

# SCAN STATISTICS OF LÉVY NOISES AND MARKED EMPIRICAL PROCESSES

ZAKHAR KABLUCHKO,\* *Georg-August-Universität Göttingen*

EVGENY SPODAREV,\*\* *Universität Ulm*

## Abstract

Let  $n$  points be chosen independently and uniformly in the unit cube  $[0, 1]^d$ , and suppose that each point is supplied with a mark, the marks being independent and identically distributed random variables independent of the location of the points. To each cube  $R$  contained in  $[0, 1]^d$  we associate its score  $\mathcal{X}_n(R)$  defined as the sum of marks of all points contained in  $R$ . The scan statistic is defined as the maximum of  $\mathcal{X}_n(R)$ , taken over all cubes  $R$  contained in  $[0, 1]^d$ . We show that if the marks are nonlattice random variables with finite exponential moments, having negative mean and assuming positive values with nonzero probability, then the appropriately normalized distribution of the scan statistic converges as  $n \rightarrow \infty$  to the Gumbel distribution. We also prove a corresponding result for the scan statistic of a Lévy noise with negative mean. The more elementary cases of zero and positive mean are also considered.

*Keywords:* Scan statistic with variable window size; marked empirical process; independently scattered Lévy random measure; extremes; Pickands' double sum method

2000 Mathematics Subject Classification: Primary 60G70

Secondary 60F05

## 1. Introduction

Let  $\{U_i, i = 1, \dots, n\}$  be  $n$  points chosen independently and uniformly from the  $d$ -dimensional unit cube  $[0, 1]^d$ . Suppose that to each point  $U_i$  a mark  $X_i$  is attached, the marks being independent and identically distributed (i.i.d.) real-valued random variables independent of  $U_1, \dots, U_n$ . The collection  $\{(U_i, X_i), i = 1, \dots, n\}$  is called the *marked empirical process*. A natural problem is how to detect inhomogeneities, e.g. clustering of unusually big marks, in the marked empirical process. To this end, we may consider the *scan statistic*, whose definition we now recall (see [9] and [10]). For a set  $R \subset [0, 1]^d$ , define its score  $\mathcal{X}_n(R)$  as the sum of marks of all points contained in  $R$ , that is,

$$\mathcal{X}_n(R) = \sum_{i \in \{1, \dots, n\}: U_i \in R} X_i. \quad (1.1)$$

Then the scan statistic is defined as  $\sup_{R \in \mathcal{R}(1)} \mathcal{X}_n(R)$ , where  $\mathcal{R}(1)$  is some collection of subsets ('windows') of  $[0, 1]^d$ . Since no a priori assumptions about the size of the clusters are made, it is natural to require  $\mathcal{R}(1)$  to contain windows of all sizes. In this paper we take  $\mathcal{R}(1)$  to be the collection of all cubes contained in  $[0, 1]^d$  (a cube is a translate of the set  $[0, x]^d$  for some  $x > 0$ ).

Received 7 July 2008; revision received 17 November 2008.

\* Postal address: Institut für Mathematische Stochastik, Georg-August-Universität Göttingen, Goldschmidtstr. 7, D-37077 Göttingen, Germany. Email address: kabluch@math.uni-goettingen.de

\*\* Postal address: Institut für Stochastik, Universität Ulm, Helmholtzstr. 18, D-89069 Ulm, Germany.

The main question is then how the scan statistic is distributed as  $n \rightarrow \infty$ . To state our main result, we make the following assumptions about the distribution of the random marks.

- (X1) The logarithmic moment generating function of  $X_1$ ,  $\psi(\theta) = \log E[\exp\{\theta X_1\}]$ , exists as long as  $\theta \in [0, \theta_0)$  for some  $\theta_0 \in (0, \infty]$  which is supposed to be maximal with this property.
- (X2) The function  $\psi$  has a zero  $\theta^* \in (0, \theta_0)$ .
- (X3) The distribution of  $X_1$  is nonlattice.

Note that condition (X2) implies that  $E[X_1] = \psi'(0) < 0$ . A further corollary is that  $P[X_1 > 0] \neq 0$ . Conversely, if (i)  $E[X_1] < 0$ , (ii)  $P[X_1 > 0] \neq 0$ , and (iii) condition (X1) is satisfied for  $\theta_0 = \infty$ , then condition (X2) is fulfilled automatically.

Let  $\mathcal{R}(1)$  be the collection of all cubes contained in  $[0, 1]^d$ . Our main result reads as follows.

**Theorem 1.1.** *Let  $\{(U_i, X_i), i = 1, \dots, n\}$  be a marked empirical process satisfying conditions (X1)–(X3) above. Then*

$$\lim_{n \rightarrow \infty} P \left[ \sup_{R \in \mathcal{R}(1)} \mathcal{X}_n(R) \leq \frac{1}{\theta^*} (\log n + (d - 1) \log \log n + H^* + \tau) \right] = \exp\{-e^{-\tau}\}$$

for every  $\tau \in \mathbb{R}$ , where  $H^*$  is a constant to be specified later.

The scan statistic of Theorem 1.1 may be interpreted as a likelihood ratio test statistic in the following sense. Suppose that we are given a set of points  $\{U_i, i = 1, \dots, n\}$  in  $[0, 1]^d$ , the point  $U_i$  being marked by a number  $Y_i$ . Let  $F_0$  and  $F_1$  be two distribution functions such that the density  $p = dF_1/dF_0$  exists. Consider the following hypotheses (here,  $R \in \mathcal{R}(1)$ ).

$H_0$ :  $U_1, \dots, U_n, Y_1, \dots, Y_n$  are independent with the following distributions: the  $U_i$  are uniformly distributed in  $[0, 1]^d$  and  $Y_i \sim F_0$ .

$H_R$ :  $U_1, \dots, U_n$  are independent and uniformly distributed in  $[0, 1]^d$ , whereas, for each  $i = 1, \dots, n$ , the random variable  $Y_i$  is independent of  $\{U_1, \dots, U_n\} \setminus \{U_i\}$  and

$$Y_i \sim \begin{cases} F_1 & \text{if } U_i \in R, \\ F_0 & \text{if } U_i \notin R. \end{cases}$$

$H_1$ :  $\bigcup_{R \in \mathcal{R}(1)} H_R$ .

It is easy to see that the log-likelihood ratio statistic for testing  $H_0$  against  $H_R$  is given by  $\mathcal{X}_n(R)$  defined in (1.1) with  $X_i = \log p(Y_i)$ . Thus, the scan statistic considered in Theorem 1.1 may be interpreted as a likelihood ratio statistic for testing  $H_0$  against  $H_1$ .

In the one-dimensional case, the distribution of the scan statistic with variable window size applied to an i.i.d. sequence with negative mean was initially studied in [7] and [12], where it was shown that an analogue of Theorem 1.1 in dimension 1 with marked empirical process replaced by an i.i.d. sequence of random variables holds. This result was extended from i.i.d. to Markov-dependent sequences in [14], a version for Lévy processes was obtained in [8]. The one-dimensional scan statistic with variable window size appears in a variety of settings. For example, it may be interpreted as the statistic used in the CUSUM stopping procedure in change-point analysis, as the maximal waiting time among the first  $n$  customers in a GI/G/1 queue, as the maximum of a random walk reflected at its minimum, or, in bioinformatics, as the maximal segmental score when comparing two random sequences.

The papers cited above use fluctuation theory of random walks and Lévy processes. Fluctuation theory, giving very elegant solutions in dimension  $d = 1$ , does not allow an extension to the case in which  $d \geq 2$ . To prove Theorem 1.1, we use a completely different method introduced by Pickands [18], [19]; see [2], [16, Chapter 12], and [20] for further development of the theory, and [3] for the closely related Poisson clumping heuristic. Although Pickands' method was used originally to study extremes of Gaussian processes, it can be applied in the non-Gaussian case as well; see, e.g. [2] and [21].

A question closely related to that considered in Theorem 1.1 concerns the distribution of  $\sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R)$ , where  $\mathcal{Z}$  is an independently scattered homogeneous Lévy random measure on  $\mathbb{R}^d$  with negative mean and  $\mathcal{R}(n)$  is the collection of all cubes contained in  $[0, n]^d$ . The analogue of Theorem 1.1 in this situation, Theorem 2.1, will be stated in Section 2. In fact, it will be more convenient for us to prove Theorem 2.1 first and then to deduce Theorem 1.1 from it using a close relation between the empirical process and the Poisson process.

It will be seen in the proof of Theorem 2.1 that the main contribution to the extremes of the scan statistic is made by cubes of some 'optimal' volume  $v_n \approx c^* \log n$  for some constant  $c^*$ , as well as by cubes having a volume differing from the optimal volume by a quantity of order  $\sqrt{v_n}$ . Thus, the situation we encounter is close to that of [11], where a scan statistic applied to a fractional Brownian noise with negative mean was considered. Using a change of variables, the problem in [11] was reduced to studying extremes of a Gaussian field with nonconstant variance, the points of maximal variance corresponding to the intervals of 'optimal' size. In our case, random fields under consideration are non-Gaussian, which makes many results from the extreme-value theory of Gaussian processes inapplicable and causes some technical difficulties.

In Theorem 1.1 and Theorem 2.1, below, the distribution of the scan statistic applied to noises with *negative* mean is considered. We may ask: what happens if the mean is 0 or positive? Compared to the negative mean case, these two cases, which will be treated in Section 3, are much simpler. See Section 4 of [12] for the similar problem in the case of i.i.d. sequences and [23] for the case of fractional Brownian noise.

It seems that Theorems 1.1 and 2.1 remain true, with a different constant  $H^*$ , if, instead of cubes, the family  $\{\lambda B + x : \lambda > 0, x \in \mathbb{R}^d\}$  is used as a set of scanning windows, where  $B$  is any  $d$ -dimensional body with a piecewise-smooth boundary. However, working out the details is rather difficult and is not done here.

Finally, let us note that although we are considering only the scan statistic with *variable* window size, the same method, with considerable simplifications, can be used to obtain Erdős–Rényi-type laws in distribution for the scan statistic taken over all windows of *fixed* volume  $c \log n$  ( $c \log n/n$  in the case of the marked empirical process). The corresponding result in the case of a one-dimensional i.i.d. sequence was proved in [15] and [21]. In the case of a  $d$ -dimensional compound Poisson process, this can be deduced from [6], where large deviations estimates are proved for the scan statistic taken over a set of windows with fixed shape and size (the windows need not be cubes). Using such a statistic in applications requires a preknowledge about the size of the clusters to be discovered.

The organization of the paper is as follows. In Section 2 we state Theorem 2.1, an analogue of Theorem 1.1 for the scan statistic applied to a Lévy noise. The limiting distribution of the scan statistic in the case of zero or positive mean is considered in Section 3. The proof of Theorem 2.1 will be carried out in Section 4. Finally, in Section 5 we deduce Theorem 1.1 from Theorem 2.1.

Scan statistics of multidimensional i.i.d. arrays with negative mean were studied by Jiang [13]. Our Theorem 1.1 solves an open problem stated at the end of Jiang's paper.

### 2. The scan statistic of a Lévy noise with negative mean

Let  $\{\xi(t), t \geq 0\}$  be a Lévy process. An independently scattered homogeneous Lévy random measure on  $\mathbb{R}^d$  (Lévy noise for short) is a stochastic process  $\{\mathcal{Z}(R), R \in \mathcal{B}(\mathbb{R}^d)\}$ , indexed by the collection  $\mathcal{B}(\mathbb{R}^d)$  of Borel sets in  $\mathbb{R}^d$ , such that the following conditions are satisfied.

- (Z1)  $\mathcal{Z}(R)$  has the same distribution as  $\xi(|R|)$ , where  $|R|$  is the Lebesgue measure of a Borel set  $R$ .
- (Z2) If  $R_1, \dots, R_n$  are disjoint Borel subsets of  $\mathbb{R}^d$  then  $\mathcal{Z}(R_1), \dots, \mathcal{Z}(R_n)$  are independent and  $\mathcal{Z}(\bigcup_{i=1}^n R_i) = \sum_{i=1}^n \mathcal{Z}(R_i)$ .

The Lévy sheet  $\{\Xi(x_1, \dots, x_d), (x_1, \dots, x_d) \in [0, \infty)^d\}$  associated to the Lévy noise  $\mathcal{Z}$  is defined by

$$\Xi(x_1, \dots, x_d) = \mathcal{Z}([0, x_1] \times \dots \times [0, x_d]).$$

By [1] we may always assume that the sample paths of  $\Xi$  belong to the Skorokhod space in  $d$  dimensions, as defined in [1] and [5]. Concerning the underlying Lévy process  $\xi$ , we suppose that the following three conditions are satisfied.

- (L1) The logarithmic moment generating function of  $\xi(1)$ ,  $\varphi(\theta) = \log E[\exp\{\theta\xi(1)\}]$ , exists as long as  $\theta \in [0, \theta_0)$  for some  $\theta_0 \in (0, \infty]$  which is supposed to be maximal with this property.
- (L2) The function  $\varphi$  has a zero  $\theta^* \in (0, \theta_0)$ .
- (L3) The distribution of  $\xi(1)$  is nonlattice.

Let  $\mathcal{R}(n)$  be the collection of all cubes contained in  $[0, n]^d$ .

**Theorem 2.1.** *Let  $\{\mathcal{Z}(R), R \in \mathcal{B}(\mathbb{R}^d)\}$  be a Lévy noise on  $\mathbb{R}^d$ , as defined above, such that conditions (L1)–(L3) are satisfied. Then*

$$\lim_{n \rightarrow \infty} P \left[ \sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R) \leq \frac{1}{\theta^*} (d \log n + (d-1) \log \log n + (d-1) \log d + H^* + \tau) \right] = \exp\{-e^{-\tau}\}$$

for every  $\tau \in \mathbb{R}$ , where  $H^*$  is a constant to be specified later.

For  $d = 1$ , this theorem was proved in [8] by a method which uses fluctuation theory of Lévy processes and, thus, cannot be extended to higher dimensions.

### 3. Results in the case of zero and positive mean

In the preceding sections we considered the limiting distribution of the scan statistics of Lévy noises and marked empirical processes assuming, essentially, that the mean is negative. The case of negative mean is the most difficult case. Here, we state the corresponding results in the case of zero and positive mean, restricting ourselves, for simplicity, to Lévy noises.

First we consider the case of zero mean. Let  $\{\mathcal{W}(R), R \in \mathcal{B}([0, 1]^d)\}$  be a standard Gaussian white noise on  $[0, 1]^d$ ; we suppose that the corresponding Brownian sheet has continuous sample paths. Recall that  $\mathcal{R}(n)$  is the collection of all cubes contained in  $[0, n]^d$ . For a Borel set  $R$ , let  $|R|$  be its Lebesgue measure.

**Theorem 3.1.** *Let  $\{\mathcal{Z}(R), R \in \mathcal{B}(\mathbb{R}^d)\}$  be a Lévy noise such that*

$$E[\mathcal{Z}(R)] = 0, \quad \text{var } \mathcal{Z}(R) = \sigma^2 |R|, \quad \text{for each Borel set } R \subset \mathbb{R}^d.$$

Then the distribution of  $\sigma^{-1}n^{-d/2} \sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R)$  converges as  $n \rightarrow \infty$  to the distribution of  $\sup_{R \in \mathcal{R}(1)} \mathcal{W}(R)$ .

*Proof.* The Lévy sheet  $\Xi_{\mathcal{Z}}$  corresponding to the noise  $\mathcal{Z}$  and the Brownian sheet  $\Xi_{\mathcal{W}}$  corresponding to the noise  $\mathcal{W}$  are respectively defined by

$$\begin{aligned} \Xi_{\mathcal{Z}}(x_1, \dots, x_d) &= \mathcal{Z}([0, x_1] \times \dots \times [0, x_d]), \\ \Xi_{\mathcal{W}}(x_1, \dots, x_d) &= \mathcal{W}([0, x_1] \times \dots \times [0, x_d]). \end{aligned}$$

By the invariance principle for multidimensionally indexed random fields (see, e.g. [5]),

$$\sigma^{-1}n^{-d/2} \Xi_{\mathcal{Z}}(n \cdot) \Rightarrow \Xi_{\mathcal{W}}(\cdot) \quad \text{as } n \rightarrow \infty, \tag{3.1}$$

where ‘ $\Rightarrow$ ’ denotes the weak convergence in the Skorokhod space  $D([0, 1]^d)$ . We define a continuous functional  $F: D([0, 1]^d) \rightarrow \mathbb{R}$  by

$$F(\Xi) = \sup_{R \in \mathcal{R}(1)} \Xi(R), \quad \Xi \in D([0, 1]^d),$$

where  $\Xi(R)$  is defined in a straightforward way (so that, e.g.  $\Xi_{\mathcal{Z}}(R) = \mathcal{Z}(R)$  and  $\Xi_{\mathcal{W}}(R) = \mathcal{W}(R)$ ). It follows from (3.1) that the random variable

$$F(\sigma^{-1}n^{-d/2} \Xi_{\mathcal{Z}}(n \cdot)) = \sigma^{-1}n^{-d/2} \sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R)$$

converges in distribution to  $F(\Xi_{\mathcal{W}}(\cdot)) = \sup_{R \in \mathcal{R}(1)} \mathcal{W}(R)$  as  $n \rightarrow \infty$ . This proves the theorem.

**Theorem 3.2.** Let  $\{\mathcal{Z}(R), R \in \mathcal{B}(\mathbb{R}^d)\}$  be a Lévy noise such that, for some  $\mu > 0$  and  $\sigma^2 > 0$ ,

$$E[\mathcal{Z}(R)] = \mu|R|, \quad \text{var } \mathcal{Z}(R) = \sigma^2|R|.$$

Then the distribution of  $\sigma^{-1}n^{-d/2}(\sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R) - n^d \mu)$  converges as  $n \rightarrow \infty$  to the standard normal distribution.

*Proof.* The idea is to show that  $\sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R)$  behaves essentially like  $\mathcal{Z}([0, n]^d)$  and then to apply the central limit theorem. We show that, for every  $a > 0$ ,

$$P\left[\sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R) - \mathcal{Z}([0, n]^d) \geq an^{d/2}\right] \rightarrow 0 \quad \text{as } n \rightarrow \infty. \tag{3.2}$$

Denoting the left-hand side by  $P_n$  and taking  $\varepsilon > 0$  small, we have  $P_n \leq P'_n + P''_n$ , where

$$\begin{aligned} P'_n &= P\left[\sup_{R \in \mathcal{R}'(n)} \mathcal{Z}(R) - \mathcal{Z}([0, n]^d) \geq 0\right], \\ P''_n &= P\left[\sup_{R \in \mathcal{R}''(n)} \mathcal{Z}(R) - \mathcal{Z}([0, n]^d) \geq an^{d/2}\right], \end{aligned}$$

and  $\mathcal{R}'(n) = \{R \in \mathcal{R}(n): |R| < (1 - \varepsilon)n^d\}$  and  $\mathcal{R}''(n) = \{R \in \mathcal{R}(n): |R| \geq (1 - \varepsilon)n^d\}$ .

Define the centered noise,  $\{\mathcal{Z}_0(R), R \in \mathcal{B}(\mathbb{R}^d)\}$ , by  $\mathcal{Z}_0(R) = \mathcal{Z}(R) - \mu|R|$ . Then

$$P'_n \leq P\left[\sup_{R \in \mathcal{R}'(n)} \mathcal{Z}_0(R) - \mathcal{Z}_0([0, n]^d) \geq \mu\varepsilon n^d\right].$$

By the multidimensional invariance principle of [5], applied to  $\mathcal{Z}_0$ ,  $\lim_{n \rightarrow \infty} P'_n = 0$ . Furthermore,

$$P''_n \leq \mathbb{P} \left[ \sup_{R \in \mathcal{R}''(n)} \mathcal{Z}_0(R) - \mathcal{Z}_0([0, n]^d) \geq an^{d/2} \right].$$

Again, using the multidimensional invariance principle, we see that this converges to

$$c(\varepsilon) = \mathbb{P} \left[ \sup_{R \in \mathcal{R}(1), |R| \geq 1-\varepsilon} \mathcal{W}(R) - \mathcal{W}([0, 1]^d) \geq a \right].$$

It is easy to see that  $\lim_{\varepsilon \rightarrow 0} c(\varepsilon) = 0$ . It follows that

$$\limsup_{n \rightarrow \infty} P_n \leq \limsup_{n \rightarrow \infty} P'_n + \limsup_{n \rightarrow \infty} P''_n \leq c(\varepsilon).$$

Letting  $\varepsilon \rightarrow 0$  we obtain  $\lim_{n \rightarrow \infty} P_n = 0$ , which proves (3.2). Now, the statement of the theorem follows from the central limit theorem applied to  $\mathcal{Z}([0, n]^d)$  in combination with (3.2).

### 4. Proof of Theorem 2.1

This section is devoted to the proof of Theorem 2.1. In the first two subsections we introduce some notation and prove technical lemmas which will be used often in the sequel. In what follows,  $C > 0$  and  $\delta > 0$  denote large and small constants, respectively, whose values may change from line to line.

#### 4.1. Notation and preliminaries

4.1.1. *Normalizing constants.* Let  $\tau \in \mathbb{R}$  be fixed once and for all. For the constants  $H^*$  and  $\alpha^*$  to be specified later, define

$$u_n = \frac{1}{\theta^*} (d \log n + (d - 1) \log \log n + (d - 1) \log d + H^* + \tau), \tag{4.1}$$

$$v_n = \frac{u_n}{\alpha^*}, \quad l_n = v_n^{1/d}. \tag{4.2}$$

4.1.2. *The space of cubes.* A  $d$ -dimensional cube (denoted usually by  $R$ ) is a set of the form  $\times_{i=1}^d [x_i - x/2, x_i + x/2]$ , where  $(x_1, \dots, x_d) \in \mathbb{R}^d$  are the coordinates of the center and  $x > 0$  is the side length. The space of all cubes, denoted by  $\mathcal{R}$ , will be identified with  $\mathbb{R}^d \times (0, \infty)$ , a cube  $R$  being identified with the tuple  $(x_1, \dots, x_d; x)$ . We denote by  $|R| = x^d$  the volume of the cube  $R$ .

4.1.3. *The underlying Lévy process.* Let  $\{\xi(t), t \geq 0\}$  be a Lévy process satisfying conditions (L1)–(L3) of Section 2. The function  $\varphi$  is a real analytic, convex function on  $(0, \theta_0)$ . The zero  $\theta^*$  is necessarily unique by convexity of  $\varphi$ . It follows from condition (L2) that  $E[\xi(1)] < 0$ . Furthermore, condition (L2) implies that  $\mathbb{P}[\xi(1) > 0] \neq 0$ , and it follows that  $\xi(1)$ , being infinitely divisible, can attain arbitrarily large values. Using this, it is not difficult to show that  $\lim_{\theta \rightarrow \theta_0} \varphi'(\theta) = \infty$ . Note also that  $\varphi'$  is monotone increasing and that  $\varphi'(0) = E[\xi(1)] < 0$ . For each  $\alpha \in (E[\xi(1)], \infty)$ , let  $\theta(\alpha)$  be the unique solution of  $\varphi'(\theta(\alpha)) = \alpha$  and let  $\sigma(\alpha) = \sqrt{\varphi''(\theta(\alpha))}$ . Define the Cramér–Chernoff information function  $I: [E[\xi(1)], \infty) \rightarrow [0, \infty)$  by

$$I(\alpha) = \sup_{\theta \geq 0} (\alpha\theta - \varphi(\theta)) = \alpha\theta(\alpha) - \varphi(\theta(\alpha)). \tag{4.3}$$

Define

$$\alpha^* = \varphi'(\theta^*) \quad \text{and} \quad \sigma^* = \sqrt{\varphi''(\theta^*)}.$$

Note that  $\alpha^* > 0$ ,  $\theta(\alpha^*) = \theta^*$ , and  $\sigma(\alpha^*) = \sigma^*$ .

**Lemma 4.1.** *The function  $J : (0, \infty) \rightarrow (0, \infty)$  defined by  $J(\alpha) = I(\alpha)/\alpha$  has a unique minimum at  $\alpha = \alpha^*$ . Furthermore,*

$$J(\alpha^*) = \theta^*, \quad J'(\alpha^*) = 0, \quad \text{and} \quad J''(\alpha^*) = \frac{1}{\alpha^* \sigma^{*2}}. \tag{4.4}$$

*Proof.* Substituting  $\alpha = \alpha^*$  into (4.3) gives  $I(\alpha^*) = \alpha^* \theta^* - \varphi(\theta^*) = \alpha^* \theta^*$ , which proves that  $J(\alpha^*) = \theta^*$ . Differentiating (4.3) at  $\alpha = \alpha^*$  we obtain  $I'(\alpha^*) = \theta(\alpha^*) = \theta^*$  and  $I''(\alpha^*) = \theta'(\alpha^*) = 1/\varphi''(\theta(\alpha^*)) = 1/\sigma^{*2}$ . Substituting this into

$$J'(\alpha) = \alpha^{-2}(\alpha I'(\alpha) - I(\alpha)) \quad \text{and} \quad J''(\alpha) = \alpha^{-3}(\alpha^2 I''(\alpha) - 2\alpha I'(\alpha) + 2I(\alpha)),$$

we obtain (4.4). In order to show that  $\alpha = \alpha^*$  is the *unique* minimum of  $J$ , note that it follows from the above equation that  $J'(\alpha) = \alpha^{-2}\varphi(\theta(\alpha))$  and that  $\alpha = \alpha^*$  is the unique solution of  $\varphi(\theta(\alpha)) = 0$ ,  $\alpha > 0$ .

4.1.4. *Large deviations.* We need the following precise large deviations theorem due to Petrov [17].

**Theorem 4.1.** ([17].) *Let  $\{\xi(t), t \geq 0\}$  be a Lévy process satisfying conditions (L1) and (L3) of Section 2. Let  $\alpha \in (E[\xi(1)], \infty)$ . We have, as  $v \rightarrow \infty$ ,*

$$P\left[\frac{\xi(v)}{v} > \alpha\right] \sim \frac{1}{\sqrt{2\pi\theta(\alpha)\sigma(\alpha)}} \frac{1}{\sqrt{v}} e^{-vI(\alpha)}.$$

*Moreover, the above holds uniformly in  $\alpha$  as long as  $\alpha$  stays bounded away from  $E[\xi(1)]$  and  $+\infty$ .*

The next lemma is a simple consequence of Markov’s inequality and will be used often in the sequel.

**Lemma 4.2.** *For every  $u, v > 0$ ,*

$$P[\xi(v) > u] \leq \exp\left\{-uJ\left(\frac{u}{v}\right)\right\}.$$

*Proof.* By Markov’s inequality we have, for every  $\theta > 0$ ,

$$P[\xi(v) > u] \leq e^{-u\theta} E[e^{\theta\xi(v)}] = \exp\{-u\theta + v\varphi(\theta)\} = \exp\left\{-v\left(\frac{u\theta}{v} - \varphi(\theta)\right)\right\}.$$

Since the above is true for every  $\theta > 0$ , we obtain

$$P[\xi(v) > u] \leq \exp\left\{-vI\left(\frac{u}{v}\right)\right\} = \exp\left\{-uJ\left(\frac{u}{v}\right)\right\},$$

which completes the proof.

**Corollary 4.1.** *For every  $u, v > 0$ ,*

$$P[\xi(v) > u] \leq \exp\{-\theta^*u\}.$$

*Proof.* Use Lemma 4.2 and recall that  $J(u/v) \geq \theta^*$  by Lemma 4.1.

**4.2. Modulus of the continuity estimate**

Let  $\{\mathcal{Z}(R), R \in \mathcal{B}(\mathbb{R}^d)\}$  be a Lévy noise such that the underlying Lévy process  $\xi$  satisfies conditions (L1)–(L3) of Section 2, and let  $\Xi$  be the corresponding Lévy sheet. The next lemma gives a large deviations estimate for the supremum of  $\Xi$  over  $[0, c]^d, c > 0$ .

**Lemma 4.3.** *For every  $\theta < \theta_0$  and  $c > 0$ , there is a  $C = C(\theta, c)$  such that*

$$P\left[\sup_{(x_1, \dots, x_d) \in [0, c]^d} \Xi(x_1, \dots, x_d) > u\right] < Ce^{-\theta u} \text{ for all } u > 0. \tag{4.5}$$

*Proof.* For simplicity, we assume that  $c = 1$ . For  $d = 1$ , the lemma was proved in [22, Equation (2.1)]. We use induction over  $d$  combined with the method of [22]. Suppose that the statement of the lemma was proved in dimensions  $1, \dots, d - 1$ . Let

$$\tau_1 = \inf\{x_1 \geq 0 : \text{there exists } (x_2, \dots, x_d) \in [0, 1]^{d-1} \text{ such that } \Xi(x_1, \dots, x_d) > u\}.$$

The left-hand side of (4.5) is the probability of the event  $A = \{\tau_1 \leq 1\}$ . We have  $A = A_1 \cup A_2$ , where

$$A_1 = \left\{ \tau_1 \leq 1 \cap \sup_{(x_2, \dots, x_d) \in [0, 1]^{d-1}} \Xi(1, x_2, \dots, x_d) > u - 1 \right\},$$

$$A_2 = \left\{ \tau_1 \leq 1 \cap \sup_{(x_2, \dots, x_d) \in [0, 1]^{d-1}} \Xi(1, x_2, \dots, x_d) \leq u - 1 \right\}.$$

Now, by the induction hypothesis,

$$P[A_1] \leq P\left[\sup_{(x_2, \dots, x_d) \in [0, 1]^{d-1}} \Xi(1, x_2, \dots, x_d) > u - 1\right] \leq Ce^{-\theta u}.$$

We estimate

$$\begin{aligned} P[A_2] &\leq P\left[A \cap \inf_{(x_2, \dots, x_d) \in [0, 1]^{d-1}} (\Xi(1, x_2, \dots, x_d) - \Xi(\tau_1, x_2, \dots, x_d)) < -1\right] \\ &\leq P\left[A \cap \inf_{(x_1, x_2, \dots, x_d) \in [0, 1]^d} (\Xi(\tau_1 + x_1, x_2, \dots, x_d) - \Xi(\tau_1, x_2, \dots, x_d)) < -1\right] \\ &= P[A] P\left[\inf_{(x_1, x_2, \dots, x_d) \in [0, 1]^d} (\Xi(\tau_1 + x_1, x_2, \dots, x_d) - \Xi(\tau_1, x_2, \dots, x_d)) < -1\right] \\ &= P[A] P\left[\inf_{(x_1, x_2, \dots, x_d) \in [0, 1]^d} \Xi(x_1, x_2, \dots, x_d) < -1\right] \\ &= p P[A] \text{ for some } p < 1. \end{aligned}$$

We obtain  $P[A] = P[A_1] + P[A_2] \leq Ce^{-\theta u} + p P[A]$  for some  $p < 1$ , from which the statement of the lemma follows.

In the sequel, we will often use the following technical lemma which estimates the continuity modulus of the random field  $\{\mathcal{Z}(R), R \in \mathcal{R}\}$ .

**Lemma 4.4.** *Let  $c > 0$  be a fixed constant. Let  $x > c^d$  and  $q < cx^{1-d}$ . Define a set of cubes  $\mathcal{B} = [-q/2, q/2]^d \times [x, x + q]$ . Let  $R_0 = [-(x - q)/2, (x - q)/2]^d$  be the intersection of all cubes from  $\mathcal{B}$ , and define the random variable  $M$  by*

$$M = \sup_{R \in \mathcal{B}} \mathcal{Z}(R) - \mathcal{Z}(R_0).$$



Then, for every  $\theta < \theta_0$ , there is a constant  $C = C(c, \theta)$  such that, uniformly in  $x$  and  $q$ ,

$$P[M > t] \leq Ce^{-\theta t} \quad \text{for all } t > 0.$$

*Proof.* We show that  $E[e^{\theta M}] < C(c, \theta)$ ; the lemma then follows from Markov’s inequality. For  $h = (h_1, \dots, h_d)$ ,  $h_i > 0$ , and  $\varepsilon_1, \dots, \varepsilon_d \in \{-1, 0, 1\}$ , define a rectangle

$$R(\varepsilon_1, \dots, \varepsilon_d; h) = I(\varepsilon_1; h_1) \times \dots \times I(\varepsilon_d; h_d),$$

where

$$I(\varepsilon_i; h_i) = \begin{cases} \left[ -\frac{x-q}{2}, \frac{x-q}{2} \right] & \text{if } \varepsilon_i = 0, \\ \left[ \frac{x-q}{2}, \frac{x-q}{2} + h_i \right] & \text{if } \varepsilon_i = 1, \\ \left[ -\frac{x-q}{2} - h_i, -\frac{x-q}{2} \right] & \text{if } \varepsilon_i = -1. \end{cases}$$

Note that  $R(0, \dots, 0; h) = R_0$  (in particular, the left-hand side does not depend on  $h$ ). Let

$$M(\varepsilon_1, \dots, \varepsilon_d) = \sup_{h \in [0, 3q/2]^d} \mathcal{Z}(R(\varepsilon_1, \dots, \varepsilon_d; h)).$$

The random variables  $M(\varepsilon_1, \dots, \varepsilon_d)$  are independent and

$$M \leq \sum_{(\varepsilon_1, \dots, \varepsilon_d) \in \{-1, 0, 1\}^d \setminus (0, \dots, 0)} M(\varepsilon_1, \dots, \varepsilon_d). \tag{4.6}$$

Furthermore, if  $r$  is the number of +1s and -1s among  $\varepsilon_i$  and if  $r \neq 0$ , then the random variable  $M(\varepsilon_1, \dots, \varepsilon_d)$  has the same distribution as the supremum of an  $r$ -dimensionally indexed Lévy sheet on  $[0, \frac{3}{2}q(x-q)^{(d-r)/r}]^r$ . Since  $\frac{3}{2}q(x-q)^{(d-r)/r} \leq \frac{3}{2}qx^{d-1} < \frac{3}{2}c$  by the assumption of the lemma, we have, by Lemma 4.3,

$$E[\exp\{\theta M(\varepsilon_1, \dots, \varepsilon_d)\}] < C(c, \theta).$$

To complete the proof of the lemma, use (4.6).

### 4.3. Cubes of nearly optimal size

4.3.1. *Idea of the proof of Theorem 2.1.* Now we are ready to start the proof of Theorem 2.1. We are interested in the high-crossing probability  $P[\sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R) > u_n]$ . Intuitively, too small or too large cubes have asymptotically no chance to contribute to the above probability (for large cubes, this is due to the assumption that the mean of the Lévy noise is negative). We will see later that, asymptotically, the probability  $P[\mathcal{Z}(R) > u_n]$  achieves its maximum if the volume of cube  $R$  is equal to  $v_n$  (equivalently, if its side length is equal to  $l_n$ ). Furthermore, we will see that cubes of volume differing from the optimal volume,  $v_n$ , by a quantity of order more than  $\sqrt{v_n}$  have asymptotically no chance to contribute to the extremes of the field  $\mathcal{Z}$ .

In this subsection we are dealing with cubes of nearly optimal size, that is with cubes whose volume differs from  $v_n$  by a quantity of order  $\sqrt{v_n}$ . To be more precise, we fix a very large  $A > 0$  and define

$$l_n^- = (v_n - A\sqrt{v_n})^{1/d}, \quad l_n^+ = (v_n + A\sqrt{v_n})^{1/d}. \tag{4.7}$$

The main result of this subsection is Lemma 4.11, below, in which the limit, as  $n \rightarrow \infty$ , of  $P[\sup_{R \in \mathcal{R}_A(n)} \mathcal{Z}(R) \leq u_n]$  is calculated, where  $\mathcal{R}_A(n)$  is the set of cubes from  $\mathcal{R}(n)$  whose side lengths are in the interval  $[l_n^-, l_n^+]$ .

4.3.2. *Cubes of nearly optimal size.* First we evaluate the high crossing probability  $P[\mathcal{Z}(R) > u_n]$  for cubes  $R$  having the optimal volume,  $v_n$ .

**Lemma 4.5.** *We have, as  $n \rightarrow \infty$ ,*

$$P[\xi(v_n) > u_n] \sim \frac{\sqrt{\alpha^*}}{\sqrt{2\pi\theta^*\sigma^*}} \frac{1}{\sqrt{u_n}} \exp\{-\theta^*u_n\}.$$

*Proof.* The proof follows from Theorem 4.1.

Now we consider cubes with volume differing from the optimal volume by a quantity of order  $\sqrt{v_n}$ . Comparably to cubes of optimal volume, the high crossing probability changes by a constant factor.

**Lemma 4.6.** *We have, as  $n \rightarrow \infty$ ,*

$$P[\xi(v_n + s\sqrt{v_n}) > u_n + t] \sim e^{-\theta^*t} \exp\left\{-\frac{(\alpha^*s)^2}{2\sigma^{*2}}\right\} P[\xi(v_n) > u_n].$$

*The above holds uniformly in  $s$  and  $t$  as long as  $s = O(1)$  and  $t = o(\sqrt{u_n})$ .*

*Proof.* Let  $\alpha_n = (u_n + t)/(v_n + s\sqrt{v_n})$ . Note that  $\lim_{n \rightarrow \infty} \alpha_n = \alpha^*$ . We obtain, by Theorem 4.1,

$$P[\xi(v_n + s\sqrt{v_n}) > (u_n + t)] \sim \frac{1}{\sqrt{2\pi\theta^*\sigma^*}} \frac{1}{\sqrt{v_n}} \exp\{-(u_n + t)J(\alpha_n)\}.$$

Now, an easy calculation shows that  $\alpha_n = \alpha^*(1 - s/\sqrt{v_n} + o(1/\sqrt{v_n}))$  as  $n \rightarrow \infty$ . Using Lemma 4.1, we obtain

$$J(\alpha_n) = \theta^* + \frac{1}{2} \frac{\alpha^*s^2}{\sigma^{*2}v_n} + o\left(\frac{1}{v_n}\right).$$

It follows that

$$P[\xi(v_n + s\sqrt{v_n}) > (u_n + t)] \sim \left(\frac{1}{\sqrt{2\pi\theta^*\sigma^*}} \frac{1}{\sqrt{v_n}} \exp\{-\theta^*u_n\}\right) e^{-\theta^*t} \exp\left\{-\frac{(\alpha^*s)^2}{2\sigma^{*2}}\right\}.$$

The statement of the lemma follows by noting that the first factor on the right-hand side is asymptotically equivalent to  $P[\xi(v_n) > u_n]$  by Lemma 4.5.

Let  $q_n = l_n^{1-d} = v_n^{(1-d)/d}$ . Note that if  $d = 1$ ,  $q_n = 1$ , whereas otherwise  $\lim_{n \rightarrow \infty} q_n = 0$ . In the next lemma we consider a high-crossing probability over a set having a size of order  $q_n$  in the space of cubes.

**Lemma 4.7.** *For  $x > 0$  and a fixed  $m \in \mathbb{N}$ , define a set of cubes*

$$\mathcal{B}_x^m(n) = \left[-\frac{mq_n}{2}, \frac{mq_n}{2}\right]^d \times [x, x + mq_n].$$

*Then, for some constant  $H_m > 0$ , the following asymptotic equality holds as  $n \rightarrow \infty$  uniformly in  $x$  as long as  $x \in [l_n^-, l_n^+]$ :*

$$P\left[\sup_{R \in \mathcal{B}_x^m(n)} \mathcal{Z}(R) > u_n\right] \sim H_m P[\xi(x^d) > u_n]. \tag{4.8}$$

*Proof.* Let  $R_0 = [-(x - mq_n)/2, (x - mq_n)/2]^d$  be the intersection of all cubes from  $\mathcal{B}_x^m(n)$ . Note that  $|R_0| = |x - mq_n|^d = x^d + O(1)$  as  $n \rightarrow \infty$ . Applying Lemma 4.6 twice, we obtain

$$P[\mathcal{Z}(R_0) > u_n - t] \sim e^{\theta^* t} P[\xi(x^d) > u_n]. \tag{4.9}$$

Let

$$M_n = \sup_{R \in \mathcal{B}_x^m(n)} \mathcal{Z}(R) - \mathcal{Z}(R_0).$$

Then it is easy to see that  $\mathcal{Z}(R_0)$  and  $M_n$  are independent and that  $M_n$  converges in distribution as  $n \rightarrow \infty$  to the random variable

$$M_\infty = \sup_{(l_1, \dots, l_d; l) \in [-m/2, m/2]^d \times [0, m/2]} \sum_{i=1}^d \left( \xi_i \left( l_i + l + \frac{1}{2} \right) - \xi_i \left( l_i - l - \frac{1}{2} \right) \right),$$

where  $\xi_1(\cdot), \dots, \xi_d(\cdot)$  are independent copies of the Lévy process  $\{\xi(t), t \in \mathbb{R}\}$ . Denote the probability on the left-hand side of (4.8) by  $P_n$ . Then

$$P_n = \int_{-\infty}^{\infty} P[\mathcal{Z}(R_0) > u_n - t] dP[M_n = t].$$

Using (4.9), we obtain, at least formally, as  $n \rightarrow \infty$ ,

$$P_n \sim \left( \int_{-\infty}^{\infty} e^{\theta^* t} dP[M_\infty = t] \right) P[\xi(x^d) > u_n],$$

which proves the lemma with  $H_m = E[\exp\{\theta^* M_\infty\}]$ . In the rest of the proof we justify this step. Take  $T > 0$  large. We have

$$\begin{aligned} \frac{P_n}{P[\xi(x^d) > u_n]} &= \int_{-\infty}^{\infty} \frac{P[\mathcal{Z}(R_0) > u_n - t]}{P[\xi(x^d) > u_n]} dP[M_n = t] \\ &= \int_{-T}^T \frac{P[\mathcal{Z}(R_0) > u_n - t]}{P[\xi(x^d) > u_n]} dP[M_n = t] \\ &\quad + \int_T^{u_n^{1/3}} \frac{P[\mathcal{Z}(R_0) > u_n - t]}{P[\xi(x^d) > u_n]} dP[M_n = t] \\ &\quad + \int_{u_n^{1/3}}^{\infty} \frac{P[\mathcal{Z}(R_0) > u_n - t]}{P[\xi(x^d) > u_n]} dP[M_n = t] \\ &\quad + \int_{-\infty}^{-T} \frac{P[\mathcal{Z}(R_0) > u_n - t]}{P[\xi(x^d) > u_n]} dP[M_n = t] \\ &= I + II + III + IV. \end{aligned}$$

Since, by Lemma 4.6, the convergence in (4.9) is uniform for  $t \in [-T, T]$  and since  $M_n$  converges in distribution to  $M_\infty$ , we obtain

$$\lim_{n \rightarrow \infty} I = \int_{-T}^T e^{\theta^* t} dP[M_\infty = t].$$

The convergence in (4.9) remains uniform for  $t = o(\sqrt{u_n})$ . Using the fact that, by Lemma 4.4, applied to  $M_n$ ,  $E[\exp\{\theta M_n\}] < C(\theta, m)$  for every  $\theta < \theta_0$ , we obtain

$$II \leq \int_T^{u_n^{1/3}} C e^{\theta^* t} dP[M_n = t] < C e^{-\delta T} \quad \text{for some } \delta > 0.$$

To estimate the third term, note that  $P[Z(R_0) > u_n - t] \leq \exp\{-\theta^*(u_n - t)\}$  by Corollary 4.1 and that  $P[\xi(x^d) > u_n] \geq cu_n^{-1/2} \exp\{-\theta^*u_n\}$  by Lemma 4.6. Thus,

$$III \leq \int_{u_n^{1/3}}^\infty \frac{\exp\{-\theta^*(u_n - t)\}}{cu_n^{-1/2} \exp\{-\theta^*u_n\}} dP[M_n = t] \leq Cu_n^{1/2} \int_{u_n^{1/3}}^\infty e^{\theta^*t} dP[M_n = t].$$

The right-hand side of the above inequality converges to 0 as  $n \rightarrow \infty$  since  $E[\exp\{\theta M_n\}] < C$  for every  $\theta < \theta_0$  by Lemma 4.4. Thus,  $\lim_{n \rightarrow \infty} III = 0$ .

We estimate the last term:

$$IV \leq \frac{P[Z(R_0) > u_n + T]}{P[\xi(x^d) > u_n]}.$$

It follows from Lemma 4.6 applied twice that the right-hand side of the above inequality converges to  $e^{-\theta^*T}$  as  $n \rightarrow \infty$  and, hence,  $\limsup_{n \rightarrow \infty} IV \leq e^{-\theta^*T}$ . The statement of the lemma follows from above by letting  $n \rightarrow \infty$  first and then  $T \rightarrow \infty$ .

Let  $s_n$  be a sequence satisfying  $s_n = O(l_n)$ ,  $s_n > 1$ . In the next lemma we evaluate the high-crossing probability of the scan statistic taken over the set of all cubes of nearly optimal volume with centers contained in  $[0, s_n]^d$ .

**Lemma 4.8.** *Let  $\mathcal{L}_A(n) = [0, s_n]^d \times [l_n^-, l_n^+]$ . Then, as  $n \rightarrow \infty$ ,*

$$P\left[\sup_{R \in \mathcal{L}_A(n)} Z(R) > u_n\right] \sim H \cdot \left(\int_{-A}^A \exp\left\{-\frac{(\alpha^*t)^2}{2\sigma^{*2}}\right\} dt\right) s_n^d v_n^{d-1/2} P[\xi(v_n) > u_n]. \tag{4.10}$$

Here,  $H \in (0, \infty)$  is a constant defined by

$$H = \frac{1}{d} \lim_{m \rightarrow \infty} \frac{H_m}{m^{d+1}}. \tag{4.11}$$

*Proof.* Define

$$\mathcal{L}_A^m(n) = mq_n \mathbb{Z}^{d+1} \cap \mathcal{L}_A(n).$$

For  $R = (x_1, \dots, x_d; x) \in \mathcal{L}_A^m(n)$ , let

$$\mathcal{B}_R^m(n) = \left[x_1 - \frac{q_n}{2}, x_1 + \left(m - \frac{1}{2}\right)q_n\right] \times \dots \times \left[x_d - \frac{q_n}{2}, x_d + \left(m - \frac{1}{2}\right)q_n\right] \times [x, x + mq_n],$$

and define  $B_R^m(n) = B_{x_1, \dots, x_d; x}^m(n)$  to be the random event  $\{\sup_{Q \in \mathcal{B}_R^m(n)} Z(Q) > u_n\}$ . Denote the probability on the left-hand side of (4.10) by  $P_n$ . Then

$$P_n \leq S_1^m(n), \tag{4.12}$$

where

$$S_1^m(n) = \sum_{R \in \mathcal{L}_A^m(n)} P[B_R^m(n)] = \left(\frac{s_n}{mq_n}\right)^d \sum_{x \in mq_n \mathbb{Z} \cap [l_n^-, l_n^+]} P[B_{0, \dots, 0; x}^m(n)].$$

Applying Lemma 4.7 and then Lemma 4.6 yields

$$P[B_{0, \dots, 0; x}^m(n)] \sim H_m P[\xi(v_n) > u_n] \exp\left\{-\frac{\alpha^{*2}}{2\sigma^{*2}} \frac{(x^d - v_n)^2}{v_n}\right\}.$$

If the values of  $x$  are in  $mq_n\mathbb{Z} \cap [l_n^-, l_n^+]$  then the values of  $(x^d - v_n)/\sqrt{v_n}$  form a lattice in the interval  $[-A, A]$ , whose (variable) mesh size is asymptotically equivalent to  $md/\sqrt{v_n}$ . Thus, approximating the Riemann sum by an integral, we obtain

$$\sum_{x \in mq_n\mathbb{Z} \cap [l_n^-, l_n^+]} P[B_{0,\dots,0;x}^m(n)] \sim H_m P[\xi(v_n) > u_n] \left( \frac{\sqrt{v_n}}{md} \int_{-A}^A \exp\left\{-\frac{(\alpha^*t)^2}{2\sigma^{*2}}\right\} dt \right).$$

This shows that

$$S_1^m(n) \sim \frac{1}{dm^{d+1}} H_m \cdot \left( \int_{-A}^A \exp\left\{-\frac{(\alpha^*t)^2}{2\sigma^{*2}}\right\} dt \right) s_n^d v_n^{d-1/2} P[\xi(v_n) > u_n].$$

Since the above is true for every  $m$ , we find, by letting  $m \rightarrow \infty$ , that the left-hand side of (4.10) is asymptotically not greater than the right-hand side of (4.10). In order to prove the converse, we use the Bonferroni inequality,

$$P_n \geq S_1^m(n) - S_2^m(n), \tag{4.13}$$

where  $S_1^m(n)$  is as above and

$$S_2^m(n) = \sum_{\substack{R_1, R_2 \in \mathcal{L}_A^m(n) \\ R_1 \neq R_2}} P[B_{R_1}^m(n) \cap B_{R_2}^m(n)].$$

The sum  $S_1^m(n)$  was already treated above. The proof will be completed in Lemma 4.10, below, where it will be shown that  $S_2^m(n)$  can be asymptotically ignored as  $n \rightarrow \infty$  and  $m \rightarrow \infty$ .

So, we are going to estimate  $S_2^m(n)$ . For a cube  $R = (x_1, \dots, x_d; x)$ , let

$$\mathcal{B}_R(n) = \left[ x_1 - \frac{q_n}{2}, x_1 + \frac{q_n}{2} \right] \times \dots \times \left[ x_d - \frac{q_n}{2}, x_d + \frac{q_n}{2} \right] \times [x, x + q_n]$$

and let  $B_R(n)$  be the random event  $\{\sup_{R \in \mathcal{B}_R(n)} \mathcal{Z}(R) > u_n\}$ . Let  $R_1, R_2 \in \mathcal{L}_A(n)$  be two cubes. Denote by  $\Delta = \Delta(R_1, R_2) = |R_1 \Delta R_2|$  the volume of the symmetric difference of  $R_1$  and  $R_2$ .

**Lemma 4.9.** *For arbitrary cubes  $R_1, R_2 \in \mathcal{L}_A(n)$ , we have*

$$P[B_{R_1}(n) \cap B_{R_2}(n)] \leq C \exp\{-\delta\Delta(R_1, R_2)\} P[\xi(v_n) > u_n].$$

*Proof.* For a cube  $R = (x_1, \dots, x_d; x)$ ,  $h = (h_1, \dots, h_d)$ ,  $h_i > 0$ , and  $(\varepsilon_1, \dots, \varepsilon_d) \in \{-1, 0, 1\}^d$ , let

$$R(\varepsilon_1, \dots, \varepsilon_d; h) = I_1(\varepsilon_1; h_1) \times \dots \times I_d(\varepsilon_d; h_d),$$

where

$$I_j(\varepsilon_j; h_j) = \begin{cases} \left[ x_j - \frac{x - q_n}{2}, x_j + \frac{x - q_n}{2} \right] & \text{if } \varepsilon_j = 0, \\ \left[ x_j + \frac{x - q_n}{2}, x_j + \frac{x - q_n}{2} + h_j \right] & \text{if } \varepsilon_j = 1, \\ \left[ x_j - \frac{x - q_n}{2} - h_j, x_j - \frac{x - q_n}{2} \right] & \text{if } \varepsilon_j = -1. \end{cases}$$

Let  $R_1(\varepsilon_1, \dots, \varepsilon_d; h)$  and  $R_2(\varepsilon_1, \dots, \varepsilon_d; h)$  be defined analogously with  $R$  replaced by  $R_1$  and  $R_2$ , respectively. Let  $h(n) = (\frac{3}{2}q_n, \dots, \frac{3}{2}q_n)$ , and let

$$\bar{R}_1 = \bigcup_{(\varepsilon_1, \dots, \varepsilon_d) \in \{-1, 0, 1\}^d} R_1(\varepsilon_1, \dots, \varepsilon_d; h(n)).$$

Define

$$R'_2(\varepsilon_1, \dots, \varepsilon_d; h) = R_2(\varepsilon_1, \dots, \varepsilon_d; h) \cap \bar{R}_1,$$

$$R''_2(\varepsilon_1, \dots, \varepsilon_d; h) = R_2(\varepsilon_1, \dots, \varepsilon_d; h) \setminus \bar{R}_1.$$

Note that  $R'_0 := R'_2(0, \dots, 0; h)$  and  $R''_0 := R''_2(0, \dots, 0; h)$  do not depend on  $h$ . Let

$$M' = \sum_{\{\varepsilon_1, \dots, \varepsilon_d\} \in \{-1, 0, 1\}^d \setminus \{0, \dots, 0\}} \sup_{h \in [0, 3q_n/2]^d} \mathcal{Z}(R'_2(\varepsilon_1, \dots, \varepsilon_d; h)),$$

$$M'' = \sum_{\{\varepsilon_1, \dots, \varepsilon_d\} \in \{-1, 0, 1\}^d \setminus \{0, \dots, 0\}} \sup_{h \in [0, 3q_n/2]^d} \mathcal{Z}(R''_2(\varepsilon_1, \dots, \varepsilon_d; h)).$$

Finally, fix some small  $a > 0$  and let

$$B' = \{\mathcal{Z}(R'_0) + M' > u_n + a\Delta\}, \quad B'' = \{\mathcal{Z}(R''_0) + M'' > -a\Delta\}.$$

We trivially have

$$\begin{aligned} \mathbb{P}[B_{R_1}(n) \cap B_{R_2}(n)] &= \mathbb{P}[B_{R_1}(n) \cap B_{R_2}(n) \cap B'] + \mathbb{P}[(B_{R_1}(n) \cap B_{R_2}(n)) \setminus B'] \\ &\leq \mathbb{P}[B'] + \mathbb{P}[B_{R_1}(n) \cap B''] \\ &= \mathbb{P}[B'] + \mathbb{P}[B_{R_1}(n)]\mathbb{P}[B'']. \end{aligned}$$

By Lemma 4.7, with  $m = 1$ , we have  $\mathbb{P}[B_{R_1}(n)] \leq C \mathbb{P}[\xi(v_n) > u_n]$ . Thus, in order to prove the lemma, we need to show the following two inequalities:

$$\mathbb{P}[B'] \leq C \exp\{-\delta\Delta(R_1, R_2)\} \mathbb{P}[\xi(v_n) > u_n], \tag{4.14}$$

$$\mathbb{P}[B''] \leq C \exp\{-\delta\Delta(R_1, R_2)\}. \tag{4.15}$$

We prove (4.14). Take  $\varepsilon > 0$  sufficiently small. Then

$$\begin{aligned} \mathbb{P}[B'] &= \int_0^\infty \mathbb{P}[\mathcal{Z}(R'_0) \geq u_n + a\Delta - t] d\mathbb{P}[M' = t] \\ &= \int_0^{(1-\varepsilon)u_n} \mathbb{P}[\mathcal{Z}(R'_0) \geq u_n + a\Delta - t] d\mathbb{P}[M' = t] \\ &\quad + \int_{(1-\varepsilon)u_n}^\infty \mathbb{P}[\mathcal{Z}(R'_0) \geq u_n + a\Delta - t] d\mathbb{P}[M' = t] \\ &= I + II. \end{aligned}$$

To estimate  $I$ , suppose first that  $|R'_0| > \varepsilon'v_n$ , where  $\varepsilon'$  is much smaller than  $\varepsilon$ . As in Lemma 4.4,

we have  $E[e^{\theta M'}] < C(\theta)$  for every  $\theta < \theta_0$ . Then, by Petrov’s theorem,

$$\begin{aligned} I &= \int_0^{(1-\varepsilon)u_n} P[Z(R'_0) \geq u_n + a\Delta - t] dP[M' = t] \\ &\leq \int_0^{(1-\varepsilon)u_n} C u_n^{-1/2} \exp\{-\theta^*(u_n + a\Delta - t)\} dP[M' = t] \\ &\leq C u_n^{-1/2} \exp\{-\theta^* u_n\} e^{-\delta\Delta} \int_0^\infty e^{\theta^* t} dP[M' = t] \\ &\leq C e^{-\delta\Delta} P[\xi(v_n) > u_n]. \end{aligned}$$

Now suppose that  $|R'_0| < \varepsilon' v_n$ . Then, by Lemma 4.2, and if  $\varepsilon'$  is small enough,

$$I \leq P[Z(R'_0) > \varepsilon u_n] < \exp\{-(1 + \delta)\theta^* u_n\}.$$

To estimate  $II$ , note that, by Lemma 4.4 and if  $\varepsilon$  is sufficiently small,

$$II \leq P[M' > (1 - \varepsilon)u_n] \leq \exp\{-(1 + \delta)\theta^* u_n\}.$$

This proves (4.14). We now prove (4.15). By symmetry, we may assume that  $|R_1 \setminus R_2| \leq |R_2 \setminus R_1|$  and, hence,  $|R''_0| \geq \Delta/2 - O(1)$ . By the Markov inequality, for  $t > 0$  small,

$$\begin{aligned} P[B''] &= P[Z(R''_0) + M'' > -a\Delta] \\ &\leq e^{ta\Delta} E[\exp\{tZ(R''_0)\}] E[e^{tM''}] \\ &\leq C e^{ta\Delta} E[\exp\{tZ(R''_0)\}] \\ &= C \exp\{ta\Delta + |R''_0|\varphi(t)\} \\ &\leq C \exp\left\{\frac{\Delta(2ta + \varphi(t))}{2}\right\}. \end{aligned}$$

Now, since  $a > 0$  is small enough and  $\varphi'(0) < 0$ , we may choose  $t > 0$  so small that  $2ta + \varphi(t) < 0$ . This proves (4.15).

**Lemma 4.10.** *With the notation of Lemma 4.8 and its proof,*

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{S_2^m(n)}{s_n^d v_n^{d-1/2} P[\xi(v_n) > u_n]} = 0.$$

*Proof.* For each  $R \in \mathcal{L}_A(n)$ , the set  $\mathcal{B}_R^m(n)$  may be written as a union of  $m^{d+1}$  sets of the form  $\mathcal{B}_Q(n)$ , where  $Q \in q_n \mathbb{Z}^{d+1} \cap \mathcal{L}_A(n)$ . For two cubes  $Q_1, Q_2 \in q_n \mathbb{Z}^{d+1} \cap \mathcal{L}_A(n)$ , we write  $Q_1 \sim^m Q_2$  if there is an  $R \in \mathcal{L}_A^m(n)$  such that both  $\mathcal{B}_{Q_1}(n)$  and  $\mathcal{B}_{Q_2}(n)$  are subsets of  $\mathcal{B}_R^m(n)$ . It is not difficult to see that

$$S_2^m(n) \leq \sum_{\substack{Q_1, Q_2 \in q_n \mathbb{Z}^{d+1} \cap \mathcal{L}_A^m(n) \\ Q_1 \sim^m Q_2}} P[B_{Q_1}(n) \cap B_{Q_2}(n)].$$

Applying Lemma 4.9, we obtain

$$S_2^m(n) \leq C P[\xi(v_n) > u_n] \sum_{\substack{Q_1, Q_2 \in q_n \mathbb{Z}^{d+1} \cap \mathcal{L}_A^m(n) \\ Q_1 \sim^m Q_2}} \exp\{-\delta\Delta(Q_1, Q_2)\}.$$

If  $Q_1 = q_n w_1$  and  $Q_2 = q_n w_2$  for  $w_1, w_2 \in \mathbb{Z}^{d+1}$ , then  $\Delta(Q_1, Q_2) > c \|w_1 - w_2\|$  for some  $c > 0$ , where  $\|\cdot\|$  is any norm on  $\mathbb{R}^{d+1}$ . The lattice  $\mathbb{Z}^{d+1}$  can be decomposed into a disjoint union of discrete cubes of side length  $m$ , where the cubes have the form  $w + K_m$  for  $w \in m\mathbb{Z}^{d+1}$  and  $K_m = \{0, \dots, m - 1\}^{d+1}$ . For  $w_1, w_2 \in \mathbb{Z}^{d+1}$ , we write  $w_1 \sim^m w_2$  if  $w_1$  and  $w_2$  are contained in the same cube of the form described above. It is clear that  $Q_1 \sim^m Q_2$  if and only if  $w_1 \sim^m w_2$ . It follows that

$$S_2^m(n) \leq C \mathbb{P}[\xi(v_n) > u_n] \sum_{\substack{w_1, w_2 \in \mathbb{Z}^{d+1} \cap q_n^{-1}([0, s_n]^d \times [l_n^+, l_n^-]) \\ w_1 \sim^m w_2}} \exp\{-\delta \|w_1 - w_2\|\}.$$

The set  $\mathbb{Z}^{d+1} \cap q_n^{-1}([0, s_n]^d \times [l_n^+, l_n^-])$  contains  $O(1)s_n^d v_n^{d-1/2}$  points. If (for sufficiently large  $c$ )  $\|w_1 - w_2\| > cm$  then  $w_1 \not\sim^m w_2$ . Using this, we obtain

$$S_2^m(n) \leq C \mathbb{P}[\xi(v_n) > u_n] s_n^d v_n^{d-1/2} (I + II),$$

where

$$I = \sum_{\substack{w \in \mathbb{Z}^{d+1} \\ \|w\| \geq cm}} e^{-\delta \|w\|}, \quad II = \frac{1}{m^{d+1}} \sum_{\substack{w_1 \in K_m, w_2 \in \mathbb{Z}^{d+1} \\ \|w_1 - w_2\| \leq cm}} \exp\{-\delta \|w_1 - w_2\|\}.$$

Both  $I$  and  $II$  do not depend on  $n$ , and a straightforward calculation shows that  $\lim_{m \rightarrow \infty} I = \lim_{m \rightarrow \infty} II = 0$ . This completes the proof of Lemma 4.10.

Now we can complete the proof of Lemma 4.8. Equation (4.10) and, in particular, the existence of the limit in (4.11) follow from the Bonferroni inequalities, (4.12) and (4.13), as well as from the above asymptotic equalities for  $S_1^m(n)$  and  $S_2^m(n)$ . It remains to only show that  $H > 0$ .

We have

$$P_n \geq \sum_{R \in \mathcal{L}_A^m(n)} \mathbb{P}[B_R(n)] - \sum_{\substack{R_1, R_2 \in \mathcal{L}_A^m(n) \\ R_1 \neq R_2}} \mathbb{P}[B_{R_1}(n) \cap B_{R_2}(n)] = I - II.$$

By the above,  $I$  is asymptotically greater than  $cm^{-(d+1)} s_n^d v_n^{d-1/2} \mathbb{P}[\xi(v_n) > u_n]$  for some  $c > 0$ , whereas

$$II \leq C \mathbb{P}[\xi(v_n) > u_n] s_n^d v_n^{d-1/2} m^{-(d+1)} \sum_{w \in m\mathbb{Z}^{d+1}} e^{-\delta \|w\|}.$$

It follows that if  $m$  is sufficiently large then  $P_n \geq I - II \geq c s_n^d v_n^{d-1/2} \mathbb{P}[\xi(v_n) > u_n]$  for some  $c > 0$ . It follows that the constant  $H$  in (4.10) is positive.

**Remark 4.1.** Let the constant  $H^*$  in (4.1) be defined by  $H^* = \log H - d \log(\alpha^* \theta^*)$ . Then the statement of Lemma 4.8 may be written as

$$\mathbb{P}\left[\sup_{R \in \mathcal{L}_A(n)} \mathcal{Z}(R) > u_n\right] \sim D_A \frac{e^{-\tau}}{(n/s_n)^d},$$

where

$$D_A = \alpha^* (\sqrt{2\pi} \sigma^*)^{-1} \int_{-A}^A \exp\left\{-\frac{(\alpha^* t)^2}{2\sigma^{*2}}\right\} dt.$$

To see this, recall Lemma 4.5.



Finally, we are ready to prove the main result of this subsection.

**Lemma 4.11.** *Let  $\mathcal{R}_A(n) = [0, n]^d \times [l_n^-, l_n^+]$ . We have*

$$\lim_{n \rightarrow \infty} P \left[ \sup_{R \in \mathcal{R}_A(n)} \mathcal{Z}(R) \leq u_n \right] = \exp\{-D_A e^{-\tau}\}.$$

*Proof.* The set  $\mathcal{R}_A(n)$  can be decomposed into  $(n/s_n)^d$  translates of the set  $\mathcal{L}_A(n)$ , which was considered in Lemma 4.8 and Remark 4.1. The lemma then follows from the Poisson limit theorem: the only problem to overcome is that the events under consideration are dependent.

Let  $s_n = l_n$ . For  $(m_1, \dots, m_d) \in \mathbb{Z}^d \cap [0, n/s_n]^d$ , let

$$\mathcal{L}_A(n; m_1, \dots, m_d) = [m_1 s_n, (m_1 + 1)s_n] \times \dots \times [m_d s_n, (m_d + 1)s_n] \times [l_n^-, l_n^+],$$

and define the random event  $L_A(n; m_1, \dots, m_d) = \{\sup_{R \in \mathcal{L}_A(n; m_1, \dots, m_d)} \mathcal{Z}(R) > u_n\}$ . By Lemma 4.8 and Remark 4.1,

$$P[L_A(n; m_1, \dots, m_d)] \sim D_A \frac{e^{-\tau}}{(n/s_n)^d}. \tag{4.16}$$

Now we want to apply the Poisson limit theorem to the events

$$\left\{ L_A(n; m_1, \dots, m_d), (m_1, \dots, m_d) \in \mathbb{Z}^d \cap \left[0, \frac{n}{s_n}\right]^d \right\}.$$

Note that the events are only finite-range dependent. More precisely, the events  $L_A(n; m'_1, \dots, m'_d)$  and  $L_A(n; m''_1, \dots, m''_d)$  are independent if  $|m'_i - m''_i| > 1$  for at least one  $i = 1, \dots, d$ . In order to justify the use of the Poisson limit theorem, we have to show that

$$P[L_A(n; m'_1, \dots, m'_d) \cap L_A(n; m''_1, \dots, m''_d)] = o\left(\left(\frac{n}{s_n}\right)^{-d}\right) \text{ as } n \rightarrow \infty, \tag{4.17}$$

where the  $m''_i = m'_i + \varepsilon_i$ ,  $\varepsilon_i \in \{-1, 0, 1\}$ , are not all 0; see, e.g. Theorem 1 of [4]. To this end, we use Lemma 4.8 again, this time for  $s'_n = 3l_n$ . We obtain

$$P \left[ \bigcup_{(\varepsilon_1, \dots, \varepsilon_d) \in \{-1, 0, 1\}^d} L_A(n, m'_1 + \varepsilon_1, \dots, m'_d + \varepsilon_d) \right] \sim 3^d D_A \frac{e^{-\tau}}{(n/s_n)^d}.$$

On the other hand, by (4.16),

$$\sum_{(\varepsilon_1, \dots, \varepsilon_d) \in \{-1, 0, 1\}^d} P[L_A(n, m'_1 + \varepsilon_1, \dots, m'_d + \varepsilon_d)] \sim 3^d D_A \frac{e^{-\tau}}{(n/s_n)^d}.$$

Then (4.17) follows by the Bonferroni inequality.

#### 4.4. Cubes of nonoptimal size

In this subsection we deal with cubes whose volume differs significantly from the optimal volume,  $v_n$ . More precisely, we consider cubes with volume outside the interval  $[v_n - A\sqrt{v_n}, v_n + A\sqrt{v_n}]$ . We show that, if  $A \rightarrow \infty$  and  $n \rightarrow \infty$ , these cubes do not contribute to the extremal behavior of the random field  $\{\mathcal{Z}(R), R \in \mathcal{R}(n)\}$ . Let  $\varepsilon > 0$  be sufficiently small.

**Lemma 4.12.** *There is  $\delta > 0$  such that the following inequality holds uniformly in  $s$  and  $t$  as long as  $|s| \leq \varepsilon\sqrt{v_n}$  and  $t = o(\sqrt{u_n})$ :*

$$P[\xi(v_n + s\sqrt{v_n}) > u_n + t] \leq Ce^{-\theta^*t} \exp\{-\delta s^2\} P[\xi(v_n) > u_n].$$

*Proof.* Let  $\alpha_n = (u_n + t)/(v_n + s\sqrt{v_n})$ . Note that  $\alpha_n \in [(1 - \eta)\alpha^*, (1 + \eta)\alpha^*]$ , where  $\eta = \eta(\varepsilon)$  is small if  $\varepsilon$  is small. By Petrov’s theorem,

$$P[\xi(v_n + s\sqrt{v_n}) > u_n + t] \leq C \frac{1}{\sqrt{v_n}} \exp\{-(u_n + t)J(\alpha_n)\}. \tag{4.18}$$

An easy calculation shows that, for some  $c > 0$ ,

$$(\alpha_n - \alpha^*)^2 \geq c(s^2 - o(1))v_n^{-1} \quad \text{as } n \rightarrow \infty.$$

Applying Lemma 4.1 and recalling Lemma 4.5, we obtain, for some  $\delta > 0$ ,

$$J(\alpha_n) \geq \theta^* + \delta \frac{s^2}{v_n}.$$

Substituting this into (4.18), we obtain the statement of the lemma.

Let  $A > 0$  be large. Recall that  $l_n^-$  and  $l_n^+$  were defined in (4.7).

**Lemma 4.13.** *Define a set of cubes  $\mathcal{R}_1(n) = [0, n]^d \times [(1 - \varepsilon)l_n, l_n^-]$ , and let*

$$P_n(A) = P\left[\sup_{R \in \mathcal{R}_1(n)} \mathcal{Z}(R) > u_n\right].$$

*Then, for  $c(A) = \limsup_{n \rightarrow \infty} P_n(A)$ , we have  $\lim_{A \rightarrow \infty} c(A) = 0$ .*

*Proof.* Recall that  $q_n = l_n^{1-d}$ . Define the set  $\mathcal{B}_x(n) = [-q_n/2, q_n/2]^d \times [x, x + q_n]$ . Then

$$P_n(A) \leq \left(\frac{n}{q_n}\right)^d \sum_{x \in q_n \mathbb{Z} \cap [(1-\varepsilon)l_n, l_n^-]} P\left[\sup_{R \in \mathcal{B}_x(n)} \mathcal{Z}(R) > u_n\right]. \tag{4.19}$$

Given  $x \in [(1 - \varepsilon)l_n, l_n^-]$ , define a cube

$$R_0 = \left[-\frac{1}{2(x - q_n)}, \frac{1}{2(x - q_n)}\right]^d,$$

and define  $M = \sup_{R \in \mathcal{B}_x(n)} \mathcal{Z}(R) - \mathcal{Z}(R_0)$ . Note that  $M$  and  $\mathcal{Z}(R_0)$  are independent and that  $|R_0| = x^d + O(1)$ . We have

$$\begin{aligned} P\left[\sup_{R \in \mathcal{B}_x(n)} \mathcal{Z}(R) > u_n\right] &\leq P[\mathcal{Z}(R_0) > u_n] + \int_0^\infty P[\mathcal{Z}(R_0) > u_n - t] dP[M = t] \\ &= P[\mathcal{Z}(R_0) > u_n] + \int_0^{u_n^{1/3}} P[\mathcal{Z}(R_0) > u_n - t] dP[M = t] \\ &\quad + \int_{u_n^{1/3}}^\infty P[\mathcal{Z}(R_0) > u_n - t] dP[M = t] \\ &= I + II + III. \end{aligned}$$

Let  $s = s_n(x)$  be chosen such that  $x^d = v_n - s\sqrt{v_n}$ . By Lemma 4.12,

$$I \leq C \exp\{-\delta s^2\} P[\xi(v_n) > u_n].$$

We estimate the second term by first using Lemma 4.12 and then using Lemma 4.4:

$$\begin{aligned} II &\leq C \exp\{-\delta s^2\} P[\xi(v_n) > u_n] \int_0^{u_n^{1/3}} e^{\theta^* t} dP[M = t] \\ &\leq C \exp\{-\delta s^2\} P[\xi(v_n) > u_n]. \end{aligned}$$

Using Corollary 4.1 and then Lemma 4.4, the third term may be estimated by

$$\begin{aligned} III &\leq \int_{u_n^{1/3}}^\infty \exp\{-\theta^*(u_n - t)\} dP[M = t] \\ &= \exp\{-\theta^* u_n\} \int_{u_n^{1/3}}^\infty e^{\theta^* t} dP[M = t] \\ &\leq C \exp\{-\theta^* u_n\} \exp\{-\delta u_n^{1/3}\}. \end{aligned}$$

Bringing all three estimates together and recalling Lemma 4.5, we obtain

$$P\left[\sup_{R \in \mathcal{B}_x(n)} \mathcal{Z}(R) > u_n\right] \leq C u_n^{-1/2} \exp\{-\theta^* u_n\} \exp\{-\delta s^2\} + C \exp\{-\theta^* u_n\} \exp\{-\delta u_n^{1/3}\}.$$

It follows from (4.19) that  $P_n(A) \leq I' + II'$ , where

$$\begin{aligned} I' &= C \left(\frac{n}{q_n}\right)^d u_n^{-1/2} \exp\{-\theta^* u_n\} \sum_{x \in q_n \mathbb{Z} \cap [(1-\varepsilon)l_n, l_n^-]} \exp\{-\delta s_n(x)^2\}, \\ II' &= C \left(\frac{n}{q_n}\right)^d \exp\{-\theta^* u_n\} \exp\{-\delta u_n^{1/3}\} \sum_{x \in q_n \mathbb{Z} \cap [(1-\varepsilon)l_n, l_n^-]} 1. \end{aligned}$$

It is easy to see that  $\lim_{n \rightarrow \infty} II' = 0$ . We estimate  $I'$ . If  $x \in q_n \mathbb{Z} \cap [(1-\varepsilon)l_n, l_n^-]$  then the possible values of  $s_n(x)$  form a lattice in  $[A, \infty)$  with mesh size  $O(u_n^{-1/2})$ . Thus, estimating the Riemann sum by an integral, we obtain

$$I' \leq C \left(\frac{n}{q_n}\right)^d \exp\{-\theta^* u_n\} \int_A^\infty \exp\{-\delta s^2\} ds < C \int_A^\infty \exp\{-\delta s^2\} ds.$$

The statement of the lemma follows.

**Lemma 4.14.** Let  $\mathcal{R}_2(n) = [0, n]^d \times [0, (1-\varepsilon)l_n]$ , and define

$$P_n = P\left[\sup_{R \in \mathcal{R}_2(n)} \mathcal{Z}(R) > u_n\right].$$

Then  $\lim_{n \rightarrow \infty} P_n = 0$ .

*Proof.* The proof starts similarly to the proof of the previous lemma. For  $x \in [0, (1 - \varepsilon)l_n]$ , define a set of cubes  $\mathcal{B}_x(n) = [-q_n/2, q_n/2]^d \times [x, x + q_n]$ . Then

$$P_n \leq \left(\frac{n}{q_n}\right)^d \sum_{x \in q_n \mathbb{Z} \cap [0, (1-\varepsilon)l_n]} \mathbb{P}\left[\sup_{R \in \mathcal{B}_x(n)} \mathcal{Z}(R) > u_n\right]. \tag{4.20}$$

Define, as in the proof of the previous lemma,

$$R_0 = \left[-\frac{1}{2(x - q_n)}, \frac{1}{2(x - q_n)}\right]^d$$

and  $M = \sup_{R \in \mathcal{B}_x(n)} \mathcal{Z}(R) - \mathcal{Z}(R_0)$  (if  $x < q_n$ , we set  $R_0 = \emptyset$ ). Note that  $M$  and  $\mathcal{Z}(R_0)$  are independent. We have

$$\begin{aligned} \mathbb{P}\left[\sup_{R \in \mathcal{B}_x(n)} \mathcal{Z}(R) > u_n\right] &\leq \mathbb{P}[\mathcal{Z}(R_0) > u_n] + \int_0^\infty \mathbb{P}[\mathcal{Z}(R_0) > u_n - t] d\mathbb{P}[M = t] \\ &= \mathbb{P}[\mathcal{Z}(R_0) > u_n] + \int_0^{\varepsilon u_n/2} \mathbb{P}[\mathcal{Z}(R_0) > u_n - t] d\mathbb{P}[M = t] \\ &\quad + \int_{\varepsilon u_n/2}^\infty \mathbb{P}[\mathcal{Z}(R_0) > u_n - t] d\mathbb{P}[M = t] \\ &= I + II + III. \end{aligned}$$

To estimate the first term, we use Lemma 4.2 and the fact that  $|R_0| < (1 - \varepsilon)^d v_n$  (and, thus,  $u_n/|R_0| > (1 + \delta)\alpha^*$ ):

$$I \leq \exp\left\{-u_n J\left(\frac{u_n}{|R_0|}\right)\right\} \leq \exp\{-(1 + \delta)\theta^* u_n\}.$$

The second term is estimated analogously, using Corollary 4.1 and Lemma 4.4:

$$\begin{aligned} II &\leq \int_0^{\varepsilon u_n/2} \exp\left\{-(u_n - t) J\left(\frac{u_n - t}{|R_0|}\right)\right\} d\mathbb{P}[M = t] \\ &\leq \int_0^{\varepsilon u_n/2} \exp\{-(1 + \delta)\theta^*(u_n - t)\} d\mathbb{P}[M = t] \\ &= \exp\{-(1 + \delta)\theta^* u_n\} \int_0^{\varepsilon u_n/2} e^{(1+\delta)\theta^* t} d\mathbb{P}[M = t] \\ &\leq C \exp\{-(1 + \delta)\theta^* u_n\}. \end{aligned}$$

To estimate the third term, we again use Corollary 4.1 and Lemma 4.4:

$$\begin{aligned} III &\leq \int_{\varepsilon u_n/2}^\infty \exp\{-\theta^*(u_n - t)\} d\mathbb{P}[M = t] \\ &= \exp\{-\theta^* u_n\} \int_{\varepsilon u_n/2}^\infty e^{\theta^* t} d\mathbb{P}[M = t] \\ &\leq C \exp\{-(1 + \delta)\theta^* u_n\}. \end{aligned}$$

Bringing the three estimates together, we obtain

$$P\left[\sup_{R \in \mathcal{B}_x(n)} \mathcal{Z}(R) > u_n\right] \leq C \exp\{-(1 + \delta)\theta^* u_n\}.$$

It follows from (4.20) that

$$P_n \leq C \left(\frac{n}{q_n}\right)^d \exp\{-(1 + \delta)\theta^* u_n\} \sum_{x \in q_n \mathbb{Z} \cap [0, (1-\varepsilon)l_n]} 1,$$

which converges to 0 as  $n \rightarrow \infty$ . This completes the proof.

**Lemma 4.15.** Define  $\mathcal{R}_1^+(n) = [0, n]^d \times [l_n^+, (1 + \varepsilon)l_n]$ , and let

$$P_n(A) = P\left[\sup_{R \in \mathcal{R}_1^+(n)} \mathcal{Z}(R) > u_n\right].$$

Then, for  $c(A) = \limsup_{n \rightarrow \infty} P_n(A)$ , we have  $\lim_{A \rightarrow \infty} c(A) = 0$ .

*Proof.* The proof follows analogously to the proof of Lemma 4.13.

**Lemma 4.16.** Let  $\mathcal{R}_2^+(n) = [0, n]^d \times [(1 + \varepsilon)l_n, n]$ , and define

$$P_n = P\left[\sup_{R \in \mathcal{R}_2^+(n)} \mathcal{Z}(R) > u_n\right].$$

Then  $\lim_{n \rightarrow \infty} P_n = 0$ .

*Proof.* Let  $q_l = l^{1-d}$ , and, for  $x \in [l, l + 1]$ , define a set of cubes  $\mathcal{B}_x = [-q_l/2, q_l/2]^d \times [x, x + q_l]$ . Let  $M$  and  $R_0$  be defined as in the previous lemmas. Then  $P_n \leq P'_n + P''_n$ , where

$$P'_n = \sum_{l=[(1+\varepsilon)l_n]}^{[l_n^2]} (nq_l^{-1})^d \sum_{x \in q_l \mathbb{Z} \cap [l, l+1]} P\left[\sup_{R \in \mathcal{B}_x} \mathcal{Z}(R) > u_n\right],$$

$$P''_n = \sum_{l=[l_n^2]}^n (nq_l^{-1})^d \sum_{x \in q_l \mathbb{Z} \cap [l, l+1]} P\left[\sup_{R \in \mathcal{B}_x} \mathcal{Z}(R) > u_n\right].$$

If  $l \in [(1 + \varepsilon)l_n, l_n^2]$  then we use the estimate

$$P\left[\sup_{R \in \mathcal{B}_x} \mathcal{Z}(R) > u_n\right] \leq P[\mathcal{Z}(R_0) > u_n] + \int_0^\infty P[\mathcal{Z}(R_0) > u_n - t] dP[M = t]$$

$$= I + II.$$

The first term may be estimated using Lemma 4.2 and the fact that  $|R_0| \geq (1 + \varepsilon)v_n$ :

$$I \leq \exp\{-(1 + \delta)\theta^* u_n\}.$$

To estimate the second term, use additionally Lemma 4.4:

$$II \leq \int_0^\infty \exp\{-(1 + \delta)\theta^*(u_n - t)\} dP[M = t]$$

$$\leq \exp\{-(1 + \delta)\theta^* u_n\} \int_0^\infty e^{(1+\delta)\theta^* t} dP[M = t]$$

$$< C \exp\{-(1 + \delta)\theta^* u_n\}.$$

Using this, we obtain  $P'_n \leq Cn^d l_n^C \exp\{-(1 + \delta)\theta^* u_n\}$ , which converges to 0 as  $n \rightarrow \infty$ . Now suppose that  $l \in [l_n^2, n]$ . Then (the constant  $b$  is large)

$$P\left[\sup_{R \in \mathcal{B}_x} \mathcal{Z}(R) > u_n\right] \leq P[\mathcal{Z}(R_0) > -bu_n] + P[M > bu_n] = I + II.$$

Since  $|R_0| > cl_n^{2d} = cv_n^2$ ,  $c > 0$ , the first term may be estimated using, e.g. Petrov’s theorem:

$$P[\mathcal{Z}(R_0) > -bu_n] \leq C \exp\left\{-cv_n^2 I \left(\frac{bu_n}{v_n^2}\right)\right\} < C \exp\{-\delta v_n^2\} < C \exp\{-\delta u_n^2\} < Cn^{-D}$$

for any given  $D$ . To estimate the second term, we use Lemma 4.4:

$$P[M > bu_n] < C \exp\{-b\theta^* u_n\} < \frac{1}{n^D}$$

for any given  $D$  if  $b$  is sufficiently large. Bringing everything together, we obtain

$$P''_n \leq Cn^{d^2+1} l_n^C n^{-D},$$

which converges to 0 for large  $D$ . Thus,  $\lim_{n \rightarrow \infty} P''_n = 0$ . This completes the proof of the lemma.

**4.5. Proof of Theorem 2.1**

Now we are able to complete the proof of Theorem 2.1. Ignoring trivial boundary effects, we have, for every  $A > 0$ ,

$$P\left[\sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R) \leq u_n\right] \leq P\left[\sup_{R \in \mathcal{R}_A(n)} \mathcal{Z}(R) \leq u_n\right].$$

Letting  $n \rightarrow \infty$  and applying Lemma 4.11 to the right-hand side, we obtain

$$\limsup_{n \rightarrow \infty} P\left[\sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R) \leq u_n\right] \leq \exp\{-D_A e^{-\tau}\}.$$

Now, letting  $A \rightarrow \infty$  and using the fact that  $\lim_{A \rightarrow \infty} D_A = 1$ , we obtain

$$\limsup_{n \rightarrow \infty} P\left[\sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R) \leq u_n\right] \leq \exp\{-e^{-\tau}\}. \tag{4.21}$$

On the other hand, we have

$$P\left[\sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R) \leq u_n\right] \geq P\left[\sup_{R \in \mathcal{R}_A(n)} \mathcal{Z}(R) \leq u_n\right] - P\left[\sup_{R \in \mathcal{R}(n) \setminus \mathcal{R}_A(n)} \mathcal{Z}(R) > u_n\right].$$

Again, by Lemma 4.11, the first term on the right-hand side converges to  $\exp\{-D_A e^{-\tau}\}$  as  $n \rightarrow \infty$ . By Lemmas 4.13, 4.14, 4.15, and 4.16,

$$\lim_{A \rightarrow \infty} \limsup_{n \rightarrow \infty} P\left[\sup_{R \in \mathcal{R}(n) \setminus \mathcal{R}_A(n)} \mathcal{Z}(R) > u_n\right] = 0.$$

Thus, letting  $n \rightarrow \infty$  first and then  $A \rightarrow \infty$ , we obtain

$$\liminf_{n \rightarrow \infty} P\left[\sup_{R \in \mathcal{R}(n)} \mathcal{Z}(R) \leq u_n\right] \geq \exp\{-e^{-\tau}\}. \tag{4.22}$$

The statement of the theorem follows from (4.21) and (4.22).

### 5. Proof of Theorem 1.1

In this section we deduce Theorem 1.1 from Theorem 2.1 using a relation between marked empirical processes and compound Poisson processes stated below.

*Proof of Theorem 1.1.* First, let  $\{(U_i, X_i), i = 1, \dots, n\}$  be a marked empirical process, as in Section 1. On the other hand, let  $\{V_i, i \in \mathbb{N}\}$  be a Poisson point process on  $\mathbb{R}^d$  with unit intensity. To each point  $V_i$  we attach a mark  $Y_i$ . We suppose that the  $Y_i$  are i.i.d. with the same distribution as the marks  $X_i$  used in the construction of the marked empirical process, and that the  $Y_i$ s do not depend on the  $V_i$ s. For a Borel set  $R$ , let

$$\mathcal{Z}(R) = \sum_{i \in \mathbb{N}: V_i \in R} Y_i.$$

The compound Poisson process  $\mathcal{Z}$  is an example of Lévy noise. Now we are going to show that  $\mathcal{Z}$  satisfies conditions (L1)–(L3) provided that  $X_1$  satisfies conditions (X1)–(X3). First note that if  $\varphi$  is the logarithmic moment generating function of  $X_1$  then the logarithmic moment generating function of  $\mathcal{Z}([0, 1]^d)$  is  $\psi(t) = e^{\varphi(t)} - 1$ . Thus, if  $\varphi$  is finite on  $[0, \theta_0)$  and has a 0 at  $\theta^*$ , then the same holds for  $\psi$ . Finally, it is clear that if  $X_1$  is nonlattice then  $\mathcal{Z}([0, 1]^d)$  is also nonlattice. This shows that conditions (L1)–(L3) are satisfied.

We denote by  $N_t = \#\{i \in \mathbb{N}: V_i \in [0, t]^d\}$  the number of points of the compound Poisson process contained in the cube  $[0, t]^d$ . For  $n \in \mathbb{N}$ , let  $T_n = \inf\{t > 0: N_t = n + 1\}$ . Then we have the equality in distribution

$$\{\mathcal{X}_n(R), R \in \mathcal{R}(1)\} \sim \{\mathcal{Z}(T_n R), R \in \mathcal{R}(1)\}. \tag{5.1}$$

To see this, consider the right-hand side of (5.1) conditioned on  $\{T_n = t\}$ . Under this condition, the cube  $[0, t]^d$  contains  $n$  points of the Poisson point process  $\{V_i, i \in \mathbb{N}\}$  which have the same distribution as  $n$  points chosen independently and uniformly in  $[0, t]^d$ . Thus,

$$\{\mathcal{Z}(T_n R), R \in \mathcal{R}(1)\} \mid \{T_n = t\} \sim \{\mathcal{X}_n(R), R \in \mathcal{R}(1)\}.$$

Since this is true for every  $t$ , we obtain (5.1).

Define  $T_n^+ = (n + n^{2/3})^{1/d}$  and  $T_n^- = (n - n^{2/3})^{1/d}$ . Noting that  $E[N_{(n+n^{2/3})^{1/d}}] = \text{var}[N_{(n+n^{2/3})^{1/d}}] = n + n^{2/3}$  and using the central limit theorem, we obtain

$$\lim_{n \rightarrow \infty} P[T_n > T_n^+] = \lim_{n \rightarrow \infty} P[N_{(n+n^{2/3})^{1/d}} < n + 1] = 0.$$

Analogously, we have  $\lim_{n \rightarrow \infty} P[T_n < T_n^-] = 0$ .

We have

$$\begin{aligned} P\left[\sup_{R \in \mathcal{R}(1)} \mathcal{X}_n(R) \leq u_{n^{1/d}}\right] &= P\left[\sup_{R \in \mathcal{R}(T_n)} \mathcal{Z}(R) \leq u_{n^{1/d}}\right] \\ &\leq P\left[\sup_{R \in \mathcal{R}(T_n)} \mathcal{Z}(R) \leq u_{n^{1/d}} \cap T_n \geq T_n^-\right] + P[T_n < T_n^-] \\ &\leq P\left[\sup_{R \in \mathcal{R}(T_n^-)} \mathcal{Z}(R) \leq u_{n^{1/d}}\right] + P[T_n < T_n^-] \\ &= P\left[\sup_{R \in \mathcal{R}(T_n^-)} \mathcal{Z}(R) \leq u_{T_n^-} + o(1)\right] + P[T_n < T_n^-]. \end{aligned}$$

Now, the first term converges to  $\exp\{-e^{-\tau}\}$  by Theorem 2.1, whereas the second term was shown to converge to 0. This shows that

$$\limsup_{n \rightarrow \infty} \mathbb{P} \left[ \sup_{R \in \mathcal{R}(1)} \mathcal{X}_n(R) \leq u_{n^{1/d}} \right] \leq \exp\{-e^{-\tau}\}. \quad (5.2)$$

On the other hand,

$$\begin{aligned} \mathbb{P} \left[ \sup_{R \in \mathcal{R}(1)} \mathcal{X}_n(R) \leq u_{n^{1/d}} \right] &= \mathbb{P} \left[ \sup_{R \in \mathcal{R}(T_n)} \mathcal{Z}(R) \leq u_{n^{1/d}} \right] \\ &\geq \mathbb{P} \left[ \sup_{R \in \mathcal{R}(T_n^+)} \mathcal{Z}(R) \leq u_{n^{1/d}} \cap T_n \leq T_n^+ \right] \\ &\geq \mathbb{P} \left[ \sup_{R \in \mathcal{R}(T_n^+)} \mathcal{Z}(R) \leq u_{n^{1/d}} \right] - \mathbb{P}[T_n > T_n^+] \\ &= \mathbb{P} \left[ \sup_{R \in \mathcal{R}(T_n^+)} \mathcal{Z}(R) \leq u_{T_n^+} + o(1) \right] - \mathbb{P}[T_n > T_n^+]. \end{aligned}$$

As above, the first term converges to  $\exp\{-e^{-\tau}\}$  by Theorem 2.1, whereas the second converges to 0. This shows that

$$\liminf_{n \rightarrow \infty} \mathbb{P} \left[ \sup_{R \in \mathcal{R}(1)} \mathcal{X}_n(R) \leq u_{n^{1/d}} \right] \geq \exp\{-e^{-\tau}\}. \quad (5.3)$$

To complete the proof of Theorem 1.1, combine (5.2) with (5.3) and note that

$$u_{n^{1/d}} = \frac{1}{\theta^*} (\log n + (d-1) \log \log n + H^* + \tau) + o(1) \quad \text{as } n \rightarrow \infty.$$

### Acknowledgements

The authors are grateful to M. Schlather and the anonymous referee for numerous useful remarks. Z. Kabluchko was supported by DFG-FOR916.

### References

- [1] ADLER, R. J., MONRAD, D., SCISSORS, R. H. AND WILSON, R. (1983). Representations, decompositions and sample function continuity of random fields with independent increments. *Stoch. Process. Appl.* **15**, 3–30.
- [2] ALBIN, J. M. P. (1990). On extremal theory for stationary processes. *Ann. Prob.* **18**, 92–128.
- [3] ALDOUS, D. (1989). *Probability Approximations via the Poisson Clumping Heuristic* (Appl. Math. Sci. **77**). Springer, New York.
- [4] ARRATIA, R., GOLDSTEIN, L. AND GORDON, L. (1989). Two moments suffice for Poisson approximations: the Chen–Stein method. *Ann. Prob.* **17**, 9–25.
- [5] BICKEL, P. J. AND WICHURA, M. J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* **42**, 1656–1670.
- [6] CHAN, H. P. (2007). Maxima of moving sums in a Poisson random field. Preprint. Available at <http://arxiv.org/abs/0708.2764>.
- [7] COHEN, J. W. (1968). Extreme value distribution for the M/G/1 and the G/M/1 queueing systems. *Ann. Inst. H. Poincaré Sect. B* **4**, 83–98.
- [8] DONEY, R. A. AND MALLER, R. A. (2005). Cramér’s estimate for a reflected Lévy process. *Ann. Appl. Prob.* **15**, 1445–1450.
- [9] GLAZ, J. AND BALAKRISHNAN, N. (eds) (1999). *Scan Statistics and Applications*. Birkhäuser, Boston, MA.
- [10] GLAZ, J., NAUS, J. AND WALLENSTEIN, S. (2001). *Scan Statistics*. Springer, New York.
- [11] HÜSLER, J. AND PITERBARG, V. (2004). Limit theorem for maximum of the storage process with fractional Brownian motion as input. *Stoch. Process. Appl.* **114**, 231–250.



- [12] IGLEHART, D. L. (1972). Extreme values in the GI/G/1 queue. *Ann. Math. Statist.* **43**, 627–635.
- [13] JIANG, T. (2002). Maxima of partial sums indexed by geometrical structures. *Ann. Prob.* **30**, 1854–1892.
- [14] KARLIN, S. AND DEMBO, A. (1992). Limit distributions of maximal segmental score among Markov-dependent partial sums. *Adv. Appl. Prob.* **24**, 113–140.
- [15] KOMLÓS, J. AND TUSNÁDY, G. (1975). On sequences of “pure heads”. *Ann. Prob.* **3**, 608–617.
- [16] LEADBETTER, M. R., LINDGREN, G. AND ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.
- [17] PETROV, V. V. (1965). On the probabilities of large deviations for sums of independent random variables. *Teor. Veroyat. Primen.* **10**, 310–322.
- [18] PICKANDS, J., III (1969). Asymptotic properties of the maximum in a stationary Gaussian process. *Trans. Amer. Math. Soc.* **145**, 75–86.
- [19] PICKANDS, J., III (1969). Upcrossing probabilities for stationary Gaussian processes. *Trans. Amer. Math. Soc.* **145**, 51–73.
- [20] PITERBARG, V. I. (1996). *Asymptotic Methods in the Theory of Gaussian Processes and Fields* (Trans. Math. Mono. **148**). American Mathematical Society, Providence, RI.
- [21] PITERBARG, V. I. AND KOZLOV, A. M. (2003). On large jumps of a random walk with the Cramér condition. *Theory Prob. Appl.* **47**, 719–729.
- [22] WILLEKENS, E. (1987). On the supremum of an infinitely divisible process. *Stoch. Process. Appl.* **26**, 173–175.
- [23] ZEEVI, A. J. AND GLYNN, P. W. (2000). On the maximum workload of a queue fed by fractional Brownian motion. *Ann. Appl. Prob.* **10**, 1084–1099.