in (BL, \theta)$) if there exists a non-increasing sequence $\theta = (\theta_r)_{r \in \mathbb{N}}, \theta_r \to 0$ as $r \to \infty$, such that, for any finite disjoint sets I, $J \subset \mathbb{Z}^d$ with $\mathsf{dist}(I, J) = r$ and any functions $f \in BL(\mathsf{card}\, I)$, $g \in BL(\mathsf{card}\, J)$, one has*

$$|\mathsf{cov}(f(X_I), g(X_J))| \leq \mathsf{Lip}(f)\,\mathsf{Lip}(g)(\mathsf{card}\, I \wedge \mathsf{card}\, J)\theta_r, \tag{3}$$

where $\mathsf{dist}(I, J) = \min\{\|s - t\|_\infty \colon s \in I, t \in J\}$.

If $X = \{X(t), t \in \mathbb{Z}^d\} \in \mathsf{QA}$, then $X \in (BL, \theta)$ whenever the Cox–Grimmett coefficient

$$u_r := \sup_{s \in \mathbb{Z}^d} \sum_{t \colon \|s - t\|_\infty \geq r} |\mathsf{cov}(X_s, X_t)| \tag{4}$$

tends to zero as $r \to \infty$. In this case, one can take $\theta_r = u_r$ in (3).

For a random field $X = \{X(t), t \in \mathbb{R}^d\}$ we use (see [5]) the following extension of (3). Let $T(\Delta) = \{(j_1/\Delta, \ldots, j_d/\Delta) \colon (j_1, \ldots, j_d) \in \mathbb{Z}^d\}$, where $\Delta > 0$.

Definition 4. *A real-valued random field $X = \{X(t), t \in \mathbb{R}^d\}$ is called (BL, θ)-dependent if there exists a non-increasing function $\theta = (\theta_r)_{r>0}$, $\theta_r \to 0$ as $r \to \infty$, such that, for all Δ large enough and any finite disjoint sets I, $J \subset T(\Delta)$ with $\mathsf{dist}(I, J) = r$, and any functions $f \in BL(\mathsf{card}\, I)$, $g \in BL(\mathsf{card}\, J)$, one has*

$$|\mathsf{cov}(f(X_I), g(X_J))| \le \mathsf{Lip}(f)\,\mathsf{Lip}(g)(\mathsf{card}\, I \wedge \mathsf{card}\, J)\Delta^d \theta_r. \tag{5}$$

In many cases, one can use the integral analog of (4) for θ_r. Thus for a (wide-sense) stationary random field $X = \{X(t), t \in \mathbb{R}^d\} \in \mathsf{QA}$, having covariance function $R(t)$, $t \in \mathbb{R}^d$, absolutely directly integrable in the Riemann sense (i.e., when $d = 1$; see, e.g., Feller [11], page 362. For $d > 1$, the definition is quite similar. One takes the partition of $\mathbb{R}^d$ generated by partitions of each coordinate axis and forms the corresponding upper and lower Riemann sums.), relation (5) holds with

$$\theta_r = 2 \int_{\|x\|_\infty \ge r} |R(t)|\, \mathrm{d}t, \qquad r > 0; \tag{6}$$

see [5]. We shall also write $\theta(X) = \theta_r(X)$ to emphasize that θ in (3) or (5) refers to the field X.

2.2. Excursion sets

Now we recall the definition of an excursion set and illustrate it by Figure 1.

For a real-valued random field $X = \{X(t), t \in \mathbb{R}^d\}$, we assume the measurability of $X(\cdot)$ as a function on $\mathbb{R}^d \times \Omega$ endowed with the σ-algebra $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{F}$.

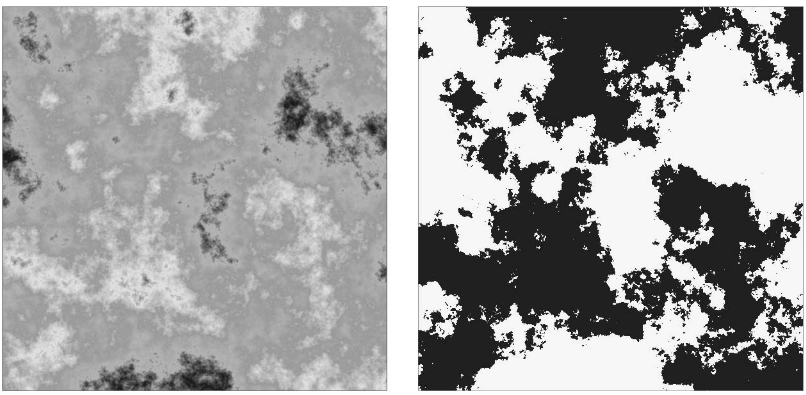

Figure 1. Realization of a stationary centered Gaussian random field X with covariance function $\mathsf{cov}(X(0), X(t)) = \exp(-\|t\|_2)$ (left figure), bright colours indicate high values of X. The excursion set $A_u(X, T)$ for $u = 0$ is shown in black (right figure).

Definition 5. *Let X be a measurable real-valued function on $\mathbb{R}^d$ and $T \subset \mathbb{R}^d$ be a (Lebesgue) measurable subset. Then, for each $u \in \mathbb{R}$,*

$$A_u(X, T) = \{t \in T\colon\ X(t) \geq u\}$$

is called the excursion set of X in T over the level u.

Let $\nu_d(B)$ be the volume (i.e., the Lebesgue measure) of a measurable set $B \subset \mathbb{R}^d$ and $\mathbb{I}\{C\}$ denote the indicator of a set C.

Since X is measurable, the volume of the excursion set

$$\nu_d(A_u(X, T)) = \int_T \mathbb{I}\{X(t) \geq u\}\,\mathrm{d}t$$

is a random variable for each $u \in \mathbb{R}$ and any measurable set $T \subset \mathbb{R}^d$.

2.3. Growing sets

Denote the boundary of a set $B \subset \mathbb{R}^d$ by ∂B. The *Minkowski sum* of two sets, $A, B \subset \mathbb{R}^d$, is given by $A \oplus B = \{x + y\colon\ x \in A, y \in B\}$. The following concept of "regular growth" for a family of subsets in $\mathbb{R}^d$ will be used in the sequel.

Definition 6. *A sequence, $(W_n)_{n\in\mathbb{N}}$, of bounded measurable sets, $W_n \subset \mathbb{R}^d$, tends to infinity in the Van Hove sense (VH-growing) if, for any $\varepsilon > 0$, one has*

$$\nu_d(W_n) \to \infty \quad \text{and} \quad \frac{\nu_d(\partial W_n \oplus B_\varepsilon(0))}{\nu_d(W_n)} \to 0 \tag{7}$$

as $n \to \infty$, where $B_\varepsilon(0) = \{x \in \mathbb{R}^d\colon\ \|x\|_2 \leq \varepsilon\}$ is the closed ball in $\mathbb{R}^d$ with center at the origin $0 \in \mathbb{R}^d$ and radius ε.

If $W_n = (a(n), b(n)] = (a_1(n), b_1(n)] \times \cdots \times (a_d(n), b_d(n)]$ is a parallelepiped, then $W_n \to \infty$ in the Van Hove sense if and only if $b_k(n) - a_k(n) \to \infty$ as $n \to \infty$ for $k = 1, \ldots, d$.

Definition 7. *A sequence of finite sets $U_n \subset \mathbb{Z}^d$ tends to infinity in a regular way if*

$$\frac{\mathsf{card}\,\delta U_n}{\mathsf{card}\,U_n} \to 0 \qquad \text{as } n \to \infty; \tag{8}$$

cf. (7). *Here $\delta U_n = \{j \in \mathbb{Z}^d \setminus U_n\colon\ \mathsf{dist}(j, U_n) = 1\}$ and $\mathsf{dist}(j, U_n) = \min_{k\in U_n} \|j - k\|_\infty$.*

3. Central limit theorem

Now we state and prove a CLT for the volume of excursion sets of random fields. Ivanov and Leonenko [12] studied stationary and isotropic Gaussian random fields. In our approach, the isotropy of Gaussian fields is not required. Moreover, we consider a more general class of random fields possessing the quasi-association property. To avoid long formulations, we introduce the following two conditions for a random field $X = \{X(t), t \in \mathbb{R}^d\}$.

(A) X is quasi-associated and strictly stationary such that $X(0)$ has a bounded density. Assume that the covariance function of X is continuous and there exists some $\alpha > 3d$ such that

$$|\operatorname{cov}(X(0), X(t))| = \mathrm{O}(\|t\|_2^{-\alpha}) \qquad \text{as } \|t\|_2 \to \infty. \tag{9}$$

(B) X is Gaussian and stationary. Suppose that its continuous covariance function satisfies (9) for some $\alpha > d$.

Notice that continuity of the covariance function of X implies the existence of a measurable modification of this field. We consider only such versions of X. We exclude the trivial case when $X(t) = \text{const}$ a.s. for all $t \in \mathbb{R}^d$.

3.1. Quasi-associated random fields

To prove the CLT for the volume of excursion sets of a random field satisfying condition (A), we need the following auxiliary result.

Lemma 1 ([7], Lemma 7.3.4). *Let $\{U, V\} \in \mathsf{QA}$, where random variables U and V are square-integrable and have densities bounded by $a > 0$. Then*

$$|\operatorname{cov}(\mathbb{I}\{U \geq u\}, \mathbb{I}\{V \geq v\})| \leq 3 \cdot 2^{2/3} a^{2/3} |\operatorname{cov}(U, V)|^{1/3}$$

for arbitrary $u, v \in \mathbb{R}$.

Theorem 1. *Let $X = \{X(t), t \in \mathbb{R}^d\}$ be a random field satisfying condition* (A). *Then, for any sequence of VH-growing sets $W_n \subset \mathbb{R}^d$ and each $u \in \mathbb{R}$, one has*

$$\frac{\nu_d(A_u(X, W_n)) - \nu_d(W_n)\mathsf{P}(X(0) \geq u)}{\sqrt{\nu_d(W_n)}} \xrightarrow{d} Y_u \sim \mathcal{N}(0, \sigma^2(u)), \qquad n \to \infty. \tag{10}$$

Here $\xrightarrow{d}$ denotes convergence in distribution, Y_u being a Gaussian random variable with mean zero and variance

$$\sigma^2(u) = \int_{\mathbb{R}^d} \operatorname{cov}\big(\mathbb{I}\{X(0) \geq u\}, \mathbb{I}\{X(t) \geq u\}\big)\, \mathrm{d}t \in \mathbb{R}_+. \tag{11}$$

Proof. Fix any $u \in \mathbb{R}$ and transform a random field X into a field $Z = \{Z(j), j \in \mathbb{Z}^d\}$, setting

$$Z(j) = \int_{Q_j} \mathbb{I}\{X(t) \geq u\}\, \mathrm{d}t - \mathsf{P}\big(X(0) \geq u\big), \tag{12}$$

where the unit cubes

$$Q_j = \{x = (x_1, \dots, x_d)^\top \in \mathbb{R}^d\colon\ 0 < x_k \le 1, k = 1, \dots, d\} \oplus \{j\}, \qquad j = (j_1, \dots, j_d)^\top \in \mathbb{Z}^d.$$

The Fubini theorem implies $\mathsf{E}Z(j) = 0$ for any $j \in \mathbb{Z}^d$. It is easily seen that the field Z is strictly stationary and square-integrable. Introduce

$$J_n^- = \{j \in \mathbb{Z}^d\colon\ Q_j \subset W_n\}, \qquad J_n^+ = \{j \in \mathbb{Z}^d\colon\ Q_j \cap W_n \neq \varnothing\} \tag{13}$$

and

$$W_n^- = \bigcup_{j \in J_n^-} Q_j, \qquad W_n^+ = \bigcup_{j \in J_n^+} Q_j.$$

Due to (7), we conclude (see [7], Lemma 3.1.2) that

$$\nu_d(W_n^-) \to \infty \quad \text{and} \quad \nu_d(W_n^-)/\nu_d(W_n^+) \to 1 \qquad \text{as } n \to \infty. \tag{14}$$

Write

$$\begin{aligned} &\frac{\nu_d(A_u(X, W_n)) - \nu_d(W_n)\mathsf{P}(X(0) \ge u)}{\sqrt{\nu_d(W_n)}} \\ &\quad = \frac{\nu_d(A_u(X, W_n^-)) - \nu_d(W_n^-)\mathsf{P}(X(0) \ge u)}{\sqrt{\nu_d(W_n)}} \\ &\qquad + \frac{\nu_d(A_u(X, W_n)) - \nu_d(A_u(X, W_n^-)) - (\nu_d(W_n) - \nu_d(W_n^-))\mathsf{P}(X(0) \ge u)}{\sqrt{\nu_d(W_n)}}. \end{aligned} \tag{15}$$

We prove that the second term on the right-hand side in (15) tends to zero in probability. By Chebyshev's inequality, it suffices to show that

$$\mathsf{var}\big(\nu_d(A_u(X, W_n)) - \nu_d(A_u(X, W_n^-))\big)/\nu_d(W_n) \to 0, \qquad n \to \infty.$$

Set

$$Y_n(j) = \int_{Q_j \cap W_n} \mathbb{I}\{X(t) \ge u\}\,\mathrm{d}t - \nu_d(Q_j \cap W_n)\mathsf{P}\big(X(0) \ge u\big) \qquad \text{for } j \in \mathbb{Z}^d,\ n \in \mathbb{N}.$$

Note that $Y_n(j) = Z(j)$ for $j \in J_n^-$ and $n \in \mathbb{N}$ (clearly $Y_n(j)$ and $Z(j)$ depend on u as well). Applying the Fubini theorem and Lemma 1, we get

$$\begin{aligned} &\mathsf{var}\big(\nu_d(A_u(X, W_n)) - \nu_d(A_u(X, W_n^-))\big) \\ &\quad = \mathsf{var}\Bigg(\sum_{j \in J_n^+ \setminus J_n^-} Y_n(j)\Bigg) \\ &\quad \le \sum_{j,m \in J_n^+ \setminus J_n^-} \int_{Q_j \times Q_m} \big|\mathsf{cov}\big(\mathbb{I}\{X(s) \ge u\}, \mathbb{I}\{X(t) \ge u\}\big)\big|\,\mathrm{d}s\,\mathrm{d}t \end{aligned} \tag{16}$$

$$\le \nu_d(W_n^+ \setminus W_n^-) \sum_{j \in \mathbb{Z}^d} \int_{Q_0 \times Q_j} C_1 |\operatorname{cov}(X(s), X(t))|^{1/3} \, \mathrm{d}s \, \mathrm{d}t$$

$$\le \nu_d(W_n^+ \setminus W_n^-) \left(C_2 + C_3 \sum_{r=r_0}^{\infty} \sum_{j \in \mathbb{Z}^d:\ \|j\|_\infty = r} \int_{Q_0 \times Q_j} \|s-t\|_2^{-\alpha/3} \, \mathrm{d}s \, \mathrm{d}t \right)$$

$$\le \nu_d(W_n^+ \setminus W_n^-) \left(C_2 + C_4 \sum_{r=1}^{\infty} r^{d-1} r^{-\alpha/3} \right) = C_5 \nu_d(W_n^+ \setminus W_n^-)$$

for some $r_0 > 0$ and all $n \in \mathbb{N}$. The factors C_i do not depend on n. We used the inequality $|\operatorname{cov}(X(s), X(t))| \le \tau^2$ for all $s, t \in \mathbb{R}^d$, which is satisfied as $\operatorname{var} X(t) = \tau^2$ for any $t \in \mathbb{R}^d$. We also took into account that

$$\operatorname{card}\{j \in \mathbb{Z}^d:\ \|j\|_\infty = r\} \le C_6 r^{d-1}$$

for each $r \in \mathbb{N}$ and employed the inequality (9) with $\alpha > 3d$.

By (14), (16) and in view of the relation $\nu_d(W_n^-) \le \nu_d(W_n) \le \nu_d(W_n^+)$, we get

$$\operatorname{var}\Big(\big(\nu_d(A_u(X, W_n)) - \nu_d(A_u(X, W_n^-))\big) / \sqrt{\nu_d(W_n)}\Big) \to 0, \qquad n \to \infty.$$

Now we show that J_n^-, introduced in (13), tends to infinity in a regular way as $n \to \infty$. Indeed, $J_n^- \subset J_n \subset J_n^+$, where $J_n := W_n \cap \mathbb{Z}^d$, $n \in \mathbb{Z}^d$. Due to [7], Lemma 3.1.5, J_n tends to infinity in a regular way. Thus, it suffices to mention that $\delta J_n^- \subset \delta J_n \cup (J_n \setminus J_n^-)$ and apply the relations $\operatorname{card} \delta J_n / \operatorname{card} J_n \to 0$ and $\operatorname{card} J_n^+ / \operatorname{card} J_n^- \to 1$ as $n \to \infty$. Lemma 3.1.6 of [7] implies that $W_n^- = \bigcup_{j \in J_n^-} Q_j$ tends to infinity in the Van Hove sense as $n \to \infty$.

So, while establishing (10), we can assume w.l.g. that $W_n = W_n^-$, that is, W_n is a finite union of cubes Q_j $(n \in \mathbb{N})$ and the sequence $(W_n)_{n \in \mathbb{N}}$ is VH-growing.

Observe that

$$\frac{\nu_d(A_u(X, W_n)) - \nu_d(W_n)\mathsf{P}(X(0) \ge u)}{\sqrt{\nu_d(W_n)}} = \frac{\sum_{j \in W_n \cap \mathbb{Z}^d} Z(j)}{\sqrt{\nu_d(W_n)}} := S_n.$$

As $X = \{X(t), t \in \mathbb{R}^d\} \in \mathsf{QA}$, it follows from (6) and (9) that $X \in (BL, \theta)$ with

$$\theta_r(X) = \mathrm{O}(r^{-\alpha+d}) \qquad \text{as } r \to \infty (r > 0).$$

For $\gamma > 0$ (and u fixed) introduce the Lipschitz functions $h_\gamma : \mathbb{R} \to \mathbb{R}$ by the formula

$$h_\gamma(x) = \begin{cases} 0, & \text{if } x \le u - \gamma, \\ (x - u + \gamma)/\gamma, & \text{if } u - \gamma < x \le u, \\ 1, & \text{otherwise.} \end{cases} \tag{17}$$

Superposition of two Lipschitz functions is also a Lipschitz one. Thus, for $n \in \mathbb{N}$ and $\gamma > 0$, the random field $Z_{n,\gamma} = \{Z_{n,\gamma}(j),\ j = (j_1, \dots, j_d)^\top \in \mathbb{Z}^d\} \in (BL, \theta)$, where

$$Z_{n,\gamma}(j) = \frac{1}{n^d} \sum_{k_1,\dots,k_d=1}^{n} h_\gamma\left(X\left(j_1 + \frac{k_1}{n}, \dots, j_d + \frac{k_d}{n}\right)\right) - \mathsf{E}h_\gamma(X(0)) \tag{18}$$

and the terms of a sequence $\theta(Z_{n,\gamma})$ admit the estimate

$$\theta_r(Z_{n,\gamma}) \le C_7 \gamma^{-2} r^{-\alpha+d}, \qquad r \in \mathbb{N}, \tag{19}$$

with C_7 depending neither on γ nor on n.

It is not difficult to verify that the finite-dimensional distributions of the fields $Z_{n,\gamma}$ weakly converge to the corresponding ones of the field Z_γ as $n \to \infty$, where

$$Z_\gamma(j) = \int_{Q_j} h_\gamma(X(t))\,\mathrm{d}t - \mathsf{E}h_\gamma(X(0)), \qquad j \in \mathbb{Z}^d. \tag{20}$$

Consequently (see [7], Lemma 1.5.16), we can claim that $Z_\gamma \in (BL, \theta)$ and $\theta_r(Z_\gamma)$ is bounded by the right-hand side of inequality (19). Theorem 3.1.12 of [7], guarantees that, for each $\gamma > 0$,

$$S_n(\gamma) := \frac{\sum_{j \in W_n \cap \mathbb{Z}^d} Z_\gamma(j)}{\sqrt{\nu_d(W_n)}} \xrightarrow{d} Y_{u,\gamma} \sim \mathcal{N}(0, \sigma^2(u,\gamma)), \qquad n \to \infty, \tag{21}$$

where

$$\sigma^2(u,\gamma) = \sum_{j \in \mathbb{Z}^d} \mathsf{cov}(Z_\gamma(0), Z_\gamma(j)) = \int_{\mathbb{R}^d} \mathsf{cov}(h_\gamma(X(0)), h_\gamma(X(t)))\,\mathrm{d}t \in \mathbb{R}_+.$$

Therefore, to prove (10), two steps remain. First of all, we estimate the difference of the characteristic functions of the random variables $S_n(\gamma)$ and S_n and show that it tends to zero as $\gamma \to 0+$. After that, we verify that

$$\sigma^2(u,\gamma) \to \sigma^2(u) \qquad \text{as } \gamma \to 0+\,. \tag{22}$$

Set $h(x) = \mathbb{I}\{x \ge u\}$ and $H_\gamma(x) = h_\gamma(x) - h(x)$, where $x \in \mathbb{R}$ (and $u \in \mathbb{R}$ is fixed). Then, for each $\lambda \in \mathbb{R}$, one has

$$\left|\mathsf{E}e^{\mathrm{i}\lambda S_n(\gamma)} - \mathsf{E}e^{\mathrm{i}\lambda S_n}\right| \le |\lambda| \mathsf{E}|S_n(\gamma) - S_n| \le |\lambda| \left(\frac{V_n(\gamma)}{\nu_d(W_n)}\right)^{1/2}, \tag{23}$$

where $\mathrm{i}^2 = -1$ and

$$V_n(\gamma) = \mathsf{E}\left(\sum_{j \in W_n \cap \mathbb{Z}^d} \int_{Q_j} \big(H_\gamma(X(t)) - \mathsf{E}H_\gamma(X(t))\big)\,\mathrm{d}t\right)^2.$$

It is easily seen that

$$V_n(\gamma) \le \nu_d(W_n) \int_{\mathbb{R}^d} |\operatorname{cov}(H_\gamma(X(0)), H_\gamma(X(t)))| \, dt. \tag{24}$$

Furthermore, we have

$$|\operatorname{cov}(H_\gamma(X(0)), H_\gamma(X(t)))| \le (\mathsf{E}(H_\gamma(X(0)))^2 \mathsf{E}(H_\gamma(X(t)))^2)^{1/2} \le a\gamma,$$

where a is a constant that bounds the density of $X(0)$. If $|\operatorname{cov}(X(0), X(t))|^{1/3} \le \gamma$, then reasoning similar to that proving Lemma 1 leads to the inequality

$$|\operatorname{cov}(H_\gamma(X(0)), H_\gamma(X(t)))| \le C(a) |\operatorname{cov}(X(0), X(t))|^{1/3} \tag{25}$$

with some $C(a) > 0$. Write $\alpha = 3(d+\mu)$, $\mu > 0$, and take $R = c\gamma^{-1/(d+\mu)}$, where $c > 0$. Then, in view of (9) and due to the appropriate choice of c, one can conclude that for all $\gamma > 0$ small enough,

$$\begin{aligned} F(\gamma) &:= \int_{\mathbb{R}^d} |\operatorname{cov}(H_\gamma(X(0)), H_\gamma(X(t)))| \, dt \\ &\le a\gamma\omega_d R^d + C_8 \int_{\|t\|_2 \ge R} \|t\|_2^{-\alpha/3} \, dt \le C_9 \gamma^{\mu/(d+\mu)}, \end{aligned} \tag{26}$$

where $\omega_d = \pi^{d/2}/(\Gamma(d/2+1))$ is the volume of the unit ball in $\mathbb{R}^d$ with the Euclidean norm. Consequently, inequalities (23), (24) and (26) imply that the laws of $S_n(\gamma)$ and S_n are close for all n large enough if $\gamma > 0$ is small enough.

Next, we proceed to (22). By the arguments leading to (26) and invoking Lemma 1, we deduce that $\sigma^2(u) < \infty$. Similar to (25), one shows that if $|\operatorname{cov}(X(s), X(t))|^{1/3} \le \gamma$, then

$$|\operatorname{cov}(h(X(s)), H_\gamma(X(t)))| \le D(a) |\operatorname{cov}(X(s), X(t))|^{1/3} \tag{27}$$

with $D(a) > 0$ depending on a only. The absolute value of $\sigma^2(u,\gamma) - \sigma^2(u)$ does not exceed the following expression:

$$F(\gamma) + \int_{\mathbb{R}^d} |\operatorname{cov}(h(X(0)), H_\gamma(X(t)))| \, dt + \int_{\mathbb{R}^d} |\operatorname{cov}(H_\gamma(X(0)), h(X(t)))| \, dt.$$

Taking into account the above upper bound and relations (26) and (27), we complete the proof of (22). The asymptotic (finite) variances $\sigma^2(u,\gamma)$ are non-negative, whence one concludes that $\sigma^2(u) \ge 0$.

In view of (21)–(23), the proof is complete. □

Now we turn to the multidimensional CLT for random vectors,

$$S_{\vec{u}}(X, W_n) = (\nu_d(A_{u_1}(X, W_n)), \dots, \nu_d(A_{u_r}(X, W_n)))^\top, \qquad n \in \mathbb{N}, \tag{28}$$

where $\vec{u} = (u_1, \dots, u_r)^\top \in \mathbb{R}^r$.

Theorem 2. *Let $X = \{X(t), t \in \mathbb{R}^d\}$ be a random field satisfying condition* (A). *Then, for each $\vec{u} = (u_1, \ldots, u_r)^\top \in \mathbb{R}^r$ and any VH-growing sequence $(W_n)_{n\in\mathbb{N}}$ of subsets of $\mathbb{R}^d$, one has*

$$\nu_d(W_n)^{-1/2}\big(S_{\vec{u}}(X, W_n) - \nu_d(W_n)P(\vec{u})\big) \xrightarrow{d} V_{\vec{u}} \sim \mathcal{N}(0, \Sigma(\vec{u})) \qquad \text{as } n \to \infty, \tag{29}$$

where

$$P(\vec{u}) = \big(\mathsf{P}\big(X(0) \geq u_1\big), \ldots, \mathsf{P}\big(X(0) \geq u_r\big)\big)^\top$$

and $\Sigma(\vec{u}) = (\sigma_{lm}(\vec{u}))_{l,m=1}^r$ is an $(r \times r)$-matrix having the elements

$$\sigma_{lm}(\vec{u}) = \int_{\mathbb{R}^d} \mathsf{cov}\big(\mathbb{I}\{X(0) \geq u_l\}, \mathbb{I}\{X(t) \geq u_m\}\big)\, \mathrm{d}t. \tag{30}$$

Proof. Observe that the convergence of all r^2 integrals in (30) is proved in the same way as that of the integral representing $\sigma^2(u)$ in the one-dimensional case. The result follows by using the Cramér–Wold device. We omit further details that are quite similar to those in the proof of Theorem 1. □

The last theorem entails:

Corollary 1. *Let $X = \{X(t), t \in \mathbb{R}^d\}$ be a random field satisfying condition* (A). *Assume that $\Sigma(\vec{u})$ is non-degenerate for some $\vec{u} \in \mathbb{R}^r$. Then, for this $\vec{u}$ and any sequence $(W_n)_{n\in\mathbb{N}}$ of VH-growing subsets of $\mathbb{R}^d$, one has*

$$\nu_d(W_n)^{-1/2}\Sigma(\vec{u})^{-1/2}\big(S_{\vec{u}}(X, W_n) - \nu_d(W_n)P(\vec{u})\big) \xrightarrow{d} V \sim \mathcal{N}(0, \mathrm{I}) \qquad \text{as } n \to \infty;$$

here I denotes the unit $(r \times r)$-matrix.

3.2. Shot noise random fields

We verify the conditions of Theorem 1 for shot noise random fields. These fields appear naturally in the theory of disordered structures. Let $\mathcal{B}(\mathbb{R}^d)$ (resp., $\mathcal{B}_0(\mathbb{R}^d)$) be the family of all (bounded) Borel sets in $\mathbb{R}^d$. A shot noise random field $X = \{X(t), t \in \mathbb{R}^d\}$ is defined by the relation

$$X(t) = \sum_{i\in\mathbb{N}} \xi_i \varphi(t - x_i),$$

where $\{\xi_i\}$ is a family of i.i.d. non-negative random variables and $\{x_i\}$ is a homogeneous Poisson point process in $\mathbb{R}^d$ with intensity $\lambda \in (0, \infty)$, that is, $\{x_i\}$ is the support set of a random Poisson counting measure $\{N_B, B \in \mathcal{B}(\mathbb{R}^d)\}$, where $N_B = \#\{i\colon x_i \in B\}$ has the following properties:

(i) $N_{B_1}, N_{B_2}, \ldots$ are independent for pairwise disjoint $B_1, B_2, \ldots \in \mathcal{B}_0(\mathbb{R}^d)$,
(ii) $N_B \sim \mathrm{Pois}(\lambda\nu_d(B))$ for all $B \in \mathcal{B}_0(\mathbb{R}^d)$.

Suppose that $\{\xi_i\}$, $N_{(\cdot)}$ are independent, $\mathsf{E}\xi_i^2 < \infty$ and $\varphi : \mathbb{R}^d \to \mathbb{R}_+$ is a Borel function.

For the shot-noise field X introduced above, we impose the condition:

(C) $X(0)$ has a bounded density and for a function φ bounded and uniformly continuous on $\mathbb{R}^d$,

$$\varphi(t) \le \varphi_0(\|t\|_2) = \mathrm{O}(\|t\|_2^{-\alpha}) \qquad \text{as } \|t\|_2 \to \infty, \tag{31}$$

where $\alpha > 3d$ and $\varphi_0 : \mathbb{R}_+ \to \mathbb{R}_+$.

Proposition 1. *The statement of Theorem 1 holds for a random field X satisfying condition* (C).

Proof. By [7], Theorem 1.3.8, X is associated and hence quasi-associated. Moreover, it is strictly stationary with covariance function given, for example, in [7], Theorem 2.3.6. The continuity of the covariance function follows from the inequality

$$|\operatorname{cov}(X(0), X(s)) - \operatorname{cov}(X(0), X(t))| \le \lambda \mathsf{E}\xi_1^2 \sup_{y\in\mathbb{R}^d} |\varphi(t-y) - \varphi(s-y)| \int_{\mathbb{R}^d} |\varphi(y)|\,\mathrm{d}y$$

and the uniform continuity of φ. Corollary 2.3.7 of [7] yields the desired bound for the covariance function in condition (A). The proof is complete. □

Note that the characteristic function of $X(0)$, provided by [7], Lemma 1.3.7, is integrable if

$$\int_{\mathbb{R}} \left| \exp\left\{ \lambda \int_{\mathbb{R}^d} \big(\varphi_\xi(s\varphi(t)) - 1\big)\,\mathrm{d}t \right\} \right| \mathrm{d}s < \infty. \tag{32}$$

Thus, (32) guarantees the existence of the bounded density of $X(0)$.

Condition (32) can be easily verified in a number of special cases; for instance, if $\xi_1 = \text{const} > 0$ a.s. and $\varphi(t) = a\exp\{-b\|t\|_2\}$ or $\varphi(t) = a\min\{1, \|t\|_2^{-b}\}$ with $a, b > 0$.

3.3. Gaussian random fields

In contrast to Lemma 1, we obtain a sharper estimate for the covariance of indicator functions in the Gaussian case. Our result extends formula (2.7.1) of [12]. Let Φ and Ψ stand for the cumulative distribution function and the tail distribution function of a standard Gaussian random variable, respectively.

Lemma 2. *Let $(U, V)^\top$ be a Gaussian random vector in $\mathbb{R}^2$ such that $U \sim \mathcal{N}(a, \tau^2)$, $V \sim \mathcal{N}(a, \tau^2)$, where $a \in \mathbb{R}$, $\tau > 0$ and correlation coefficient $\operatorname{corr}(U, V) = \rho$. Then, for any $u, v \in \mathbb{R}$ and $\rho \in (-1, 1)$, the following equality holds*:

$$\begin{aligned}&\operatorname{cov}(\mathbb{I}\{U \ge u\}, \mathbb{I}\{V \ge v\}) \\ &\quad = \frac{1}{2\pi} \int_0^\rho \frac{1}{\sqrt{1-r^2}} \exp\left\{ -\frac{(u-a)^2 - 2r(u-a)(v-a) + (v-a)^2}{2\tau^2(1-r^2)} \right\} \mathrm{d}r.\end{aligned} \tag{33}$$

In particular, for $u = v$, *one has*

$$\mathsf{cov}(\mathbb{I}\{U \ge u\}, \mathbb{I}\{V \ge u\}) = \frac{1}{2\pi} \int_0^\rho \frac{1}{\sqrt{1-r^2}} \exp\left\{-\frac{(u-a)^2}{\tau^2(1+r)}\right\} \mathrm{d}r.$$

Moreover, for any $u, v \in \mathbb{R}$ *and* $\rho \in [-1, 1]$, *one has the inequality*

$$|\,\mathsf{cov}(\mathbb{I}\{U \ge u\}, \mathbb{I}\{V \ge v\})| \le |\rho|/4. \tag{34}$$

Proof. Using the transformation $x \mapsto (x-a)/\tau$, $x \in \mathbb{R}$, we can assume w.l.g. that $U \sim \mathcal{N}(0,1)$ and $V \sim \mathcal{N}(0,1)$. Let $\rho \in (-1, 1)$. The probability density

$$f_{U,V}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right\}$$

of the bivariate Gaussian random variable $(U, V)^\top$ is invariant under the transformation $x \mapsto -x$ and $y \mapsto -y$, $(x, y)^\top \in \mathbb{R}^2$. Therefore,

$$\mathsf{cov}(\mathbb{I}\{U \ge u\}, \mathbb{I}\{V \ge v\}) = \mathsf{cov}(\mathbb{I}\{U \le -u\}, \mathbb{I}\{V \le -v\}), \qquad u, v \in \mathbb{R}.$$

It is well known (see, e.g., [10], formulae (21.12.5) and (21.12.6)) that

$$f_{U,V}(x, y) = \sum_{k=0}^{\infty} \frac{\Phi^{(k+1)}(x)\Phi^{(k+1)}(y)}{k!} \rho^k, \qquad x, y \in \mathbb{R},$$

where $\Phi^{(k)}(x) = \mathrm{d}^k\Phi(x)/\mathrm{d}x^k$ and, for any $u, v \in \mathbb{R}$,

$$\int_{-\infty}^{u} \int_{-\infty}^{v} f_{U,V}(x, y)\, \mathrm{d}x\, \mathrm{d}y = \sum_{k=0}^{\infty} \frac{\Phi^{(k)}(u)\Phi^{(k)}(v)}{k!} \rho^k.$$

Hence, for each $u, v \in \mathbb{R}$,

$$\begin{aligned}
\mathsf{E}\mathbb{I}\{U \le -u\}\mathbb{I}\{V \le -v\} &= \int_{-\infty}^{-u} \int_{-\infty}^{-v} f_{U,V}(x, y)\, \mathrm{d}x\, \mathrm{d}y = \sum_{k=0}^{\infty} \frac{\Phi^{(k)}(-u)\Phi^{(k)}(-v)}{k!} \rho^k \\
&= \Phi(-u)\Phi(-v) + \sum_{k=1}^{\infty} \frac{\Phi^{(k)}(-u)\Phi^{(k)}(-v)}{k!} \rho^k \\
&= \Phi(-u)\Phi(-v) + \int_0^\rho \sum_{k=0}^{\infty} \frac{\Phi^{(k+1)}(-u)\Phi^{(k+1)}(-v)}{k!} r^k\, \mathrm{d}r \\
&= \Phi(-u)\Phi(-v) + \int_0^\rho f_{U(r),V(r)}(u, v)\, \mathrm{d}r,
\end{aligned}$$

where $(U(r), V(r))^\top$ is a centered bivariate Gaussian vector with $\mathsf{E}U(r)^2 = \mathsf{E}V(r)^2 = 1$ and $\mathsf{cov}(U(r), V(r)) = r$. Consequently, we get

$$\begin{aligned}\mathsf{cov}(\mathbb{I}\{U \le -u\}, \mathbb{I}\{V \le -v\}) &= \int_0^\rho f_{U(r),V(r)}(u, v)\,\mathrm{d}r \\ &= \frac{1}{2\pi}\int_0^\rho \frac{1}{\sqrt{1-r^2}} \mathrm{e}^{-(u^2-2ruv+v^2)/(2(1-r^2))}\,\mathrm{d}r.\end{aligned}$$

Passing to random variables U and V with arbitrary mean a and variance $\tau^2 > 0$ gives the formula (33).

To prove inequality (34) for $\rho \in (-1, 1)$, write

$$|\mathsf{cov}(\mathbb{I}\{U \le -u\}, \mathbb{I}\{V \le -v\})| \le \frac{1}{2\pi}\int_0^{|\rho|} \frac{1}{\sqrt{1-r^2}}\,\mathrm{d}r \le \frac{1}{2\pi}\arcsin|\rho|$$

and notice that $\arcsin|\rho| \le \pi|\rho|/2$.

The case $|\rho| = 1$ is trivial, as $|\mathsf{P}(A \cap B) - \mathsf{P}(A)\mathsf{P}(B)| \le 1/4$ for any $A, B \in \mathcal{F}$. □

The following result generalizes the corresponding one established in [12] (see page 80), where the isotropy of the Gaussian random field was assumed. A central limit theorem for nonlinear transformations of a homogeneous Gaussian random field was used there.

Theorem 3. *Let $X = \{X(t), t \in \mathbb{R}^d\}$ be a Gaussian stationary random field satisfying condition* (B) *and $X(0) \sim \mathcal{N}(a, \tau^2)$. Then, for each $u \in \mathbb{R}$ and any sequence of VH-growing sets $W_n \subset \mathbb{R}^d$, one has*

$$\frac{\nu_d(A_u(X, W_n)) - \nu_d(W_n)\Psi((u-a)/\tau)}{\sqrt{\nu_d(W_n)}} \xrightarrow{d} Y_u \sim \mathcal{N}(0, \sigma^2(u))$$

as $n \to \infty$. The variance $\sigma^2(u)$ introduced in (11) *can be written in the following form*:

$$\sigma^2(u) = \frac{1}{2\pi}\int_{\mathbb{R}^d}\int_0^{\rho(t)} \frac{1}{\sqrt{1-r^2}} \exp\left\{-\frac{(u-a)^2}{\tau^2(1+r)}\right\}\mathrm{d}r\,\mathrm{d}t, \tag{35}$$

where $\rho(t) = \mathsf{corr}(X(0), X(t))$. In particular,

$$\sigma^2(a) = \frac{1}{2\pi}\int_{\mathbb{R}^d} \arcsin(\rho(t))\,\mathrm{d}t.$$

Proof. For the Gaussian field X, we have $\mathsf{P}(X(0) \ge u) = \Psi((u-a)/\tau)$. Now we apply the upper bound (34) to obtain $|\mathsf{cov}(X(0), X(t))|$ instead of $|\mathsf{cov}(X(0), X(t))|^{1/3}$ in the estimates used in the proof of Theorem 1. This leads to the hypothesis that $\alpha > d$ in (9), whereas in condition (A) we assumed $\alpha > 3d$. Note that Gaussian fields are quasi-associated [20].

Finally, we express $\sigma^2(u)$ (see (11)) in terms of the covariance function of X as in the proof of Lemma 2, and this yields (35). The proof is complete. □

Theorem 4. *Let $X = \{X(t), t \in \mathbb{R}^d\}$ be a random field satisfying condition* (B) *and $X(0) \sim \mathcal{N}(a, \tau^2)$. Then, for each $\vec{u} = (u_1, \ldots, u_r)^\top \in \mathbb{R}^r$ and any sequence $(W_n)_{n\in\mathbb{N}}$ of VH-growing subsets of $\mathbb{R}^d$, one has*

$$\nu_d(W_n)^{-1/2}\big(S_{\vec{u}}(X, W_n) - \nu_d(W_n)\Psi(\vec{u})\big) \xrightarrow{d} V_{\vec{u}} \sim \mathcal{N}(0, \Sigma(\vec{u})) \qquad \text{as } n \to \infty. \tag{36}$$

Here, $\Psi(\vec{u}) = (\Psi((u_1 - a)/\tau), \ldots, \Psi((u_r - a)/\tau))^\top$ and $\Sigma(\vec{u}) = (\sigma_{lm}(\vec{u}))_{l,m=1}^r$ is a matrix having the elements

$$\sigma_{lm}(\vec{u}) = \frac{1}{2\pi}\int_{\mathbb{R}^d}\int_0^{\rho(t)} g(r)\,\mathrm{d}r\,\mathrm{d}t, \tag{37}$$

where

$$g(r) = \frac{1}{\sqrt{1-r^2}}\exp\left\{-\frac{(u_l - a)^2 - 2r(u_l - a)(u_m - a) + (u_m - a)^2}{2\tau^2(1-r^2)}\right\}$$

and $\rho(t) = \mathsf{corr}(X(0), X(t))$. If $\Sigma(\vec{u})$ is non-degenerate, we obtain by virtue of (36)

$$\nu_d(W_n)^{-1/2}\Sigma(\vec{u})^{-1/2}\big(S_{\vec{u}}(X, W_n) - \nu_d(W_n)\Psi(\vec{u})\big) \xrightarrow{d} \mathcal{N}(0, \mathrm{I}), \qquad n \to \infty,$$

where I is the unit $(r \times r)$-matrix.

Proof. Employing Lemma 2, one can repeat the reasoning proving Theorem 2. Clearly, $\mathsf{P}(X(0) \geq u_l) = \Psi((u_l - a)/\tau)$, $l = 1, \ldots, r$. The matrix elements $\sigma_{lm}(\vec{u})$ for $l, m = 1, \ldots, r$ can be calculated by way of (33). □

Formulae in the isotropic case

For the isotropic case, we use the change of variables (passing from $t = (t_1, \ldots, t_d)^\top$ to spherical coordinates) in integrals (35) and (37) to obtain the following statement.

Corollary 2. *Let the random field $X = \{X(t), t \in \mathbb{R}^d\}$ satisfying the conditions of Theorem 3 be isotropic ($d \geq 2$). Then*

$$\sigma^2(u) = \frac{\mathrm{d}\omega_d}{2\pi}\int_0^\infty v^{d-1}\int_0^{\rho(v)} \frac{1}{\sqrt{1-r^2}}\exp\left\{-\frac{(u-a)^2}{\tau^2(1+r)}\right\}\mathrm{d}r\,\mathrm{d}v,$$

where $\rho(v) = \mathsf{corr}(X(0), X(t))$ if $|t| = v$. For $u = a$, one has

$$\sigma^2(a) = \frac{\mathrm{d}\omega_d}{2\pi}\int_0^\infty v^{d-1}\arcsin(\rho(v))\,\mathrm{d}v.$$

In the multivariate case, (37) *can be written as follows:*

$$\sigma_{lm}(\vec{u}) = \frac{\mathrm{d}\omega_d}{2\pi}\int_0^\infty v^{d-1}\int_0^{\rho(v)} g(r)\,\mathrm{d}r\,\mathrm{d}v$$

for $l, m = 1, \ldots, r$.

4. Statistical version of the CLT

Now we provide a statistical version of the CLT involving random self-normalization. Let $r \in \mathbb{N}$ be the number of levels to observe.

Theorem 5. *Let $X = \{X(t), t \in \mathbb{R}^d\}$ be a random field satisfying condition* (A). *Let $u_k \in \mathbb{R}$, $k = 1, \ldots, r$ and $(W_n)_{n\in\mathbb{N}}$ be a sequence of VH-growing sets. Furthermore, let $\hat{C}_n = (\hat{c}_{nlm})_{l,m=1}^r$ be statistical estimates for non-degenerate asymptotic covariance matrix Σ with elements σ_{lm} given by* (30). *Assume that $\hat{c}_{nlm} \xrightarrow{p} \sigma_{lm}$ as $n \to \infty$ for any $l, m = 1, \ldots, r$, where $\xrightarrow{p}$ denotes convergence in probability. Then*

$$\hat{C}_n^{-1/2} \nu_d(W_n)^{-1/2}\big(S(W_n) - \nu_d(W_n)P(\vec{u})\big) \xrightarrow{d} \mathcal{N}(0, \mathrm{I}) \qquad \text{as } n \to \infty.$$

Proof. It suffices to use Theorem 2 and elementary properties of the convergence in probability and in law for random vectors. □

One feasible estimator for the asymptotic covariance matrix Σ that arose in the multivariate CLT, see Theorem 2, can be called a *subwindow estimator* [18] and is constructed as follows. Let $(V_n)_{n\in\mathbb{N}}$ and $(W_n)_{n\in\mathbb{N}}$ be sequences of VH-growing sets (not necessarily rectangles) such that $V_n \subset W_n$, $n \geq 1$. Consider $N(n)$ subwindows $V_{n,1}, \ldots, V_{n,N(n)}$, where $(N(n))_{n\in\mathbb{N}}$ is an increasing sequence of integers with $\lim_{n\to\infty} N(n) = \infty$, and $V_{n,j} = V_n \oplus \{h_{n,j}\}$ are subwindows that are translated by certain vectors $h_{n,j} \in \mathbb{R}^d$, $j = 1, \ldots, N(n)$. Assume that $\bigcup_{j=1}^{N(n)} V_{n,j} \subseteq W_n$ for each $n \in \mathbb{N}$ and there exists some $r > 0$ such that

$$V_{n,j} \cap V_{n,i} \subset \partial V_{n,j} \oplus B_r(0) \qquad \text{for } i, j \in \{1, \ldots, N(n)\} \text{ with } i \neq j.$$

Denote by

$$\hat{\mu}_{nk}^{(j)} = \frac{1}{\nu_d(V_n)} \int_{V_{n,j}} \mathbb{I}\{X(t) \geq u_k\}\,\mathrm{d}t, \qquad j = 1, \ldots, N(n),$$

the estimator of $\mu_k = \mathsf{P}(X(0) \geq u_k)$ based on observations within $V_{n,j}$, and by

$$\bar{\mu}_{nk} = \frac{1}{N(n)} \sum_{j=1}^{N(n)} \hat{\mu}_{nk}^{(j)}, \qquad n \in \mathbb{N},\ k = 1, \ldots, r,$$

the average of these estimators. After all, we define the estimator $\hat{\Sigma}_n = (\hat{\sigma}_{nlm})_{l,m=1}^r$ for the covariance matrix Σ. Set

$$\hat{\sigma}_{nlm} = \frac{\nu_d(V_n)}{N(n) - 1} \sum_{j=1}^{N(n)} \big(\hat{\mu}_{nl}^{(j)} - \bar{\mu}_{nl}\big)\big(\hat{\mu}_{nm}^{(j)} - \bar{\mu}_{nm}\big). \tag{38}$$

We recall the following result.

Theorem 6 ([18], Theorem 3). *Let $X = \{X(t), t \in \mathbb{R}^d\}$ be a strictly stationary random field such that*

$$\int_{\mathbb{R}^{3d}} \left|c_{lm}^{(2,2)}(x, y, z)\right| \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z < \infty, \qquad l, m = 1, \ldots, r, \tag{39}$$

where the fourth-order cumulant function

$$c_{lm}^{(2,2)}(x, y, z) = \mathsf{E}\big([Z_l(0) - \mu_l][Z_m(x) - \mu_m][Z_l(y) - \mu_l][Z_m(z) - \mu_m]\big)$$
$$- \mathsf{cov}_{lm}(x)\,\mathsf{cov}_{lm}(z - y) - \mathsf{cov}_{ll}(y)\,\mathsf{cov}_{mm}(x - z) - \mathsf{cov}_{lm}(z)\,\mathsf{cov}_{ml}(x - y)$$

and $\mathsf{cov}_{lm}(t) = \mathsf{cov}(\mathbb{I}\{X(0) \geq u_l\}, \mathbb{I}\{X(t) \geq u_m\})$, $l, m = 1, \ldots, r$. *Then $\hat{\Sigma}$ introduced in (38) is mean-square consistent.*

Relation (39) holds for a random field X with finite dependence range. In this case, the estimator $\hat{\Sigma}_n$ is mean-square consistent. Among other estimators for the asymptotic covariance matrix, there are two worth mentioning. One estimator that, under certain assumptions, meets the conditions of Theorem 5 is introduced in [6,8] and involves local averaging. A major disadvantage is tedious calculation in the case of a large observation window. The same problem arises for an estimator based on the covariance function estimation for the underlying random field; see [18] (cf. [12], Chapter 4).

5. Discussion

A very important issue for applications of the estimator $\hat{\Sigma}_n$ is the choice of an appropriate size of the (e.g., rectangular) subwindow V_n. The subwindow size is related to both the covariance structure of the considered random field and the size of the observation window. We will discuss these problems while considering a simple example. The data used consist of 100 mutually independent realizations of stationary and centered Gaussian random field X with covariance function

$$\mathsf{cov}(X(0), X(t)) = \left(1 - \frac{3\|t\|_2}{2a} + \frac{\|t\|_2^3}{2a^3}\right)\mathbb{I}\{\|t\|_2 \in [0, a]\}, \qquad t \in \mathbb{R}^2,$$

for some $a > 0$ according to the spherical covariance model (see [21], page 244), which is often applied in geostatistics. The correlation range a in our simulation study has to be small enough in comparison with the size of the observation window to make the CLT argument work. Here, we take, for example, $a = 10$. The fields are simulated in the observation window $W = [0, 2000) \times [0, 2000)$ on the grid with mesh size one. That means every realization provides 4 million data points. To generate level sets, we consider the thresholds $u_1 = -1.0$, $u_2 = 0.0$ and $u_3 = 1.0$. Then

$$\Sigma = \begin{pmatrix} 4.6432 & & \\ 5.9938 & 10.5564 & \\ 2.7962 & 5.9938 & 4.6432 \end{pmatrix}.$$

An appropriate subwindow size can be found focusing only on the threshold $u_2 = 0.0$, since for other threshold values the obtained results differ from this one only slightly. The estimator provides the best result for Σ as the edge length of the rectangular subwindow equals 15. In general, the optimal choice of this length is an open non-trivial problem. After this preliminary step, we are able to apply the subwindow estimator to the simulated data. The following two matrices show averaged estimation results for Σ by means of $\hat{\Sigma}$. On the left-hand side, the averaged value of each estimated matrix element is computed out of 100 samples. On the right-hand side, the mean error to the theoretical value is provided.

$$\frac{1}{100}\sum_{k=1}^{100}\hat{\Sigma}_k = \begin{pmatrix} 4.6556 & & \\ 5.9710 & 10.5524 & \\ 2.8156 & 5.9934 & 4.6762 \end{pmatrix}, \qquad \mathrm{ME} = \begin{pmatrix} 0.27\% & & \\ -0.38\% & -0.04\% & \\ 0.69\% & \approx 0.00\% & 0.71\% \end{pmatrix}.$$

It would be interesting to propose a statistical hypothesis test based on the established statistical version of the CLT in order to apply it to data concerning the paper production. We will deal with this topic in a separate paper.

6. Open problems

The research area of limit theorems for level sets of random surfaces still offers an abundance of open problems. Let us mention just a few. It would be desirable to prove limit theorems for joint distributions of various surface characteristics of different classes of random fields. For instance, one could consider stable fields. Further on, one can study random fields possessing more strong dependence structure; for example, satisfying condition (A) with $\alpha \le d$. In this case, the normalizing factors have to be changed and the limiting distributions can be non-Gaussian. Certain results for problems of this type can be found in [12,15]. One could also prove a functional limit theorem for an innumerable set of thresholds. As our main result could also be called the CLT for the first Minkowski functional, it might be of interest to prove limit theorems involving other Minkowski functionals for level sets such as the boundary length or the Euler characteristics. It is worth mentioning that, for a stationary two-dimensional Gaussian field, this has already been done for the second Minkowski functional in [14].

Acknowledgements

The authors are grateful to Professor Richard A. Davis, the Associate Editor and the referees for valuable remarks and suggestions permitting them to improve the exposition of the paper. Alexander Bulinski's work was partially supported by RFBR Grant 10-01-00397-a.

References

[1] Adler, R.J. (1981). *The Geometry of Random Fields*. Chichester: Wiley. MR0611857
[2] Adler, R.J. and Taylor, J.E. (2007). *Random Fields and Geometry*. New York: Springer. MR2319516

[3] Belyaev, Y. (1972). The general formulae for the mean number of crossings for random processes and fields. In *Bursts of Random Fields* 38–45. Moscow: Moscow Univ. Press. (In Russian.)
[4] Brown, P.E., Diggle, P.J. and Henderson, R. (2003). A non-Gaussian spatial process model for opacity of flocculated paper. *Scand. J. Statist.* **30** 355–368. MR1983130
[5] Bulinski, A. (2010). Central limit theorem for random fields and applications. In *Advances in Data Analysis. Stat. Ind. Technol.* 141–150. Boston, MA: Birkhäuser. MR2641266
[6] Bulinski, A. and Kryzhanovskaya, N. (2006). Convergence rate in CLT for vector-valued random fields with self-normalization. *Probab. Math. Statist.* **26** 261–281. MR2325308
[7] Bulinski, A.V. and Shashkin, A.P. (2007). *Limit Theorems for Associated Random Fields and Related Systems*. Singapore: World Scientific.
[8] Bulinskiĭ, A.V. (2004). A statistical version of the central limit theorem for vector-valued random fields. *Math. Notes* **76** 455–464.
[9] Bulinskiĭ, A.V. and Shabanovich, È. (1998). Asymptotic behavior of some functionals of positively and negatively dependent random fields. *Fundam. Prikl. Mat.* **4** 479–492. MR1801168
[10] Cramér, H. (1946). *Mathematical Methods of Statistics. Princeton Mathematical Series* **9**. Princeton, NJ: Princeton Univ. Press. MR0016588
[11] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications. Vol. II*, 2nd ed. New York: Wiley. MR0270403
[12] Ivanov, A.V. and Leonenko, N.N. (1989). *Statistical Analysis of Random Fields. Mathematics and Its Applications (Soviet Series)* **28**. Dordrecht: Kluwer. MR1009786
[13] Joag-Dev, K. and Proschan, F. (1983). Negative association of random variables, with applications. *Ann. Statist.* **11** 286–295. MR0684886
[14] Kratz, M.F. and León, J.R. (2001). Central limit theorems for level functionals of stationary Gaussian processes and fields. *J. Theoret. Probab.* **14** 639–672. MR1860517
[15] Leonenko, N. (1990). *Limit Theorems for Random Fields with Singular Spectrum*. Dordrecht: Kluwer.
[16] Leonenko, N.N. (1987). Limit distributions of the characteristics of exceeding of a level by a Gaussian random field. *Math. Notes* **41** 339–345.
[17] Leonenko, N.N. (1988). Sharpness of the normal approximation of functionals of strongly correlated Gaussian random fields. *Math. Notes* **43** 161–171. MR0939529
[18] Pantle, U., Schmidt, V. and Spodarev, E. (2010). On the estimation of integrated covariance functions of stationary random fields. *Scand. J. Stat.* **37** 47–66. MR2675939
[19] Pitt, L.D. (1982). Positively correlated normal variables are associated. *Ann. Probab.* **10** 496–499. MR0665603
[20] Shashkin, A.P. (2002). Quasi-associatedness of a Gaussian system of random vectors. *Russ. Math. Surv.* **57** 1243–1244. MR1991881
[21] Wackernagel, H. (1998). *Multivariate Geostatistics*. Berlin: Springer.

Received May 2010 and revised August 2010