

# Limiting distribution for the maximal standardized increment of a random walk

Zakhar Kabluchko<sup>a,\*</sup>, Yizao Wang<sup>b</sup>

<sup>a</sup> *Institute of Stochastics, Ulm University, Helmholtzstr. 18, 89069 Ulm, Germany*

<sup>b</sup> *Department of Mathematical Sciences, University of Cincinnati, 4302 French Hall, 2815 Commons Way, Cincinnati, OH, 45221-0025, United States*

Received 30 September 2013; received in revised form 10 March 2014; accepted 31 March 2014

Available online 15 April 2014

## Abstract

Let  $X_1, X_2, \dots$  be independent identically distributed (i.i.d.) random variables with  $\mathbb{E}X_k = 0$ ,  $\text{Var}X_k = 1$ . Suppose that  $\varphi(t) := \log \mathbb{E}e^{tX_k} < \infty$  for all  $t > -\sigma_0$  and some  $\sigma_0 > 0$ . Let  $S_k = X_1 + \dots + X_k$  and  $S_0 = 0$ . We are interested in the limiting distribution of the *multiscale scan statistic*

$$\mathbf{M}_n = \max_{0 \leq i < j \leq n} \frac{S_j - S_i}{\sqrt{j - i}}.$$

We prove that for an appropriate normalizing sequence  $a_n$ , the random variable  $\mathbf{M}_n^2 - a_n$  converges to the Gumbel extreme-value law  $\exp\{-e^{-cx}\}$ . The behavior of  $\mathbf{M}_n$  depends strongly on the distribution of the  $X_k$ 's. We distinguish between four cases. In the *superlogarithmic* case we assume that  $\varphi(t) < t^2/2$  for every  $t > 0$ . In this case, we show that the main contribution to  $\mathbf{M}_n$  comes from the intervals  $(i, j)$  having length  $l := j - i$  of order  $a(\log n)^p$ ,  $a > 0$ , where  $p = q/(q - 2)$  and  $q \in \{3, 4, \dots\}$  is the order of the first non-vanishing cumulant of  $X_1$  (not counting the variance). In the *logarithmic* case we assume that the function  $\psi(t) := 2\varphi(t)/t^2$  attains its maximum  $m_* > 1$  at some unique point  $t = t_* \in (0, \infty)$ . In this case, we show that the main contribution to  $\mathbf{M}_n$  comes from the intervals  $(i, j)$  of length  $d_* \log n + a\sqrt{\log n}$ ,  $a \in \mathbb{R}$ , where  $d_* = 1/\varphi(t_*) > 0$ . In the *sublogarithmic* case we assume that the tail of  $X_k$  is heavier than  $\exp\{-x^{2-\varepsilon}\}$ , for some  $\varepsilon > 0$ . In this case, the main contribution to  $\mathbf{M}_n$  comes from the intervals of length  $o(\log n)$  and in fact, under regularity assumptions, from the intervals of length 1. In the remaining, fourth case, the  $X_k$ 's are *Gaussian*. This case has been studied earlier in the literature. The main contribution

\* Corresponding author. Tel.: +49 07315023527.

E-mail addresses: [sachar.k@gmx.net](mailto:sachar.k@gmx.net), [zakhar.kabluchko@uni-ulm.de](mailto:zakhar.kabluchko@uni-ulm.de) (Z. Kabluchko), [yizao.wang@uc.edu](mailto:yizao.wang@uc.edu) (Y. Wang).

comes from intervals of length  $a \log n$ ,  $a > 0$ . We argue that our results cover most interesting distributions with light tails. The proofs are based on the precise asymptotic estimates for large and moderate deviation probabilities for sums of i.i.d. random variables due to Cramér, Bahadur, Ranga Rao, Petrov and others, and a careful extreme value analysis of the random field of standardized increments by the double sum method. © 2014 Elsevier B.V. All rights reserved.

MSC: 60G50; 60G70; 60F10; 60F05

Keywords: Extreme value theory; Increments of random walks; Erdős–Rényi law; Large deviations; Moderate deviations; Multiscale scan statistic; Cramér series; Gumbel distribution; Double sum method; Subgaussian distributions; Change-point detection

## 1. Introduction and statement of results

### 1.1. Introduction

Suppose we are given a long sequence of observations. The observations are assumed to be independent identically distributed (i.i.d.) random variables with zero mean and unit variance, except, possibly, for a short interval, where the observations have positive mean. This interval may be interpreted as a signal in an i.i.d. noise. The question is how to decide whether a signal is present and if yes, how to locate it. A natural approach is to build a *multiscale scan statistic*. For every interval we compute the sum of the observations in this interval divided by the square root of the length of the interval. Large values of this normalized sum indicate the presence of a signal. Since no a priori knowledge about the location and length of the interval containing the signal is available, we take the maximum of such normalized sums over all possible intervals of all possible lengths. Scan statistics with windows of *fixed size* have been much studied; see, e.g., [16,17]. A large class of limit theorems dealing with fixed window size are the Erdős–Rényi–Shepp laws; see, e.g., [10–14,6]. The scan statistic we are interested in is built using windows of *all possible sizes*. In order to use this statistic for testing purposes we need to know its asymptotic distribution under the null hypothesis.

We arrive at the following problem. Let  $X_1, X_2, \dots$  be i.i.d. non-degenerate random variables with  $\mathbb{E}X_k = 0$ ,  $\text{Var}X_k = 1$ . Consider a random walk given by  $S_k = X_1 + \dots + X_k$ ,  $k \in \mathbb{N}$ , and  $S_0 = 0$ . For  $n \in \mathbb{N}$  define the multiscale scan statistic  $\mathbf{M}_n$  by

$$\mathbf{M}_n = \max_{0 \leq i < j \leq n} \frac{S_j - S_i}{\sqrt{j - i}}. \tag{1}$$

Following results on the asymptotic behavior of  $\mathbf{M}_n$  as  $n \rightarrow \infty$  are known. For random variables with finite exponential moments, Shao [39], confirming and extending a conjecture of Révész [38], proved that

$$\lim_{n \rightarrow \infty} \frac{\mathbf{M}_n}{\sqrt{2 \log n}} = \sqrt{m_*} \quad \text{a.s.} \tag{2}$$

Here,  $m_* \in [1, \infty]$  is a constant determined explicitly in terms of the distribution of  $X_1$ . Shao’s proof has been considerably simplified by Steinebach [43]; see also [23] for a multidimensional generalization. This describes the a.s. rate of growth of  $\mathbf{M}_n$ . But what about the limiting distribution? In the case when  $X_1, X_2, \dots$  are i.i.d. standard Gaussian, Siegmund and Venkatraman [40]

showed that for all  $\tau \in \mathbb{R}$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \mathbf{M}_n \leq \sqrt{2 \log n} + \frac{\frac{1}{2} \log \log n + \log \frac{H}{2\sqrt{\pi}} + \tau}{\sqrt{2 \log n}} \right] = \exp\{-e^{-\tau}\}. \tag{3}$$

Here,  $H > 0$  is some explicit constant. The distribution on the right-hand side is the Gumbel extreme-value law. An independent proof of the same result was given in [21]. It was shown in [20,21] that a result similar to (3), but with a different normalization, holds if we replace the Gaussian random walk by a Brownian motion. Generalizations of both results to the multidimensional setting with intervals replaced by cubes or rectangles, have been obtained in [22]. A similar problem for a totally skewed  $\alpha$ -stable Lévy process has been considered in [20]. In the case when  $X_1$  has regularly varying right tail, limit Fréchet distribution for  $\mathbf{M}_n$  has been obtained by Mikosch and Račkauskas [30]; see also Mikosch and Moser [29].

Apart from these special cases nothing has been known about the limiting distribution of  $\mathbf{M}_n$ . Our aim is to settle this problem for a broad class of random variables with light tails. It turns out that the behavior of  $\mathbf{M}_n$  depends heavily on some fine properties of the distribution of  $X_1$ . We assume that for some  $\sigma_0 > 0$ ,

$$\varphi(t) := \log \mathbb{E}e^{tX_1} < \infty \quad \text{for all } t \geq -\sigma_0. \tag{4}$$

The function  $\varphi$  (called the cumulant generating function of  $X_1$ ) is strictly increasing on  $[0, \infty)$ , strictly convex, infinitely differentiable, and vanishes at 0.

We will consider four cases depending on where the supremum of the function

$$\psi(t) := \frac{\varphi(t)}{t^2/2}, \quad t > 0, \tag{5}$$

is attained. The constant  $m_*$  in Shao’s result (2) is determined by  $m_* = \sup_{t>0} \psi(t)$ . Note that  $\lim_{t \downarrow 0} \psi(t) = 1$  since  $\varphi(t) \sim t^2/2$  as  $t \downarrow 0$ . Hence,  $m_* \geq 1$ . If  $X_1$  is standard Gaussian, we even have  $\psi(t) = 1$  identically, for all  $t \in \mathbb{R}$ . Our four cases can be roughly described as follows, see Fig. 1:

1. *Gaussian case:*  $\psi(t) = 1$  for all  $t \in \mathbb{R}$ .
2. *Superlogarithmic case:* the supremum  $m_* = 1$  is attained as  $t \downarrow 0$ .
3. *Logarithmic case:* the supremum  $m_* > 1$  is attained at some  $t = t_* \in (0, \infty)$ .
4. *Sublogarithmic case:*  $m_* = +\infty$ .

Since the Gaussian case has been fully analyzed in [40,21,22], we concentrate on the remaining three cases. Let us explain the difference between the cases. The definition of  $\mathbf{M}_n$  involves a maximum taken over intervals  $(i, j)$  of different lengths  $l := j - i$ . It turns out that different lengths make different contributions to  $\mathbf{M}_n$ . In all three cases we will single out some family of lengths which are *optimal* in the sense that the contribution of all other lengths to  $\mathbf{M}_n$  is asymptotically negligible. We will show that the optimal lengths are given as follows:

1. *Gaussian case:*  $l = a \log n, a > 0$ .
2. *Superlogarithmic case:*  $l = a \log^p n, a > 0$ , where  $p > 1$ .
3. *Logarithmic case:*  $l = d_* \log n + a\sqrt{\log n}, a \in \mathbb{R}$ , where  $d_* > 0$ .
4. *Sublogarithmic case:*  $l = o(\log n)$ , and, under more assumptions,  $l = 1$ .

Here,  $p > 1$  and  $d_* > 0$  are parameters depending on the distribution of  $X_1$ . To give exact meaning to these statements we will analyze the random variable  $\mathbf{M}_n(h_1, h_2)$  obtained by restricting the lengths over which the maximum is taken to some range  $[h_1, h_2]$ . Namely, for  $0 \leq h_1 < h_2$ ,

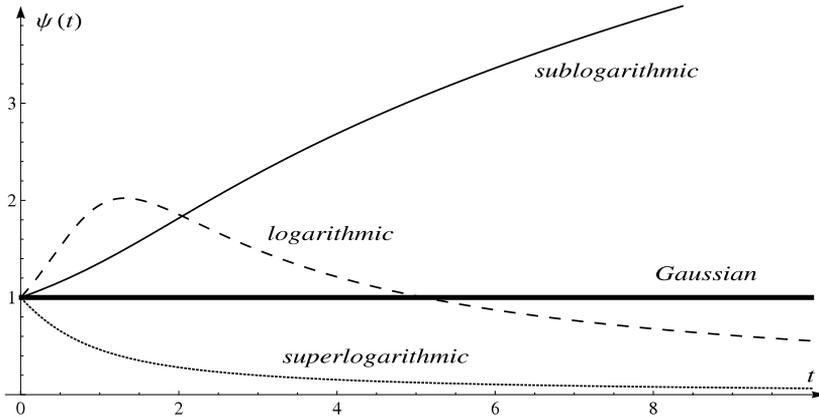


Fig. 1. The graph of  $\psi(t) = 2\varphi(t)/t^2$  in the Gaussian case (bold), superlogarithmic case (dotted), logarithmic case (dashed), sublogarithmic case (solid).

define

$$M_n(h_1, h_2) = \max_{\substack{0 \leq i < j \leq n \\ h_1 \leq j-i \leq h_2}} \frac{S_j - S_i}{\sqrt{j-i}}. \tag{6}$$

Then, the statement that in the superlogarithmic case the lengths  $a \log^p n, a > 0$ , are optimal, means that

$$\lim_{A \rightarrow +\infty} \limsup_{n \rightarrow \infty} \mathbb{P}[M_n = M_n(A^{-1} \log^p n, A \log^p n)] = 1.$$

As we will show below, analogous statements hold in all four cases.

We are now ready to state our results on the limiting distribution of the multiscale scan statistic  $M_n$ . In fact, the results will be stated in terms of  $M_n^2$  because this greatly simplifies the notation. It is easy to switch back to  $M_n$ ; see Section 1.6.1.

### 1.2. The superlogarithmic case

Here we consider random variables which are in some sense dominated by the Gaussian distribution. We assume that for all  $\varepsilon > 0$ ,

$$\sup_{t \geq \varepsilon} \frac{\varphi(t)}{t^2/2} < 1. \tag{7}$$

Equivalently,  $\psi(t) < 1$  for every  $t > 0$ , and  $\limsup_{t \rightarrow \infty} \psi(t) < 1$ . A closely related notion is the subgaussianity; see [7]. The first non-zero term in the Taylor expansion of  $\varphi(t)$  at 0 is  $t^2/2$  since we assume that  $\mathbb{E}X_k = 0, \text{Var}X_k = 1$ . Of crucial importance will be the second non-zero term in the Taylor expansion of  $\varphi(t)$ . We have, for some  $q \in \{3, 4, \dots\}$  and  $\kappa > 0$ ,

$$\varphi(t) = \frac{t^2}{2} - \kappa t^q + o(t^q), \quad t \downarrow 0. \tag{8}$$

Thus,  $q$  is the order of the first non-zero cumulant of  $X_1$ , not counting the variance. The most common value of  $q$  is 3, however, for symmetric distributions the third cumulant vanishes and we

typically have  $q = 4$ . Note that the coefficient  $\kappa$  cannot be negative, since otherwise (7) would be violated for sufficiently small  $\varepsilon > 0$ .

**Theorem 1.1.** *Let  $X_1, X_2, \dots$  be i.i.d. random variables with  $\mathbb{E}X_k = 0, \text{Var}X_k = 1$ . Suppose that conditions (4), (7), (8) are satisfied. Then, for all  $\tau \in \mathbb{R}$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{1}{2} \mathbf{M}_n^2 \leq \log(n \log^{\frac{1}{2} \cdot \frac{q-6}{q-2}} n) + \tau \right] = \exp\{-\Lambda_{q,\kappa} e^{-\tau}\}, \tag{9}$$

where  $\Lambda_{q,\kappa} = \frac{1}{\sqrt{\pi}} \Gamma(\frac{q}{q-2})(2\kappa)^{\frac{2}{q-2}}$ .

It turns out that in the superlogarithmic case, the main contribution to  $\mathbf{M}_n$  is done by intervals with length  $a \log^p n$ , where  $a \in (0, \infty)$ , and

$$p = \frac{q}{q-2} \in \left\{ 3, 2, \frac{5}{3}, \frac{3}{2}, \frac{7}{5}, \dots \right\}. \tag{10}$$

Moreover, we will even prove that the ‘‘intensity’’ with which the length  $a \log^p n$  contributes to  $\mathbf{M}_n$  is given by some explicit function  $\Lambda_{q,\kappa}(a)$ . Namely, we have the following result.

**Theorem 1.2.** *Fix arbitrary  $0 < A_1 < A_2$ . Define  $l_n^- = A_1 \log^p n$  and  $l_n^+ = A_2 \log^p n$ . Under the same assumptions as in Theorem 1.1, for every  $\tau \in \mathbb{R}$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{1}{2} \mathbf{M}_n^2(l_n^-, l_n^+) \leq \log(n \log^{\frac{1}{2} \cdot \frac{q-6}{q-2}} n) + \tau \right] = \exp \left\{ -e^{-\tau} \int_{A_1}^{A_2} \Lambda_{q,\kappa}(a) da \right\}, \tag{11}$$

where  $\Lambda_{q,\kappa} : (0, \infty) \rightarrow (0, \infty)$  is a function given by

$$\Lambda_{q,\kappa}(a) = \frac{1}{2\sqrt{\pi}a^2} \exp\{-\kappa 2^{\frac{q}{2}} a^{-\frac{q-2}{2}}\}. \tag{12}$$

Note that  $\Lambda_{q,\kappa} = \int_0^\infty \Lambda_{q,\kappa}(a) da$ , so that formally we can obtain Theorem 1.1 from Theorem 1.2 by taking  $A_1 = 0, A_2 = \infty$ . Note that  $\Lambda_{q,\kappa}(a) \rightarrow 0$  as  $a \downarrow 0$  or  $a \uparrow \infty$ . This means that too small and too large intervals make small contributions to  $\mathbf{M}_n$ . The unique maximum of the function  $a \mapsto \Lambda_{q,\kappa}(a)$  is attained at

$$a_* = 2^{\frac{q-4}{q-2}} \kappa^{\frac{2}{q-2}} (q-2)^{\frac{2}{q-2}}.$$

Thus, the largest contribution to  $\mathbf{M}_n$  comes from the intervals of length  $(a_* + o(1)) \log^p n$ .

### 1.3. The logarithmic case

We assume that (4) holds and there is  $t_* > 0$  such that

$$m_* := \frac{\varphi(t_*)}{t_*^2/2} > 1. \tag{13}$$

Moreover, we assume that  $t_*$  is the unique point of maximum of  $\psi(t) = 2\varphi(t)/t^2$  in the following uniform sense: for every  $\varepsilon > 0$ ,

$$\sup_{0 < t < t_* - \varepsilon} \frac{\varphi(t)}{t^2/2} < m_* \quad \text{and} \quad \sup_{t > t_* + \varepsilon} \frac{\varphi(t)}{t^2/2} < m_*. \tag{14}$$

Equivalently,  $\psi(t) < m_*$  for all  $t > 0, t \neq t_*$ , and  $\limsup_{t \rightarrow \infty} \psi(t) < m_*$ . Recall that a random variable  $X_1$  is lattice if there are  $a > 0, b \in \mathbb{R}$  such that  $X_1 \in a\mathbb{Z} + b$  with probability 1. Otherwise,  $X_1$  is called non-lattice.

**Theorem 1.3.** *Let  $X_1, X_2, \dots$  be i.i.d. random variables with  $\mathbb{E}X_k = 0, \text{Var}X_k = 1$ . Suppose that conditions (4), (13), (14) are satisfied and the distribution of  $X_1$  is non-lattice. Then, for every  $\tau \in \mathbb{R}$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{1}{2} \mathbf{M}_n^2 \leq m_*(\log n + \tau) \right] = \exp \{ -\Theta_* e^{-\tau} \}, \tag{15}$$

where  $\Theta_* = \frac{\sqrt{m_* H_*^2}}{\sqrt{2\beta_* \sigma_*}}, \sigma_*^2 = \varphi''(t_*)$  and

$$\beta_*^2 = \frac{(\varphi'(t_*))^4}{8m_*} \left( \frac{1}{\sigma_*^2} - \frac{1}{m_*} \right) > 0, \tag{16}$$

$$H_* = \lim_{B \rightarrow \infty} \frac{1}{B} \mathbb{E} \left[ \max_{k=0, \dots, B} e^{t_* S_k - k\varphi(t_*)} \right] \in (0, 1). \tag{17}$$

The positivity of  $\beta_*^2$  will be established in Lemma 5.2. More explicit expression for  $H_*$  will be given in Section 5.8 below. We believe that in the lattice case the convergence in Theorem 1.3 breaks down. (A similar phenomenon was observed in [26] for scan statistics with fixed window size.) However, we still have tightness.

**Theorem 1.4.** *Let  $X_1, X_2, \dots$  be i.i.d. random variables with  $\mathbb{E}X_k = 0, \text{Var}X_k = 1$ . Assume that conditions (4), (13), (14) hold. Then, the sequence of random variables  $\mathbf{M}_n^2 - 2m_* \log n, n \in \mathbb{N}$ , is tight.*

In the next theorem we compute the contribution of different lengths to  $\mathbf{M}_n$ . Let  $d_* = 1/\varphi(t_*)$ . We will show that only intervals whose length differs from  $d_* \log n$  by a quantity of order  $\sqrt{\log n}$  are relevant. Recall that  $\mathbf{M}_n(h_1, h_2)$  was defined in (6).

**Theorem 1.5.** *Fix arbitrary  $A_1 < A_2$ . Define  $l_n^- = d_* \log n + A_1 \sqrt{\log n}$  and  $l_n^+ = d_* \log n + A_2 \sqrt{\log n}$ . Under the same assumptions as in Theorem 1.3, for every  $\tau \in \mathbb{R}$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \frac{1}{2} \mathbf{M}_n^2(l_n^-, l_n^+) \leq m_*(\log n + \tau) \right] = \exp \left\{ -e^{-\tau} \int_{A_1}^{A_2} \Theta(a) da \right\}. \tag{18}$$

Here,  $\Theta : \mathbb{R} \rightarrow (0, \infty)$  is a function given by  $\Theta(a) = \frac{\sqrt{m_* H_*^2}}{2\sqrt{\pi} \sigma_*} e^{-\frac{\beta_*^2 a^2}{2}}$ , where  $\beta_*$  and  $H_*$  are as in (16) and (17).

Note that  $\Theta_* = \int_{-\infty}^{\infty} \Theta(a) da$ , so that we can obtain (at least formally) Theorem 1.3 from Theorem 1.5 by taking  $A_1 = -\infty, A_2 = \infty$ . Since  $\Theta(a)$  attains its maximum at  $a = 0$ , the maximal contribution to  $\mathbf{M}_n$  comes from the intervals with length  $d_* \log n + o(\sqrt{\log n})$ .

#### 1.4. The sublogarithmic case

In this case we consider random variables whose right tail is heavier than the standard Gaussian tail. In this case only intervals of length  $o(\log n)$  make contribution to  $\mathbf{M}_n$ , as the next theorem shows.

**Theorem 1.6.** Let  $X_1, X_2, \dots$  be i.i.d. random variables with  $\mathbb{E}X_k = 0, \text{Var}X_k = 1$  and such that  $\varphi(t) = \log \mathbb{E}e^{tX_1}$  is finite on  $[0, t_0)$ , for some  $t_0 > 0$ . Assume that for some  $\alpha < 2$  we have  $\mathbb{P}[X_1 > x] > e^{-x^\alpha}$ , for sufficiently large  $x$ . Then, for every  $a > 0$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathbf{M}_n = \mathbf{M}_n(1, a \log n)] = 1. \tag{19}$$

Under some regularity assumptions on the tail of  $X_1$  it is possible to show that only intervals of length 1 (that is, only individual observations) contribute to  $\mathbf{M}_n$ . In this case, the study of  $\mathbf{M}_n$  is equivalent to the study of the maximum  $U_n = \max\{X_1, \dots, X_n\}$ . We assume that for some  $\alpha \in [1, 2)$  and  $D > 0$ ,

$$\lim_{x \rightarrow +\infty} \frac{1}{x^\alpha} \log \mathbb{P}[X_1 > x] = -D. \tag{20}$$

**Theorem 1.7.** Let  $X_1, X_2, \dots$  be i.i.d. random variables with  $\mathbb{E}X_k = 0, \text{Var}X_k = 1$  and such that (20) holds. Then,  $\lim_{n \rightarrow \infty} \mathbb{P}[\mathbf{M}_n = \mathbf{U}_n] = 1$ .

The above results show that the square-root normalization we used in (1) is not very natural in the sublogarithmic case. See [43,27,41,30,29] for other types of normalization which can be used in the case of sublogarithmic (or even heavier) tails.

### 1.5. Examples

In this section we show that most classical families of distributions considered in the probability theory belong to one of the four cases considered above.

#### 1.5.1. Symmetric Bernoulli

Let  $X_1, X_2, \dots$  be i.i.d. with symmetric Bernoulli distribution, that is  $\mathbb{P}[X_k = \pm 1] = 1/2$ . We have

$$\varphi(t) = \log \cosh(t) = \frac{t^2}{2} - \frac{t^4}{12} + o(t^4), \quad t \rightarrow 0.$$

We are in the superlogarithmic case. Indeed, all coefficients of the Taylor series

$$\cosh t - e^{t^2/2} = \frac{1}{2} \sum_{k=2}^{\infty} t^{2k} \left( \frac{1}{(2k)!} - \frac{1}{2^k k!} \right)$$

are negative. This shows that  $\varphi(t) < t^2/2$  for every  $t > 0$ . Since we also have  $\lim_{t \rightarrow \infty} \varphi(t)/t^2 = 0$ , it follows that condition (7) is fulfilled. We are in the superlogarithmic case with  $q = 4$ . The optimal lengths are  $a \log^2 n, a > 0$ .

#### 1.5.2. Non-symmetric Bernoulli

Fix  $p \in (0, 1)$  and consider i.i.d. Bernoulli random variables  $Y_1, Y_2, \dots$  with  $\mathbb{P}[Y_k = +1] = p$  and  $\mathbb{P}[Y_k = -1] = 1 - p$ . Consider also the normalized random variables

$$X_k = \frac{1}{\sigma}(Y_k - (2p - 1)), \quad \sigma = \sqrt{4p(1 - p)}.$$

Then,  $\mathbb{E}X_k = 0$  and  $\text{Var}X_k = 1$ . The cumulant generating function of  $X_k$  is given by

$$\varphi(t) = -\frac{2pt}{\sigma} + \log(1 + pe^{\frac{2t}{\sigma}} - p) = \frac{t^2}{2} + \frac{1 - 2p}{3\sigma}t^3 + o(t^3), \quad t \rightarrow 0. \tag{21}$$

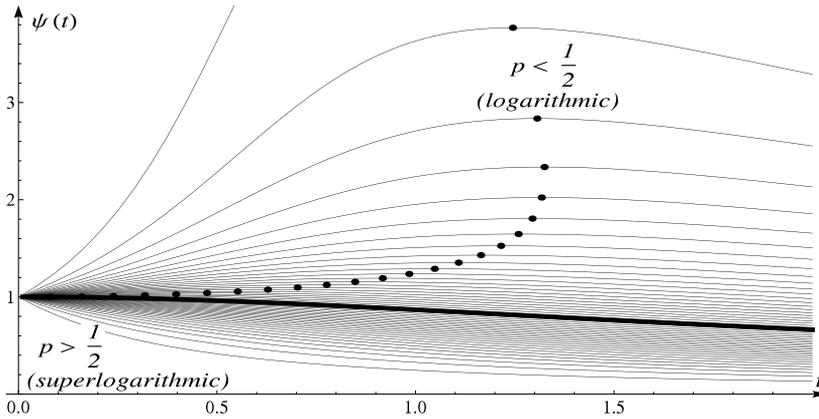


Fig. 2. The graph of  $\psi(t) = 2\varphi(t)/t^2$  for the (generally, non-symmetric) Bernoulli distributions in dependence on the parameter  $p = 1/50, \dots, 49/50$ . The bold black boundary corresponds to  $p = 1/2$  (superlogarithmic case). The curves below the boundary correspond to  $p > 1/2$  (superlogarithmic case), the curves above the boundary correspond to  $p < 1/2$  (logarithmic case). On each curve above the boundary the point of maximum  $(t_*, m_*)$  is shown.

The graph of the function  $\psi(t) = 2\varphi(t)/t^2$  in dependence on the parameter  $p$  is shown in Fig. 2. As already shown above, for  $p = 1/2$  we are in the superlogarithmic case with  $q = 4$ , and the optimal lengths are  $a \log^2 n, a > 0$ .

**Proposition 1.8.** *If  $p \in (1/2, 1)$ , then  $\varphi(t) < t^2/2$  for all  $t > 0$ . We are in the superlogarithmic case with  $q = 3$ . The optimal lengths are  $a \log^3 n, a \geq 0$ .*

For  $p < 1/2$  the coefficient of  $t^3$  in the Taylor series (21) is positive. This implies that  $m_* = \sup_{t>0} \psi(t) > 1$ . Also, it follows from (21) that  $\lim_{t \rightarrow \infty} \psi(t) = 0$ , hence the maximum  $m_*$  is finite. From Fig. 2 we see that the maximum is attained at a unique point. (We were not able to prove this fact rigorously. It seems that the proof requires tedious computations with transcendental functions). Thus, we should be in the logarithmic case. The optimal lengths are  $d_* \log n + a\sqrt{\log n}, a \in \mathbb{R}$ . Here,  $d_* = 1/\varphi(t_*)$ , where  $t_*$  is the solution of the transcendental equation  $2t_*\varphi'(t_*) = \varphi(t_*)$ .

1.5.3. Binomial

If some distribution satisfies the superlogarithmic or the logarithmic assumptions, then the same is true for its convolution powers. More precisely, this means the following. Let  $X_1, X_2, \dots$  be i.i.d. random variables with distribution function  $F(z)$ . Let  $\tilde{X}_1, \tilde{X}_2, \dots$  be i.i.d. random variables with distribution function  $F^{*m}(z/\sqrt{m})$ , where  $m \in \mathbb{N}$  is fixed, and  $F^{*m}$  denotes the  $m$ -th convolution power of  $F$ . Then, the cumulant generating function of  $\tilde{X}_k$  is given by  $\tilde{\varphi}(t) = m\varphi(t/\sqrt{m})$ . The equality

$$\tilde{\psi}(\sqrt{mt}) := \frac{\tilde{\varphi}(\sqrt{mt})}{(\sqrt{mt})^2/2} = \frac{\varphi(t)}{t^2/2} = \psi(t)$$

entails that if  $X_1$  satisfies the superlogarithmic or the logarithmic conditions, then the same holds for the variable  $\tilde{X}_1$ . For example, our results on the Bernoulli distributions imply that the binomial distribution  $\text{Bin}(k, p)$  (after standardization) belongs to the superlogarithmic case for  $p \geq 1/2$  and to the logarithmic case for  $p < 1/2$ .

1.5.4. Uniform

Let  $X_1, X_2, \dots$  be i.i.d. random variables with uniform distribution on the interval  $[-\sqrt{3}, \sqrt{3}]$ , so that  $\mathbb{E}X_k = 0$  and  $\text{Var}X_k = 1$ . We have

$$\varphi(t) = \log \frac{\sinh(\sqrt{3}t)}{\sqrt{3}t} = \frac{t^2}{2} - \frac{t^4}{20} + o(t^4), \quad t \rightarrow 0.$$

We are in the superlogarithmic case with  $q = 4$ . To see this note that

$$\frac{\sinh(\sqrt{3}t)}{\sqrt{3}t} - e^{t^2/2} = \sum_{k=2}^{\infty} t^{2k} \left( \frac{3^k}{(2k+1)!} - \frac{1}{2^k k!} \right).$$

All coefficients are negative, as one easily verifies by induction. The optimal lengths are  $a \log^2 n$ ,  $a > 0$ .

1.5.5. Gamma, negative binomial, Poisson

The former two distributions (including *exponential* and *geometric* as special cases) are covered by [Theorems 1.6](#) and [1.7](#). The Poisson distribution is covered by [Theorem 1.6](#). Although it does not satisfy the assumptions of [Theorem 1.7](#), it is easy to check that the conclusion of this theorem remains valid in the Poisson case. The square root normalization in the definition of  $\mathbf{M}_n$ , see (1), is thus not natural for these distributions. See [\[43,41\]](#) for alternative normalizations.

1.6. Remarks

We sketch some possible extensions and modifications of our results.

1.6.1.  $\mathbf{M}_n$  versus  $\mathbf{M}_n^2$

In order to simplify the formulas, we stated our results for  $\mathbf{M}_n^2$  instead of  $\mathbf{M}_n$ . It is easy to translate everything to  $\mathbf{M}_n$ : if  $\mathbf{M}_n^2 - a_n$  converges weakly to some distribution  $G$  for some sequence  $a_n \rightarrow +\infty$ , then

$$2\sqrt{a_n}(\mathbf{M}_n - \sqrt{a_n}) \xrightarrow[n \rightarrow \infty]{d} G. \tag{22}$$

Here is the proof of this implication. Note that  $\mathbb{P}[\mathbf{M}_n \leq 0]$  goes to 0 as  $n \rightarrow \infty$  since it can be estimated above by  $\mathbb{P}[\mathbf{U}_n \leq 0]$ , where  $\mathbf{U}_n = \max\{X_1, \dots, X_n\}$ . Hence, for every  $\tau \in \mathbb{R}$ ,

$$\mathbb{P} \left[ \mathbf{M}_n \leq \sqrt{a_n} + \frac{\tau}{2\sqrt{a_n}} \right] = \mathbb{P} \left[ \mathbf{M}_n^2 \leq a_n + \tau + \frac{\tau^2}{4a_n} \right] + o(1), \quad n \rightarrow \infty.$$

By our assumption, the right-hand side goes to  $G(\tau)$ , for every  $\tau \in \mathbb{R}$  where  $G$  is continuous. This yields (22). Similar argumentation applies to [Theorems 1.2](#) and [1.5](#).

1.6.2. Hitting times

It is possible to state our main results, [Theorems 1.1](#) and [1.3](#) in terms of the hitting time

$$\mathbf{T}(u) = \min\{n \in \mathbb{N} : \mathbf{M}_n > u\}, \quad u > 0, \tag{23}$$

rather than in terms of  $\mathbf{M}_n$ . This approach was used in [\[40\]](#). There,  $\mathbf{T}(u)$  was introduced as a stopping rule for a sequential change-point detection. It turns out that  $\mathbf{T}(u)$  has limiting exponential distribution, as  $u \rightarrow \infty$ . The Gaussian case was analyzed in [\[40\]](#). In the non-Gaussian case we have the following two results.

**Proposition 1.9.** *Let the assumptions of Theorem 1.1 be fulfilled. Let  $\alpha = \frac{1}{2} \cdot \frac{q-6}{q-2}$ . Then, for every  $y > 0$ ,*

$$\lim_{u \rightarrow \infty} \mathbb{P} \left[ 2^{-\alpha} u^{2\alpha} e^{-\frac{u^2}{2}} \mathbf{T}(u) > y \right] = \exp\{-\Lambda_{q,\kappa} y\}. \tag{24}$$

**Proposition 1.10.** *Let the assumptions of Theorem 1.3 be fulfilled. For every  $y > 0$ ,*

$$\lim_{u \rightarrow \infty} \mathbb{P} \left[ e^{-\frac{u^2}{2m_*}} \mathbf{T}(u) > y \right] = \exp\{-\Theta_* y\}. \tag{25}$$

**Proof of Propositions 1.9 and 1.10.** Fix  $y > 0$ . Let  $n = 2^\alpha u^{-2\alpha} e^{u^2/2} y$  (for Proposition 1.9) or  $n = e^{u^2/(2m_*)} y$  (for Proposition 1.10). Note that  $n$  need not be an integer. With  $\tau = -\log y$ , we have, as  $u, n \rightarrow \infty$ ,

$$\frac{u^2}{2} = \log([n] \log^\alpha [n]) + \tau + o(1), \quad \text{resp.} \quad \frac{u^2}{2} = m_*(\log [n] + \tau). \tag{26}$$

Recall from Section 1.6.1 that  $\lim_{n \rightarrow \infty} \mathbb{P}[\mathbf{M}_n < 0] = 0$ . In the case of Proposition 1.9,

$$\mathbb{P} \left[ 2^{-\alpha} u^{2\alpha} e^{-\frac{u^2}{2}} \mathbf{T}(u) > y \right] = \mathbb{P}[\mathbf{M}_{[n]} \leq u] = \mathbb{P} \left[ \frac{1}{2} \mathbf{M}_{[n]}^2 \leq \frac{u^2}{2} \right] + o(1). \tag{27}$$

Taking into account (26) and applying to the right-hand side of (27) Theorem 1.1 we obtain that the limit of the right-hand side of (27) is  $e^{-\Lambda_{q,\kappa} y}$ . Proposition 1.10 is proven analogously.  $\square$

### 1.6.3. Two-sided version of $\mathbf{M}_n$

If in the signal detection problem mentioned at the beginning of the paper we do not know whether the signal has positive or negative mean, it is natural to consider  $|\mathbf{M}_n| = \max\{\mathbf{M}_n^+, \mathbf{M}_n^-\}$  as a test statistic, where

$$\mathbf{M}_n^+ = \mathbf{M}_n = \max_{0 \leq i < j \leq n} \frac{S_j - S_i}{\sqrt{j - i}}, \quad \mathbf{M}_n^- = - \min_{0 \leq i < j \leq n} \frac{S_j - S_i}{\sqrt{j - i}}.$$

Large values of  $\mathbf{M}_n^+$  (resp.,  $\mathbf{M}_n^-$ ) indicate the presence of a signal with positive (resp., negative) mean. Since  $\mathbf{M}_n^-$  is obtained from  $\mathbf{M}_n$  by the substitution  $X_k \mapsto -X_k$ , our results (under appropriate assumptions) yield limiting distributions for both  $\mathbf{M}_n^+$  and  $\mathbf{M}_n^-$ . Moreover,  $\mathbf{M}_n^+$  and  $\mathbf{M}_n^-$  become asymptotically independent as  $n \rightarrow \infty$ . We leave this fact without a proof, but note that for i.i.d. random variables it is well-known that the maximum and the minimum become asymptotically independent as the sample size goes to  $\infty$ . If the  $X_k$ 's have symmetric distribution and if  $\mathbf{M}_n$  has a limiting distribution of the form  $\exp\{-be^{-c\tau}\}$ , for some constants  $b, c > 0$ , the asymptotic independence implies that  $|\mathbf{M}_n|$  has limiting distribution of the form  $\exp\{-2be^{-c\tau}\}$ . For non-symmetric  $X_k$ , it is possible that  $\mathbf{M}_n^+$  and  $\mathbf{M}_n^-$  belong to different cases. If this happens, the case with the larger normalizing sequence determines the behavior of  $|\mathbf{M}_n|$ .

### 1.6.4. Non-unique maximum

Among the distributions satisfying (4) there are some exotic examples which are not covered by our results. For example, it is possible that the supremum of  $\psi$  (which is strictly larger than 1) is attained at several points  $t_1, \dots, t_m > 0$  simultaneously. In this case, Theorem 1.3 still

holds, but the constant  $\Theta_*$  in (15) has to be replaced by  $\Theta_1 + \dots + \Theta_m$ , where the summands  $\Theta_i$  correspond to the contributions of the different  $t_i$ 's. It is however not possible that the maximum of  $\psi$  is attained at some interval (or some set having a limit point in  $[0, \infty)$ ). This follows from the uniqueness theorem for analytic functions. (Note that  $\psi$  can be extended analytically to the right half-plane). It is also possible that the maximum of  $\psi$  is equal to 1, but is attained at  $t_1 = 0$  and some other point  $t_2 > 0$ . The first point is described by Theorem 1.1 with normalization sequence  $a_{1,n} = \log(n \log^{\frac{1}{2} \cdot \frac{q-6}{q-2}} n)$ , the second point is described by Theorem 1.3 with normalization sequence  $a_{2,n} = \log n$ . If  $q < 6$  (which is usually the case), then  $a_{2,n} - a_{1,n} \rightarrow +\infty$  and the contribution of  $t_1 = 0$  is asymptotically negligible. Our results do not cover the situation in which  $\psi(t) < 1$  for all  $t > 0$ , but  $\lim_{t \rightarrow \infty} \psi(t) = 1$ . It is, however, difficult to find a distribution with these properties.

1.6.5. Strong approximation

The first naïve attempt to obtain the limiting distribution for  $\mathbf{M}_n$  is to approximate the random walk  $S_n$  by a Gaussian random walk  $W_n$  using the strong invariance principle of Komlós–Major–Tusnády [10]. We will now explain why this approach fails. If we exclude the case in which  $S_n$  is itself Gaussian, the best possible rate of strong approximation is  $|W_n - S_n| = O(\log n)$  a.s.; see [10]. Given  $0 \leq i < j \leq n$  with  $l = j - i$  we obtain for the difference

$$\frac{W_j - W_i}{\sqrt{j - i}} - \frac{S_j - S_i}{\sqrt{j - i}}$$

the estimate  $O(\log n / \sqrt{l})$ . If we want to apply this to show that the weak limit theorem satisfied by  $\mathbf{M}_n$  has the same form in the Gaussian and in the non-Gaussian case, the approximation error should be of smaller order than the fluctuations of  $\mathbf{M}_n$ , which are of order  $1/\sqrt{\log n}$ ; see, e.g., (3). Thus, we obtain a sufficiently accurate strong approximation if  $l$  is of larger order than  $\log^3 n$ . However, our results show that the behavior of  $\mathbf{M}_n$  is determined by the intervals of length at most  $O(\log^3 n)$ . In this domain the strong approximation is too inaccurate. The best one can prove using a direct strong approximation argument is the following result: for any sequence  $l_n$  such that  $l_n / \log^3 n \rightarrow \infty$  but  $l_n / n \rightarrow 0$ , there is a sequence  $b_n$  not depending on the distribution of  $X_1$  such that  $\mathbf{M}_n^2(l_n, n) - b_n$  converges to the Gumbel distribution. In the Gaussian case this can be proved by the methods of [21]. Then, the strong approximation implies that the same limit theorem holds for any  $X_1$  satisfying (4).

2. Notation and strategy of the proof

First we fix some notation which will be used throughout the paper. Let  $\{X_k, k \in \mathbb{Z}\}$  be non-degenerate i.i.d. random variables with  $\mathbb{E}X_k = 0, \text{Var}X_k = 1$ . We always assume that (4) holds. Define the two-sided random walk  $\{S_k, k \in \mathbb{Z}\}$  by

$$S_k = X_1 + \dots + X_k, \quad S_0 = 0, \quad S_{-k} = X_{-1} + \dots + X_{-k}, \quad k \in \mathbb{N}.$$

Consider the set  $\mathbb{I} = \{(i, j) \in \mathbb{Z}^2 : i < j\}$ . Our main object of study is the *standardized increments random field*  $\mathcal{Z} = \{\mathcal{Z}_{i,j}, (i, j) \in \mathbb{I}\}$  defined by

$$\mathcal{Z}_{i,j} = \frac{S_j - S_i}{\sqrt{j - i}}, \quad (i, j) \in \mathbb{I}. \tag{28}$$

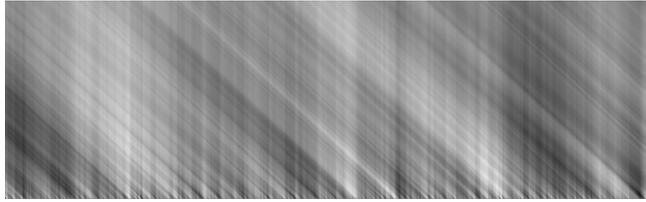


Fig. 3. A realization of the random field  $Z_{i,i+l}$  for  $1 \leq i \leq 3000$  (on the horizontal axis) and  $1 \leq l \leq 1000$  (on the vertical axis). Dark points correspond to large values of the field.

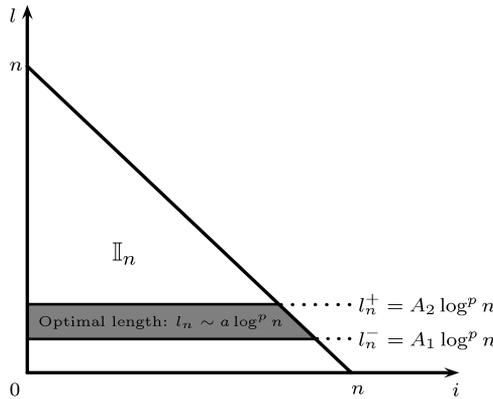


Fig. 4. The set  $\mathbb{I}_n$ . On the horizontal axis, we put the initial point  $i$  of the interval  $(i, j)$ . On the vertical axis, we put its length  $l = j - i$ . The strip shows the intervals whose length  $l$  is between  $l_n^- = A_1 \log^p n$  and  $l_n^+ = A_2 \log^p n$  (the optimal lengths in the superlogarithmic case).

See Fig. 3 for a realization of this random field. Elements  $(i, j) \in \mathbb{I}$  will be called *intervals*. Any interval  $(i, j) \in \mathbb{I}$  will be identified with  $Z_{i,j}$ , the corresponding standardized increment, as well as with the set  $\{i + 1, \dots, j\}$ . We call  $l := j - i \in \mathbb{N}$  the *length* of  $(i, j)$ .

We are interested in the random variable  $\mathbf{M}_n = \max_{(i,j) \in \mathbb{I}_n} Z_{i,j}$ , where  $\mathbb{I}_n \subset \mathbb{I}$  is a set of intervals given by

$$\mathbb{I}_n = \{(i, j) \in \mathbb{Z}^2 : 0 \leq i < j \leq n\}, \quad n \in \mathbb{N}; \tag{29}$$

see Fig. 4. Clearly,  $\mathbf{M}_n$  is a maximum of random variables, but these are neither independent, nor identically distributed. We prove our results by a careful extreme-value analysis of the field  $\mathcal{Z}$ . We will take an appropriate threshold  $u_n \rightarrow \infty$  and compute the limit of the *exceedance probability*  $\mathbb{P}[\mathbf{M}_n > u_n]$ . Our method can be subdivided into 3 steps.

STEP 1. We start by computing the *individual probability*  $\mathbb{P}[Z_{i,j} > u_n]$  for intervals  $(i, j)$  having “optimal length”  $l_n$ . The optimal length is chosen as in Section 1.1. For example, in the superlogarithmic case the optimal lengths are of the form  $a \log^p n$ ,  $a > 0$ , meaning that (as we will prove in Section 4.5) the contribution of all other lengths is negligible:

$$\begin{aligned} & \lim_{A_1 \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}[\mathbf{M}_n(1, A_1 \log^p n) > u_n] \\ &= \lim_{A_2 \rightarrow +\infty} \limsup_{n \rightarrow \infty} \mathbb{P}[\mathbf{M}_n(A_2 \log^p n, n) > u_n] = 0. \end{aligned}$$

To compute the individual probability, we need classical limit theorems on large and moderate deviations which will be recalled in Section 3. The key results of Step 1 are Lemma 4.2 (in the superlogarithmic case) and Lemma 5.2 (in the logarithmic case).

STEP 2. In the second step we compute the *local probability*  $\mathbb{P}[\max_{(i,j) \in \mathbb{T}_n} \mathcal{Z}_{i,j}]$ , where  $\mathbb{T}_n \subset \mathbb{I}$  is a discrete square with side length of order  $q_n$ . Here,  $q_n$  is chosen to be the “extremal decorrelation length” of  $\mathcal{Z}$ . This means that the exceedances of  $\mathcal{Z}$  at two points of  $\mathbb{I}$  at distance of order  $q_n$  retain non-trivial asymptotic dependence in the large  $n$  limit (that is, they become neither completely dependent, nor completely independent). There is a way to characterize  $q_n$  using the language of the Poisson clumping heuristic; see [1]. The intervals  $(i, j)$  for which  $\mathcal{Z}_{i,j} > u_n$  form small clumps distributed randomly in  $\mathbb{I}$ . Then, the linear size of these clumps is of order  $q_n$ . From this part of the proof it will be clear why it is not possible to choose the “optimal length” as large as possible (say, of order  $n$ ). Namely, long intervals are strongly dependent (meaning that the extremal decorrelation length is large there); see also Fig. 3. Therefore, long intervals make only a small contribution to  $\mathbf{M}_n$ . The key results of Step 2 are Lemma 4.3 (in the superlogarithmic case) and Lemma 5.3 (in the logarithmic case).

STEP 3. The final step is to compute the exceedance probability over a domain of size much larger than  $q_n$ . Such domain can be decomposed into many small domains of size  $q_n$ , see Fig. 6, and the exceedance events over these small domains are asymptotically independent due to the extremal decorrelation. The asymptotic independence is shown by estimating the *double sum* appearing in the Bonferroni inequality. Thus, we can apply the Poisson limit theorem for weakly dependent events. The key steps of the third step are Lemma 4.8 (in the superlogarithmic case) and Lemma 5.4 (in the logarithmic case).

On a rigorous level, there are several (closely related) powerful methods to analyze extremes of random fields; see [28,36,4,1]. We use a modification of the double sum method of Pickands [34]; see also Leadbetter et al. [28, Chapter 12], Piterbarg [36, Chapter D]. Originally, the method was used to analyze extremes of Gaussian processes, but it can be applied to non-Gaussian scan statistics as well; see [6,37,35,8]. These references deal with fixed window size, for an example with variable window size see [24]. A related method was also used by [41].

Throughout the paper  $C_1, C_2, \dots$  and  $c_1, c_2, \dots$  are positive constants which may change from line to line. They may depend on the distribution of  $X_1$  and parameters specified in the text. We write  $a_n \sim b_n$  if  $\lim_{n \rightarrow \infty} a_n/b_n = 1$ . Let  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ .

### 3. Results on large and moderate deviations

In our proofs we will make a heavy use of the exact asymptotic results for the probabilities of large and moderate deviations of sums of i.i.d. random variables. Recall from (4) that  $\varphi$  is the cumulant generating function of the  $X_k$ 's. Define the Legendre–Fenchel transform of  $\varphi$ :

$$I(s) = \sup_{t \geq 0} (st - \varphi(t)), \quad s \geq 0. \quad (30)$$

Let  $s_\infty = \sup\{s \in \mathbb{R} : \mathbb{P}[X_1 \leq s] < 1\} > 0$  be the right endpoint of  $X_1$ . Then,  $I$  is a finite, strictly convex, strictly increasing, infinitely differentiable function on  $[0, s_\infty)$ , and  $I(s) = +\infty$  for  $s > s_\infty$ . Also, we have  $s_\infty = \limsup_{t \rightarrow \infty} \varphi(t)/t$ .

The next theorem on the probability of “moderate” deviations proved originally by Cramér [9] has been subsequently strengthened by Feller [15], Petrov [31] and Höglund [18]; see also [33] and [19].

**Theorem 3.1.** Let  $x_k > 0$  be a sequence such that  $x_k \rightarrow \infty$  but  $x_k = o(\sqrt{k})$  as  $k \rightarrow \infty$ . Then, as  $k \rightarrow \infty$ ,

$$\mathbb{P} \left[ \frac{S_k}{\sqrt{k}} > x_k \right] \sim \frac{1}{\sqrt{2\pi x_k}} \exp \left\{ -kI \left( \frac{x_k}{\sqrt{k}} \right) \right\}. \tag{31}$$

Often, it is more convenient to introduce the Cramér series  $\lambda(y) = y^{-3}(\frac{y^2}{2} - I(y))$  and state (31) in the following equivalent form (which is valid without the requirement  $x_k \rightarrow \infty$ ):

$$\mathbb{P} \left[ \frac{S_k}{\sqrt{k}} > x_k \right] \sim \bar{\Phi}(x_k) \cdot \exp \left\{ \frac{x_k^3}{\sqrt{k}} \cdot \lambda \left( \frac{x_k}{\sqrt{k}} \right) \right\}. \tag{32}$$

Here,  $\bar{\Phi}(x)$  is the tail function of the standard normal distribution. Relations (31) and (32) are equivalent since  $\bar{\Phi}(x) \sim \frac{1}{\sqrt{2\pi x}} e^{-x^2/2}$  as  $x \rightarrow \infty$ . Note that for  $x_k$  being constant the central limit theorem is recovered.

The next theorem due to Bahadur and Ranga Rao [3] and Petrov [32] deals with the probabilities of “large” deviations of  $S_k$ .

**Theorem 3.2.** Assume that  $X_1$  is non-lattice. Let  $x_k > 0$  be a sequence such that  $x_k \sim \alpha\sqrt{k}$  for some  $\alpha > 0$ , as  $k \rightarrow \infty$ . Then, as  $k \rightarrow \infty$ ,

$$\mathbb{P} \left[ \frac{S_k}{\sqrt{k}} > x_k \right] \sim \frac{1}{\sqrt{2\pi k}} \frac{1}{I'(\alpha)\sigma(\alpha)} \exp \left\{ -kI \left( \frac{x_k}{\sqrt{k}} \right) \right\}. \tag{33}$$

Here,  $\sigma^2(\alpha) = \varphi''(I'(\alpha))$ .

For lattice variables the theorem should be modified; see [32]. In fact, Theorems 3.1 and 3.2 can be included as special cases in a general result; see [18]. We will need just the following inequality; see [31,18]. It is valid both for “moderate” and “large” deviations, both in the lattice and in the non-lattice case.

**Theorem 3.3.** For every  $A \in (0, s_\infty)$  there is a constant  $C = C(A)$  such that for all  $k \in \mathbb{N}$ ,  $x \in (0, A\sqrt{k})$ ,

$$\mathbb{P} \left[ \frac{S_k}{\sqrt{k}} > x \right] \leq \frac{C}{x} \exp \left\{ -kI \left( \frac{x}{\sqrt{k}} \right) \right\}. \tag{34}$$

The next lemma is elementary and well-known. It is weaker than Theorem 3.3, but valid without restriction on  $x$ . Instead of (4) we assume that  $\varphi(t) = \log \mathbb{E}e^{tX_1}$  is finite on  $[0, t_\infty)$ , for some  $t_\infty > 0$ .

**Lemma 3.4.** For every  $k \in \mathbb{N}$  and  $x > 0$ , we have

$$\mathbb{P} \left[ \frac{S_k}{\sqrt{k}} > x \right] \leq \exp \left\{ -kI \left( \frac{x}{\sqrt{k}} \right) \right\}.$$

**Proof.** By Markov’s inequality  $\mathbb{P}[S_k > \sqrt{k}x] \leq e^{-t\sqrt{k}x+k\varphi(t)}$ , for every  $t \geq 0$ . Take the minimum over  $t \geq 0$ .  $\square$

### 4. Proof in the superlogarithmic case

In this section we prove [Theorems 1.1](#) and [1.2](#). It will be convenient to pass from conditions [\(7\)](#) and [\(8\)](#) to their Legendre–Fenchel conjugates. We will assume that for every  $\varepsilon > 0$ ,

$$\inf_{s \geq \varepsilon} \frac{I(s)}{s^2/2} > 1. \tag{35}$$

We also need the Taylor expansion of  $I$  at 0: with  $\kappa > 0$ ,

$$I(s) = \frac{s^2}{2} + \kappa s^q + o(s^q), \quad s \downarrow 0. \tag{36}$$

**Proposition 4.1.** *Assume that [\(4\)](#) holds. Then, conditions [\(7\)](#) and [\(35\)](#) are equivalent. Also, conditions [\(8\)](#) and [\(36\)](#) are equivalent.*

**Proof.** Assume that [\(7\)](#) holds. Let  $\varepsilon > 0$  be arbitrary. By [\(7\)](#) we can find  $c = c(\varepsilon) < 1$  such that  $\varphi(t) \leq ct^2/2$  for all  $t \geq \varepsilon$ . It follows that for every  $s \geq \varepsilon$ ,

$$I(s) = \sup_{t \geq 0} (st - \varphi(t)) \geq \sup_{t \geq \varepsilon} (st - \varphi(t)) \geq \sup_{t \geq \varepsilon} \left( st - \frac{ct^2}{2} \right) = \frac{s^2}{2c}.$$

Note that the last equality holds since the supremum of  $st - \frac{ct^2}{2}$  is attained at  $t = \frac{s}{c} > \varepsilon$ . It follows that [\(35\)](#) holds. The proof that [\(35\)](#) implies [\(7\)](#) is analogous.

Assume now that [\(8\)](#) holds. Note that  $\varphi$  is analytic in a neighborhood of zero. Taking the derivative, we obtain  $\varphi'(t) = t - q\kappa t^{q-1} + o(t^{q-1})$ . Now, by Legendre–Fenchel duality,  $I'$  is the inverse function of  $\varphi'$ . Taking the inverse function of  $\varphi'$  we obtain  $I'(s) = s + q\kappa s^{q-1} + o(s^{q-1})$ . Integrating, we obtain [\(36\)](#). The proof of the converse implication is analogous.  $\square$

Recall the notation  $p = \frac{q}{q-2}$ . Fix some  $\tau \in \mathbb{R}$  and define a normalizing sequence  $u_n = u_n(\tau) > 0$  by

$$u_n^2 = 2 \log(n \log^{\frac{3}{2}-p} n) + 2\tau. \tag{37}$$

Throughout the remainder of Section 4 we assume that conditions [\(35\)](#) and [\(36\)](#) are satisfied. Sections 4.1–4.4 are devoted to the proof of [Theorem 1.2](#). In Section 4.5 we complete the proof of [Theorem 1.1](#).

#### 4.1. Individual probability

The first step is to compute the probability that the random variable  $Z_{i,j}$  exceeds some large threshold at some individual point  $(i, j) \in \mathbb{I}$ . We will consider intervals  $(i, j)$  whose length  $l_n$  is optimal, that is  $l_n \sim a \log^p n$ , for some  $a > 0$ .

**Lemma 4.2.** *Let  $l_n \in \mathbb{N}$  be a sequence such that  $a := \lim_{n \rightarrow \infty} l_n / \log^p n \in (0, \infty)$ . Let  $s \in \mathbb{R}$  be fixed. Then, as  $n \rightarrow \infty$ ,*

$$P_n(s) := \mathbb{P} \left[ \frac{S_{l_n}}{\sqrt{l_n}} > u_n - \frac{s}{u_n} \right] \sim \frac{1}{2\sqrt{\pi}} \cdot e^{s - \kappa 2^{\frac{q}{2}} a^{-\frac{q-2}{2}}} \cdot \frac{e^{-\tau}}{n \log^{2-p} n}.$$

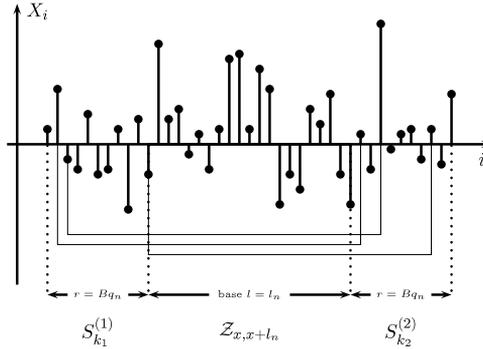


Fig. 5. The structure of the set  $\mathbb{T}_n = \mathbb{T}_{Bq_n}(x, x + l_n)$ . All the intervals share the common base  $(x, x + l_n)$ .

**Proof.** By (37), we have  $u_n \sim \sqrt{2 \log n}$ . Note that  $u_n - \frac{s}{u_n} = o(\sqrt{l_n})$  since  $p > 1$ . By Theorem 3.1,

$$P_n(s) \sim \frac{1}{\sqrt{2\pi}u_n} \exp \left\{ -l_n I \left( \frac{u_n - \frac{s}{u_n}}{\sqrt{l_n}} \right) \right\}. \tag{38}$$

We evaluate the term under the sign of the exponential. Since  $l_n \sim a \log^p n$  and  $p = \frac{q}{q-2}$ , we have

$$\left( \frac{u_n - \frac{s}{u_n}}{\sqrt{l_n}} \right)^q = 2^{\frac{q}{2}} a^{-\frac{q-2}{2}} l_n^{-1} + o(l_n^{-1}).$$

Using the Taylor expansion in (36), we obtain

$$l_n I \left( \frac{u_n - \frac{s}{u_n}}{\sqrt{l_n}} \right) = \frac{1}{2} \left( u_n - \frac{s}{u_n} \right)^2 + \kappa 2^{\frac{q}{2}} a^{-\frac{q-2}{2}} + o(1).$$

Inserting this into (38) yields

$$P_n(s) \sim \frac{e^s}{\sqrt{2\pi}u_n} \exp \left\{ -\frac{u_n^2}{2} - \kappa 2^{\frac{q}{2}} a^{-\frac{q-2}{2}} \right\}. \tag{39}$$

To complete the proof of Lemma 4.2 note that  $e^{u_n^2/2} = e^{\tau n \log^{\frac{3}{2}-p} n}$  by (37).  $\square$

#### 4.2. Local probability

The next step is to compute the exceedance probability over a small discrete square in the space of intervals. Given an interval  $(x, y) \in \mathbb{I}$  of length  $l := y - x$  and a “length fluctuation”  $r \in \mathbb{N}$ , let  $\mathbb{T}_r(x, y)$  be the set of all intervals  $(i, j) \in \mathbb{I}$  satisfying

$$x - r < i \leq x \quad \text{and} \quad y \leq j < y + r.$$

Note that all intervals from the set  $\mathbb{T}_r(x, y)$  are extensions of the “base” interval  $(x, y)$ ; see Fig. 5. We can view  $\mathbb{T}_r(x, y)$  as a discrete square with side length  $r$  in the grid  $\mathbb{I} \subset \mathbb{Z}^2$ . The base interval  $(x, y)$  corresponds to the right bottom vertex of this square. The cardinality of  $\mathbb{T}_r(x, y)$  is  $r^2$ . Fix any sequence  $q_n \in \mathbb{N}$  satisfying  $q_n \sim \log^{p-1} n$  as  $n \rightarrow \infty$ . Note that  $q_n \rightarrow \infty$  since  $p > 1$ .

**Lemma 4.3.** Let  $l_n \in \mathbb{N}$  be a sequence such that  $a := \lim_{n \rightarrow \infty} l_n / \log^p n \in (0, \infty)$ . Fix  $B \in \mathbb{N}$  and  $x \in \mathbb{Z}$ . Write  $\mathbb{T}_n = \mathbb{T}_{Bq_n}(x, x + l_n)$ ; see Fig. 5. Then, as  $n \rightarrow \infty$ , we have

$$Q_n := \mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_n} Z_{i,j} > u_n \right] \sim P_n(0) \cdot \left\{ 1 + H^2 \left( \frac{B}{a} \right) \right\}, \tag{40}$$

where  $P_n(0)$  is as in Lemma 4.2, the function  $H : (0, \infty) \rightarrow (0, \infty)$  is defined by

$$H(B) = \mathbb{E} \left[ \sup_{t \in [0, B]} e^{\sqrt{2}W(t)-t} \right], \quad B > 0, \tag{41}$$

and  $\{W(t), t \geq 0\}$  is a standard Brownian motion.

**Proof.** The idea of the proof is to represent the field  $Z_{i,j}$  as its value at the base interval  $(x, x + l_n)$  plus some incremental process. In order to exceed the level  $u_n$  over  $\mathbb{T}_n$  either the value  $Z_{x,x+l_n}$  at the base interval should be larger than  $u_n$ , or this value should be of the form  $u_n - u_n^{-1}s, s \geq 0$ , and the supremum of the incremental process should be larger than  $u_n^{-1}s$ . We will show that the incremental process converges to the sum of two independent Brownian motions.

Let  $X_k^{(1)}, X_k^{(2)}, k \in \mathbb{N}$ , be i.i.d. random variables with the same distribution as the  $X_k$ 's, and are independent of the  $X_k$ 's. Define two independent random walks

$$S_k^{(i)} = X_1^{(i)} + \dots + X_k^{(i)}, \quad k \in \mathbb{N}_0, i = 1, 2. \tag{42}$$

Let  $V_n$  be a random variable defined by  $Z_{x,x+l_n} = u_n - u_n^{-1}V_n$ . Then, since any interval from  $\mathbb{T}_n$  has the form  $(x - k_1, x + l_n + k_2)$  with some integers  $0 \leq k_1, k_2 < Bq_n$ , we have

$$Q_n = \mathbb{P} \left[ \max_{0 \leq k_1, k_2 < Bq_n} \frac{(u_n - \frac{V_n}{u_n})\sqrt{l_n} + S_{k_1}^{(1)} + S_{k_2}^{(2)}}{\sqrt{l_n + k_1 + k_2}} > u_n \right].$$

By taking  $k_1 = k_2 = 0$  we see that the maximum on the right-hand side exceeds  $u_n$  if  $V_n < 0$ . Note that  $\mathbb{P}[V_n < 0] = P_n(0)$ , the probability which was evaluated in Lemma 4.2. The random variables  $V_n, S_{k_1}^{(1)}, S_{k_2}^{(2)}$  are independent. Conditioning on  $V_n = s$  and integrating over  $s \geq 0$ , we obtain

$$Q_n = P_n(0) + \int_0^\infty G_n(s) d\mu_n(s), \tag{43}$$

where  $\mu_n$  is the probability distribution of  $V_n$  and  $G_n$  is a non-increasing function defined by

$$G_n(s) = \mathbb{P} \left[ \max_{0 \leq k_1, k_2 < Bq_n} \left( S_{k_1}^{(1)} + S_{k_2}^{(2)} - u_n(\sqrt{l_n + k_1 + k_2} - \sqrt{l_n}) - \frac{s\sqrt{l_n}}{u_n} \right) > 0 \right].$$

By Lemma 4.2, for every  $s \geq 0$ ,

$$\lim_{n \rightarrow \infty} \frac{\mu_n([0, s])}{P_n(0)} = \lim_{n \rightarrow \infty} \frac{P_n(s)}{P_n(0)} = e^s. \tag{44}$$

We are going to compute  $\lim_{n \rightarrow \infty} G_n(s)$  for  $s \geq 0$ . In fact, to be able to use Lemma 4.5, see below, we need a slightly stronger result. Let  $s_n$  be any sequence converging to  $s$ . We will compute  $\lim_{n \rightarrow \infty} G_n(s_n)$ . We have

$$G_n(s_n) = \mathbb{P} \left[ \sup_{t_1, t_2 \in [0, B]} (W_n^{(1)}(t_1) + W_n^{(2)}(t_2) - f_n(t_1, t_2)) > 0 \right], \tag{45}$$

where  $\{W_n^{(i)}(t), t \in [0, B]\}, i = 1, 2$ , are stochastic processes and  $f_n : [0, B]^2 \rightarrow \mathbb{R}$  is a function given by

$$W_n^{(i)}(t) = \frac{S_{tq_n}^{(i)}}{\sqrt{q_n}}, \quad f_n(t_1, t_2) = \frac{u_n(\sqrt{l_n + t_1q_n + t_2q_n} - \sqrt{l_n}) + \frac{s_n\sqrt{l_n}}{u_n}}{\sqrt{q_n}}$$

if  $t, t_1, t_2 \in q_n^{-1}\mathbb{Z} \cap [0, B)$ , and by linear interpolation otherwise. On the last interval of length  $q_n^{-1}$  we agree to use constant interpolation. Recall that  $u_n \sim \sqrt{2 \log n}, l_n \sim a \log^p n, q_n \sim \log^{p-1} n$ , as  $n \rightarrow \infty$ . Elementary calculus shows that uniformly in  $t_1, t_2 \in [0, B]$ ,

$$f_n(t_1, t_2) \rightarrow f(t_1, t_2) := \frac{t_1 + t_2}{\sqrt{2a}} + \sqrt{\frac{a}{2}}s \quad \text{as } n \rightarrow \infty.$$

Given a compact metric space  $K$  let  $C(K)$  be the space of continuous functions on  $K$  endowed with the sup-metric. By Donsker’s invariance principle, as  $n \rightarrow \infty$ , the processes  $\{W_n^{(i)}(t), t \in [0, B]\}, i = 1, 2$ , converge weakly on the space  $C[0, B]$  to two independent standard Brownian motions  $\{W_i(t), t \in [0, B]\}, i = 1, 2$ . Consider the map

$$\Psi : C[0, B] \times C[0, B] \times C([0, B]^2) \rightarrow \mathbb{R}$$

defined by

$$\Psi(w_1, w_2, f) = \sup_{t_1, t_2 \in [0, B]} (w_1(t_1) + w_2(t_2) - f(t_1, t_2)).$$

The map  $\Psi$  is continuous in the product topology. By the continuous mapping theorem, see Theorem 3.27 in [25], it follows that the sequence of random variables

$$\Psi(W_n^{(1)}, W_n^{(2)}, f_n) = \sup_{t_1, t_2 \in [0, B]} (W_n^{(1)}(t_1) + W_n^{(2)}(t_2) - f_n(t_1, t_2))$$

converges in distribution to the random variable

$$\Psi(W_1, W_2, f) = \sup_{t_1, t_2 \in [0, B]} \left( W_1(t_1) + W_2(t_2) - \frac{t_1 + t_2}{\sqrt{2a}} - \sqrt{\frac{a}{2}}s \right).$$

By the scaling property of the Brownian motion, the latter variable has the same distribution as  $\sqrt{\frac{a}{2}}(M_1 + M_2 - s)$ , where  $M_1, M_2$  are two independent copies of the random variable

$$M := \sup_{t \in [0, a^{-1}B]} (\sqrt{2}W(t) - t) \geq 0.$$

Here,  $\{W(t), t \geq 0\}$  is a standard Brownian motion. Note that the random variable  $M$  (and hence,  $M_1 + M_2$ ) has continuous distribution function. It follows that for every sequence  $s_n$  converging to  $s \geq 0$ ,

$$\lim_{n \rightarrow \infty} G_n(s_n) = \mathbb{P}[M_1 + M_2 > s]. \tag{46}$$

Taking (44) and (46) together, we obtain, formally,

$$\lim_{n \rightarrow \infty} \int_0^\infty G_n(s) \frac{d\mu_n(s)}{P_n(0)} = \int_0^\infty \mathbb{P}[M_1 + M_2 > s] e^s ds = \mathbb{E}e^{M_1+M_2} = (\mathbb{E}e^M)^2. \tag{47}$$

Recalling (43) we obtain the statement of the lemma. The first equality in (47) will be justified in Remark 4.7 after some technical preparations have been done.  $\square$

The next lemma gives a somewhat weaker statement than Lemma 4.3, but this statement is valid under more general assumptions. The lemma will be used later to estimate the exceedance probability over the non-optimal lengths. Essentially, it states that the local exceedance probability can be estimated by the individual exceedance probability times some constant.

**Lemma 4.4.** Fix constants  $B_1, B_2 > 0$ . Then, for all  $x \in \mathbb{Z}, l, r \in \mathbb{N}$  and all  $u > 0$  such that  $B_1 l > u^2$  and  $r < B_2 l u^{-2}$ , we have

$$Q(l, r, u) := \mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_r(x,x+l)} \frac{S_j - S_i}{\sqrt{l}} > u \right] < C_1 u^{-1} \exp \left\{ -\frac{u^2}{2} - C_2 u^q l^{-\frac{q-2}{2}} \right\},$$

where the constants  $C_1$  and  $C_2$  depend on  $B_1$  and  $B_2$  but do not depend on  $x, l, r, u$ .

**Proof.** Define two independent random walks  $S_k^{(i)}, i = 1, 2$  as in (42). Let  $V_{l,u}$  be a random variable defined by  $Z_{x,x+l} = u - u^{-1} V_{l,u}$ . Then, since any interval from  $\mathbb{T}_r(x, x + l)$  has the form  $(x - k_1, x + l + k_2)$  with some integers  $0 \leq k_1, k_2 < r$ , we have

$$Q(l, r, u) = \mathbb{P} \left[ \max_{0 \leq k_1, k_2 < r} \frac{S_{k_1}^{(1)} + S_{k_2}^{(2)}}{\sqrt{l}} > \frac{V_{l,u}}{u} \right].$$

Taking  $k_1, k_2 = 0$  we see that the maximum on the right-hand side is non-negative. Conditioning on  $V_{l,u} = s$  and considering the cases  $s < 0$  and  $s \geq 0$  separately, we obtain

$$Q(l, r, u) = F_{l,u}(0) + \int_0^\infty G_{l,r,u}(s) dF_{l,u}(s), \tag{48}$$

where  $F_{l,u}$  is the distribution function of  $V_{l,u}$  and

$$G_{l,r,u}(s) = \mathbb{P} \left[ \max_{0 \leq k_1, k_2 < r} (S_{k_1}^{(1)} + S_{k_2}^{(2)}) > \frac{s\sqrt{l}}{u} \right] \leq 2\mathbb{P} \left[ \max_{0 \leq k < r} S_k > \frac{s\sqrt{l}}{2u} \right]. \tag{49}$$

We estimate  $F_{l,u}(s)$  for  $s \in [0, \frac{3}{4}u^2]$ . Write  $v = u - \frac{s}{u}$ . By the assumption  $v < \sqrt{B_1 l}$  we can apply Theorem 3.3 and (35), (36) to get

$$F_{l,u}(s) = \mathbb{P} \left[ \frac{S_l}{\sqrt{l}} \geq v \right] \leq c_1 v^{-1} \exp \left\{ -\frac{v^2}{2} - c_2 v^q l^{-\frac{q-2}{2}} \right\}.$$

Here, the constants  $c_1, c_2, \dots$  depend on  $B_1, B_2$  but do not depend on  $x, l, r, u$ . If  $s \in [0, \frac{3}{4}u^2]$ , then  $v \geq u/4$  and we obtain

$$F_{l,u}(s) \leq c_3 u^{-1} e^s \exp \left\{ -\frac{u^2}{2} - c_4 u^q l^{-\frac{q-2}{2}} \right\}. \tag{50}$$

It is however easy to see that this inequality continues to hold for  $s \geq 3u^2/4$ . Indeed, if  $c_4$  is sufficiently small, then the assumption  $B_1 l > u^2$  implies that  $c_4 u^q l^{-(q-2)/2} \leq u^2/8$ . Hence, if  $c_3$  is sufficiently large, the right-hand side of (50) is greater than 1 and (50) holds.

We estimate  $G_{l,r,u}(s)$  for  $s \geq 0$ . Applying to the right-hand side of (49) the inequality stated in Theorem 2.4 on p. 52 in [33], we obtain

$$G_{l,r,u}(s) \leq 4\mathbb{P} \left[ S_r > \frac{s\sqrt{l}}{2u} - \sqrt{2r} \right] \leq 4\mathbb{P} \left[ \frac{S_r}{\sqrt{r}} > c_5 s - \sqrt{2} \right] \leq 4e^{-rI\left(\frac{c_5 s - \sqrt{2}}{\sqrt{r}}\right)}.$$

In the second inequality, we used the assumption  $r < B_2lu^{-2}$ . In the third inequality we used Lemma 3.4. From (35) we obtain

$$G_{l,r,u}(s) \leq c_6 e^{-c_7 s^2}. \tag{51}$$

Strictly speaking, this is valid only as long as  $c_5 s \geq \sqrt{2}$ , however, we can choose the constant  $c_6$  so large that (51) continues to hold in the case  $c_5 s \leq \sqrt{2}$ . It follows from (48), (50), (51) that

$$\begin{aligned} Q(l, r, u) &\leq F_{l,u}(0) + \sum_{k=0}^{\infty} G_{l,r,u}(k) F_{l,u}(k+1) \\ &\leq c_8 \left( 1 + \sum_{k=0}^{\infty} e^{-c_7 k^2} e^k \right) u^{-1} \exp \left\{ -\frac{u^2}{2} - c_4 u^q l^{-\frac{q-2}{2}} \right\}. \end{aligned}$$

The proof of Lemma 4.4 is complete.  $\square$

**Lemma 4.5.** *Let  $\nu, \nu_n, n \in \mathbb{N}$ , be measures on  $[0, \infty)$  which are finite on compact intervals. Let  $G, G_n, n \in \mathbb{N}$ , be measurable functions on  $[0, \infty)$  which are uniformly bounded on compact intervals. Assume that*

1.  $\nu_n$  converges to  $\nu$  weakly on every interval  $[0, t], t \geq 0$ ;
2. for  $\nu$ -a.e.  $s \geq 0$  and for every sequence  $s_n \rightarrow s$  we have  $\lim_{n \rightarrow \infty} G_n(s_n) = G(s)$ ;
3.  $\lim_{T \rightarrow +\infty} \int_T^\infty |G_n| d\nu_n = 0$  uniformly over  $n \in \mathbb{N}$ .

Then,  $\lim_{n \rightarrow \infty} \int_0^\infty G_n d\nu_n = \int_0^\infty G d\nu$ .

**Proof.** Write  $F_n(s) = \nu_n([0, s])$  and  $F(s) = \nu([0, s]), s \geq 0$ . By the first assumption,  $\lim_{n \rightarrow \infty} F_n(T) = F(T)$  for all  $T \geq 0, T \notin \mathcal{D}$ , where  $\mathcal{D}$  is the set of (at most countable) discontinuities of  $F$ . By the third assumption it suffices to show that for every  $T > 0, T \notin \mathcal{D}$ ,

$$\lim_{n \rightarrow \infty} \int_0^T G_n d\nu_n = \int_0^T G d\nu. \tag{52}$$

Fix some  $T \notin \mathcal{D}$ . By the first assumption and by Skorokhod’s representation theorem we can construct (generally, dependent)  $[0, T]$ -valued random variables  $\xi_n$  and  $\xi$  on a common probability space such that  $\xi_n \rightarrow \xi$  a.s.,  $\mathbb{P}[\xi_n \in B] = \nu_n(B)/F_n(T)$  and  $\mathbb{P}[\xi \in B] = \nu(B)/F(T)$ , for every Borel set  $B \subset [0, T]$ . It follows from the second condition that  $G_n(\xi_n) \rightarrow G(\xi)$  a.s. Since for uniformly bounded random variables the a.s. convergence implies the convergence of expectations, we obtain (52).  $\square$

**Remark 4.6.** We will verify the last condition of Lemma 4.5 by showing that  $\nu_n([0, s]) \leq c_1 e^{c_2 s}$  and  $|G_n(s)| \leq c_3 e^{-c_4 s^2}$  for all (large)  $n \in \mathbb{N}$ . Then,

$$\int_T^\infty |G_n| d\nu_n \leq \sum_{k=[T]}^\infty \nu_n([k, k+1]) \sup_{s \in [k, k+1]} |G_n(s)| \leq \sum_{k=[T]}^\infty c_5 e^{c_2 k} e^{-c_4 k^2},$$

which converges to 0 uniformly in  $n \in \mathbb{N}$ , as  $T \rightarrow +\infty$ .

**Remark 4.7.** We are in position to justify the first equality in (47). The first two assumptions of Lemma 4.5, with  $\nu_n = \mu_n/P_n(0)$ , are fulfilled by (44) and (46). To verify the last assumption we use Remark 4.6. For  $s \geq 0$  we have, as established in (50) and Lemma 4.2,

$$\mu_n([0, s]) \leq P_n^{-1}(0) u_n^{-1} e^s \exp \left\{ -\frac{u_n^2}{2} - c u_n^q l_n^{-\frac{q-2}{2}} \right\} \leq c_1 e^s.$$

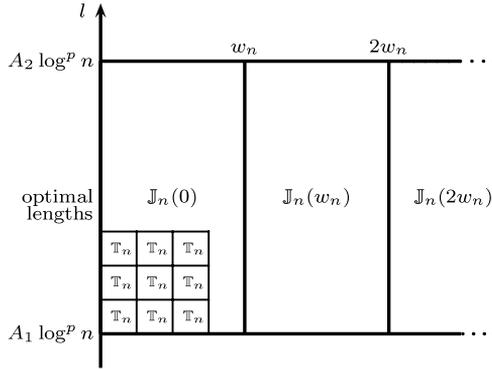


Fig. 6. On the horizontal axis we put the initial point  $i$  of the interval  $(i, j)$ , on the vertical axis we put its length  $l = j - i$ . The strip is the set of intervals whose length  $l$  is between  $l_n^-$  and  $l_n^+$ . The figure shows the decomposition of this strip into disjoint “blocks” of the form  $\mathbb{J}_n(z)$ . Lemma 4.8 computes the exceedance probability over one block  $\mathbb{J}_n(z)$  by decomposing it into disjoint squares of the form  $\mathbb{T}_n = \mathbb{T}_{Bq_n}(x, y)$ . The exceedance probability over a single square  $\mathbb{T}_n$  is computed in Lemma 4.3. In Section 4.4 it is shown that the exceedance events over different blocks of the form  $\mathbb{J}_n(z)$  are approximatively independent.

For  $G_n$  we obtain from (51) the estimate  $G_n(s) \leq G_{l_n, q_n, u_n}(s) \leq c_2 e^{-c_3 s^2}$ . Now, Remark 4.6 can be applied.

4.3. Estimating the double sum

Given  $0 < A_1 < A_2$  we define  $l_n^- = A_1 \log^p n$  and  $l_n^+ = A_2 \log^p n$ . Recall that  $q_n$  is any sequence such that  $q_n \sim \log^{p-1} n$  as  $n \rightarrow \infty$ .

**Lemma 4.8.** *Let  $w_n$  be any integer sequence such that  $w_n/q_n \rightarrow \infty$  but  $w_n = O(\log^p n)$ , as  $n \rightarrow \infty$ . For  $z \in \mathbb{Z}$  let  $\mathbb{J}_n(z)$  be the set of all intervals  $(i, j) \in \mathbb{I}$  such that  $z \leq i < z + w_n$  and  $j - i \in [l_n^-, l_n^+]$ ; see Fig. 6. Then, as  $n \rightarrow \infty$ ,*

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{J}_n(z)} \mathcal{Z}_{i,j} > u_n \right] \sim e^{-\tau} \frac{w_n}{n} \int_{A_1}^{A_2} \Lambda(a) da, \tag{53}$$

where  $\Lambda(a) = \Lambda_{q,\kappa}(a) = \frac{1}{2\sqrt{\pi}a^2} \exp\{-\kappa 2^{\frac{q}{2}} a^{-\frac{q-2}{2}}\}$ ,  $a > 0$ , is as in (12).

**Proof.** By translation invariance we may take  $z = 0$ . Take some  $B \in \mathbb{N}$  and recall that  $q_n \in \mathbb{N}$  is a sequence satisfying  $q_n \sim \log^{p-1} n$ . To get rid of the boundary effects we introduce two sequences  $\varepsilon_n$  and  $\delta_n$  such that  $\varepsilon_n/q_n \rightarrow \infty$  and  $\delta_n/q_n \rightarrow \infty$ , but  $\varepsilon_n = o(w_n)$  and  $\delta_n = o(\log^p n)$ , as  $n \rightarrow \infty$ . Introduce the following two-dimensional discrete grids with mesh size  $q_n$ :

$$\mathcal{J}_n(B) = \{(x, y) \in Bq_n\mathbb{Z}^2 : x \in [-\varepsilon_n, w_n + \varepsilon_n], y - x \in [l_n^- - \delta_n, l_n^+ + \delta_n]\}, \tag{54}$$

$$\mathcal{J}'_n(B) = \{(x, y) \in Bq_n\mathbb{Z}^2 : x \in [\varepsilon_n, w_n - \varepsilon_n], y - x \in [l_n^- + \delta_n, l_n^+ - \delta_n]\}. \tag{55}$$

Note that  $\mathcal{J}'_n(B) \subset \mathcal{J}_n(B)$ . The discrete squares  $\{\mathbb{T}_{Bq_n}(x, y)\}_{(x,y) \in \mathcal{J}_n(B)}$  (which were defined in Section 4.2) are disjoint and cover the set  $\mathbb{J}_n(0)$ . Similarly, the discrete squares  $\{\mathbb{T}_{Bq_n}(x, y)\}_{(x,y) \in \mathcal{J}'_n(B)}$  are disjoint and contained in  $\mathbb{J}_n(0)$ . By the Bonferroni inequality, we have, for every  $B \in \mathbb{N}$ ,

$$S'_n(B) - S''_n(B) \leq \mathbb{P} \left[ \max_{(i,j) \in \mathbb{J}_n(0)} \mathcal{Z}_{i,j} > u_n \right] \leq S_n(B), \tag{56}$$

where

$$S_n(B) = \sum_{(x,y) \in \mathcal{J}_n(B)} \mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_{Bq_n}(x,y)} \mathcal{Z}_{i,j} > u_n \right], \tag{57}$$

$$S'_n(B) = \sum_{(x,y) \in \mathcal{J}'_n(B)} \mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_{Bq_n}(x,y)} \mathcal{Z}_{i,j} > u_n \right], \tag{58}$$

$$S''_n(B) = \sum \mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_{Bq_n}(x_1,y_1)} \mathcal{Z}_{i,j} > u_n, \max_{(i,j) \in \mathbb{T}_{Bq_n}(x_2,y_2)} \mathcal{Z}_{i,j} > u_n \right] \tag{59}$$

and in (59) the sum is taken over all pairs  $(x_1, y_1) \in \mathcal{J}'_n(B)$  and  $(x_2, y_2) \in \mathcal{J}'_n(B)$  such that  $(x_1, y_1) \neq (x_2, y_2)$ . The statement of Lemma 4.8 follows by letting  $n \rightarrow \infty$  and then  $B \rightarrow \infty$  in (56) and applying Lemmas 4.9, 4.10 and 4.16 which we will prove below.  $\square$

**Lemma 4.9.** *Let  $S_n(B)$  be defined as in (57). We have*

$$\lim_{B \rightarrow \infty} \limsup_{n \rightarrow \infty} n w_n^{-1} S_n(B) \leq e^{-\tau} \int_{A_1}^{A_2} \Lambda(a) da. \tag{60}$$

**Proof.** Since the probability on the right-hand side of (57) depends only on  $l := y - x$  by translation invariance, we have

$$S_n(B) \leq \frac{w_n + o(w_n)}{Bq_n} \sum_{l \in \mathcal{L}_n(B)} \mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_{Bq_n}(0,l)} \mathcal{Z}_{i,j} > u_n \right],$$

where  $\mathcal{L}_n(B) = Bq_n \mathbb{Z} \cap [l_n^- - \delta_n, l_n^+ + \delta_n]$ . The idea is now to apply to each probability Lemma 4.3 and replace Riemann sums by integrals. Introduce the function

$$\lambda_{n,B}(a) = n \log^{2-p} n \cdot \mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_{Bq_n}(0, l_{n,B}(a))} \mathcal{Z}_{i,j} > u_n \right], \quad a > 0,$$

where  $l_{n,B}(a) = \max\{l \in Bq_n \mathbb{Z} : l \leq a \log^p n\}$ . The function  $\lambda_{n,B}(a)$  is locally constant and its constancy intervals have length  $Bq_n / \log^p n \sim B / \log n$ . It follows that

$$S_n(B) \leq \frac{w_n + o(w_n)}{B^2 n} \int_{A_1 - \frac{2\delta_n}{\log^p n}}^{A_2 + \frac{2\delta_n}{\log^p n}} \lambda_{n,B}(a) da. \tag{61}$$

For every fixed  $a > 0$ , the sequence  $l_n = l_{n,B}(a)$  satisfies the assumption of Lemma 4.3. Hence, by Lemmas 4.2 and 4.3, we have the pointwise convergence

$$\lim_{n \rightarrow \infty} \lambda_{n,B}(a) = e^{-\tau} \Lambda_B(a), \quad \Lambda_B(a) = \frac{1}{2\sqrt{\pi}} \cdot e^{-\kappa 2^{\frac{q}{2}} a^{-\frac{q-2}{2}}} \left( 1 + H^2 \left( \frac{B}{a} \right) \right). \tag{62}$$

Also, by Lemma 4.4,  $\lambda_{n,B}(a)$  is bounded by a constant not depending on  $a, n$ , as long as  $a$  stays bounded away from 0 and  $\infty$ . Applying the dominated convergence theorem to (61) we obtain

$$\limsup_{n \rightarrow \infty} n w_n^{-1} S_n(B) \leq e^{-\tau} \int_{A_1}^{A_2} B^{-2} \Lambda_B(a) da. \tag{63}$$

Now we let  $B \rightarrow \infty$ . Recall that  $H$  is a function defined by (41). It is known that  $\lim_{B \rightarrow \infty} H(B)/B = 1$ ; see [34, p. 72] or [28, p. 232]. It follows from (62) that uniformly in  $a \in [A_1, A_2]$ ,

we have

$$\lim_{B \rightarrow \infty} B^{-2} \Lambda_B(a) = \frac{1}{2\sqrt{\pi}a^2} e^{-\kappa 2^{\frac{q}{2}} a^{-\frac{q-2}{2}}} = \Lambda(a).$$

To complete the proof let  $B \rightarrow \infty$  in (63).  $\square$

**Lemma 4.10.** *Let  $S'_n(B)$  be defined as in (58). We have*

$$\lim_{B \rightarrow \infty} \liminf_{n \rightarrow \infty} n w_n^{-1} S'_n(B) \geq e^{-\tau} \int_{A_1}^{A_2} \Lambda(a) da. \tag{64}$$

**Proof.** Analogous to the proof of Lemma 4.9.  $\square$

**Remark 4.11.** It follows from (60) and (64) that in both equations we can replace inequality by equality.

The next lemma is needed to estimate the “double sum”  $S''_n(B)$ . It states that the exceedance events over different intervals become asymptotically independent with exponential decorrelation speed as the symmetric difference of the intervals gets larger. Consider two intervals  $K_1 = (i_1, j_1) \in \mathbb{I}$  and  $K_2 = (i_2, j_2) \in \mathbb{I}$  with lengths  $k_1 := j_1 - i_1$  and  $k_2 := j_2 - i_2$  satisfying  $k_1, k_2 \in [l_n^-, l_n^+]$  and such that  $i_1, i_2, j_1, j_2 \in q_n \mathbb{Z}$ . Let  $k \in \mathbb{N}_0$  be the length of the intersection  $K := K_1 \cap K_2$ . (More precisely,  $K$  is the intersection of the sets  $\{i_1 + 1, \dots, j_1\}$  and  $\{i_2 + 1, \dots, j_2\}$ ). Assume without restriction of generality that  $k_1 \leq k_2$  and write  $\Delta = \Delta(K_1, K_2) = k_2 - k$ . In some sense,  $\Delta$  measures the distance between the intervals  $K_1$  and  $K_2$ .

**Lemma 4.12.** *Given an interval  $K_0 = (i_0, j_0) \in \mathbb{I}$  define a random event*

$$F_n(i_0, j_0) = \left\{ \max_{(i,j) \in \mathbb{T}_{q_n}(i_0, j_0)} \mathcal{Z}_{i,j} > u_n \right\}.$$

There exist constants  $C_1, C_2 > 0$  (depending on  $A_1, A_2$  but not depending on  $K_1, K_2, n$ ) such that for every  $K_1, K_2$  as above,

$$P_n(K_1, K_2) := \mathbb{P}[F_n(K_1) \cap F_n(K_2)] \leq C_1 u_n^{-1} e^{-u_n^2/2} e^{-C_2 \Delta(K_1, K_2)/q_n}.$$

**Remark 4.13.** Note that  $\mathbb{P}[F_n(K_\varepsilon)] \sim \text{const} \cdot u_n^{-1} e^{-u_n^2/2}$ ,  $\varepsilon = 1, 2$ , by Lemma 4.3 and (39). The factor  $e^{-C_2 \Delta(K_1, K_2)/q_n}$  provides an estimate for the dependence between the events  $F_n(K_1)$  and  $F_n(K_2)$ .

**Proof of Lemma 4.12.** Given a finite set  $I \subset \mathbb{Z}$  let  $S_I = \sum_{m \in I} X_m$ . Any interval  $(i_0, j_0) \in \mathbb{I}$  will be identified with the finite set  $\{i_0 + 1, \dots, j_0\}$ . In particular, we need the random variables  $S_{K_1} = S_{j_1} - S_{i_1}$  and  $S_{K_2} = S_{j_2} - S_{i_2}$ . Introduce the random variables

$$\begin{aligned} D_1^- &= \max_{0 \leq k < q_n} (S_{i_1} - S_{i_1-k}), & D_1^+ &= \max_{0 \leq k < q_n} (S_{j_1+k} - S_{j_1}), \\ D_2^- &= \max_{0 \leq k < q_n} (S_{i_2} - S_{i_2-k}), & D_2^+ &= \max_{0 \leq k < q_n} (S_{j_2+k} - S_{j_2}). \end{aligned}$$

These random variables are corrections appearing when we extend the base intervals  $K_1$  and  $K_2$  by small intervals of length at most  $q_n$ . With this notation we have an inclusion of events

$$F_n(K_\varepsilon) \subset \{S_{K_\varepsilon} + D_\varepsilon^+ + D_\varepsilon^- > \sqrt{k_\varepsilon} u_n\}, \quad \varepsilon = 1, 2.$$

Denote by  $\bar{K}_1 = (i_1 - q_n, j_1 + q_n)$  the extended version of the interval  $K_1$ . Note that the interval  $\bar{K}_1$  contains all intervals from  $\mathbb{T}_{q_n}(i_1, j_1)$ . Let  $\bar{K} = K_2 \cap \bar{K}_1$  be the extended intersection of  $K_1$  and  $K_2$ , and denote its length by  $\bar{k} = |\bar{K}| < k + 2q_n$ . Let  $\bar{\Delta} = k_2 - \bar{k}$  be the length of  $K_2 \setminus \bar{K} = K_2 \setminus \bar{K}_1$ . Fix  $\varepsilon > 0$ . Introduce the following random events

$$E_1 = \left\{ S_{\bar{K}} + D_2^- + D_2^+ > \sqrt{\bar{k}}u_n + 2\sqrt{2q_n} + \varepsilon \frac{\bar{\Delta}\sqrt{k_2 + 2q_n}}{q_n u_n} \right\},$$

$$E_2 = \left\{ S_{K_1} + D_1^- + D_1^+ > \sqrt{k_1}u_n \right\},$$

$$E_3 = \left\{ S_{K_2} - S_{\bar{K}} > (\sqrt{k_2} - \sqrt{\bar{k}})u_n - 2\sqrt{2q_n} - \varepsilon \frac{\bar{\Delta}\sqrt{k_2 + 2q_n}}{q_n u_n} \right\}.$$

Note that  $F_n(K_1) \subset E_2$ . We have  $F_n(K_1) \cap F_2(K_2) \subset E_1 \cup (E_2 \cap E_3)$ . By construction, the events  $E_2$  and  $E_3$  are independent. We will estimate the probabilities of  $E_1, E_2, E_3$ . Bringing these estimates together will complete the proof of the lemma. First we estimate the probability of  $E_2$ . By Lemma 4.4,

$$\mathbb{P}[E_2] = \mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_{q_n}(i_1, j_1)} \frac{S_j - S_i}{\sqrt{k_1}} > u_n \right] \leq C u_n^{-1} e^{-u_n^2/2}. \tag{65}$$

We now estimate  $\mathbb{P}[E_3]$ . Consider the case  $\bar{\Delta} \geq 10\sqrt{A}q_n$ . Since  $\bar{k} \leq k_2 \leq A \log^p n, u_n \sim \sqrt{2 \log n}, q_n \sim \log^{p-1} n$ , we can choose  $\varepsilon > 0$  so small that the following inequality is valid:

$$(\sqrt{k_2} - \sqrt{\bar{k}})u_n - \varepsilon \frac{\bar{\Delta}\sqrt{k_2 + 2q_n}}{q_n u_n} = \frac{u_n \bar{\Delta}}{\sqrt{k_2} + \sqrt{\bar{k}}} - 2\sqrt{2q_n} - \varepsilon \frac{\bar{\Delta}\sqrt{k_2 + 2q_n}}{q_n u_n} \geq \frac{\varepsilon \bar{\Delta}}{\sqrt{q_n}}.$$

Note that  $\Delta \leq 3\bar{\Delta}$  since  $\Delta \leq \bar{\Delta} + 2q_n$  and  $\bar{\Delta} \geq q_n$ . With Lemma 3.4 and (35) it follows that

$$\mathbb{P}[E_3] \leq \mathbb{P} \left[ S_{\bar{\Delta}} > \frac{\varepsilon \bar{\Delta}}{\sqrt{q_n}} \right] \leq \exp \left\{ -\bar{\Delta} I \left( \frac{\varepsilon}{\sqrt{q_n}} \right) \right\} \leq c_1 \exp \left\{ -\frac{c_2 \Delta}{q_n} \right\}. \tag{66}$$

In the case  $\bar{\Delta} < 10\sqrt{A}q_n$  we can use the trivial estimate  $\mathbb{P}[E_3] \leq 1$  and (66) remains valid provided that  $c_1$  is sufficiently large.

We estimate  $\mathbb{P}[E_1]$ . First we have to get rid of the correction terms  $D_2^-$  and  $D_2^+$ . By the inequality stated in Theorem 2.4 on p. 52 in [33], for every  $t \in \mathbb{R}$  we have

$$\mathbb{P}[D_2^\pm \geq t] \leq 2\mathbb{P}[S_{q_n} \geq t - \sqrt{2q_n}].$$

If  $X$  is any random variable which is independent of  $D_2^+$  and has distribution function  $F$ , then

$$\begin{aligned} \mathbb{P}[X + D_2^+ \geq t] &= \int_{\mathbb{R}} \mathbb{P}[D_2^+ \geq t - s] dF(s) \\ &\leq 2 \int_{\mathbb{R}} \mathbb{P}[S_{q_n} \geq t - s - \sqrt{2q_n}] dF(s) \\ &= 2\mathbb{P}[X + S_{q_n} \geq t - \sqrt{2q_n}]. \end{aligned}$$

Applying this trick twice to  $D_2^+$  and  $D_2^-$  we obtain

$$\mathbb{P}[E_1] \leq 4\mathbb{P} \left[ S_{\bar{k}+2q_n} \geq \sqrt{\bar{k}}u_n + \varepsilon \frac{\bar{\Delta}\sqrt{k_2 + 2q_n}}{q_n u_n} \right] \leq 4\mathbb{P} \left[ \frac{S_{\bar{k}+2q_n}}{\sqrt{\bar{k} + 2q_n}} \geq u_n + \varepsilon \frac{\bar{\Delta}}{q_n u_n} \right],$$

where in the second inequality we used the relations  $q_n \sim \log^{p-1} n, u_n \sim \sqrt{2 \log n}, q_n \leq \bar{k} \leq k_2$ . (The case  $\bar{k} = 0$  can be excluded since Lemma 4.16 holds trivially in this case due to the independence of  $F_n(K_1)$  and  $F_n(K_2)$ ). Applying to the right-hand side Theorem 3.3 and then (35), we obtain

$$\mathbb{P}[E_1] \leq C u_n^{-1} \exp \left\{ -\frac{1}{2} \left( u_n + \varepsilon \frac{\bar{\Delta}}{q_n u_n} \right)^2 \right\} \leq C_1 u_n^{-1} e^{-u_n^2/2} e^{-C_2 \Delta/q_n}. \tag{67}$$

The proof of the lemma is completed by recalling that  $F_n(K_1) \cap F_n(K_2) \subset E_1 \cup (E_2 \cap E_3)$ , where  $E_2$  and  $E_3$  are independent, and applying (65), (66), (67).  $\square$

Given  $B \in \mathbb{N}$  consider a discrete  $d$ -dimensional cube  $Q_B = \{0, \dots, B - 1\}^d$ . We can decompose the lattice  $\mathbb{Z}^d$  into disjoint discrete cubes of the form  $Bu + Q_B, u \in \mathbb{Z}^d$ . Given  $v_1, v_2 \in \mathbb{Z}^d$  we write  $v_1 \sim v_2$  if  $v_1$  and  $v_2$  are in the same cube in this decomposition, that is if there is  $u \in \mathbb{Z}^d$  such that  $v_1, v_2 \in Bu + Q_B$ . Even though this is not explicit in our notation, the relation  $\sim$  depends on  $B$ . Given a set  $V \subset \mathbb{Z}^d$  let  $\partial_B V$  be the  $B$ -boundary of  $V$  defined as the union of discrete cubes  $Bu + Q_B, u \in \mathbb{Z}^d$ , which have a non-empty intersection with both  $V$  and  $\mathbb{Z}^d \setminus V$ . Let  $\|\cdot\|$  be the sup-norm on  $\mathbb{R}^d$  and  $|V|$  the cardinality of a finite set  $V$ .

**Lemma 4.14.** *Let  $\mathbb{V}_n \subset \mathbb{Z}^d$  be a sequence of finite sets such that for every  $B \in \mathbb{N}, |\partial_B \mathbb{V}_n| = o(|\mathbb{V}_n|)$ , as  $n \rightarrow \infty$ . Then, for every  $\varepsilon > 0$ ,*

$$\lim_{B \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{|\mathbb{V}_n|} \sum_{\substack{v_1, v_2 \in \mathbb{V}_n \\ v_1 \sim v_2}} e^{-\varepsilon \|v_1 - v_2\|} = 0.$$

**Proof.** Take any sequence  $k_B \in \mathbb{N}$  such that  $k_B \rightarrow \infty$  but  $k_B = o(B)$  as  $B \rightarrow \infty$ . We have a disjoint decomposition of the cube  $Q_B$  into the “kernel”  $Q'_B$  and the “shell”  $Q''_B$  defined by

$$Q'_B = \{k_B, \dots, B - k_B\}^d, \quad Q''_B = Q_B \setminus Q'_B.$$

We have also a disjoint decomposition  $\mathbb{Z}^d = Z'_B \cup Z''_B$ , where  $Z'_B = \cup_{u \in \mathbb{Z}^d} (Bu + Q'_B)$  and  $Z''_B = \mathbb{Z}^d \setminus Z'_B$ .

Consider any  $v_1 \in Z'_B$ . Then, for every  $v_2 \in \mathbb{Z}^d$  such that  $v_1 \sim v_2$  we have  $\|v_2 - v_1\| > k(B)$ . Hence,

$$\sum_{v_2 \in \mathbb{V}_n: v_1 \sim v_2} e^{-\varepsilon \|v_1 - v_2\|} \leq \sum_{v_2 \in \mathbb{Z}^d: \|v_2\| > k_B} e^{-\varepsilon \|v_2\|} =: C_B. \tag{68}$$

Consider any  $v_1 \in Z''_B$ . There is a constant  $C < \infty$  (not depending on  $v_1$ ) such that

$$\sum_{v_2 \in \mathbb{V}_n: v_1 \sim v_2} e^{-\varepsilon \|v_1 - v_2\|} \leq \sum_{v_2 \in \mathbb{Z}^d} e^{-\varepsilon \|v_1 - v_2\|} = \sum_{v_2 \in \mathbb{Z}^d} e^{-\varepsilon \|v_2\|} = C. \tag{69}$$

It follows from  $|\partial_B \mathbb{V}_n| = o(|\mathbb{V}_n|)$  that  $|Z'_B \cap \mathbb{V}_n|/|\mathbb{V}_n| \rightarrow 1, |Z''_B \cap \mathbb{V}_n|/|\mathbb{V}_n| \rightarrow 0$ , and  $|\mathbb{V}_n| \rightarrow \infty$ , as  $n \rightarrow \infty$ . Using (68) and (69) we obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{|\mathbb{V}_n|} \sum_{\substack{v_1, v_2 \in \mathbb{V}_n \\ v_1 \sim v_2}} e^{-\varepsilon \|v_1 - v_2\|} \leq \limsup_{n \rightarrow \infty} \left( C_B \frac{|Z'_B \cap \mathbb{V}_n|}{|\mathbb{V}_n|} + C \frac{|Z''_B \cap \mathbb{V}_n|}{|\mathbb{V}_n|} \right) = C_B.$$

To complete the proof note that  $\lim_{B \rightarrow \infty} C_B = 0$  by (68).  $\square$

Introduce the finite set

$$\mathbb{V}_n = \mathbb{Z}^2 \cap \left\{ (x, y) \in \mathbb{R}^2 : x \in \left[ 0, \frac{w_n}{q_n} \right], y - x \in \left[ \frac{l_n^-}{q_n}, \frac{l_n^+}{q_n} \right] \right\}. \tag{70}$$

**Lemma 4.15.** *There is  $c > 0$  such that for all  $v_1, v_2 \in \mathbb{V}_n$  we have  $\Delta := \Delta(v_1, v_2) \geq c \|v_1 - v_2\|$ , where  $\Delta(v_1, v_2)$  was introduced before Lemma 4.12.*

**Proof.** Let  $v = (x_1, y_1), v_2 = (x_2, y_2)$ , where  $x_1 < y_1$  and  $x_2 < y_2$ . Then,  $\|v_1 - v_2\| = \max(|x_1 - x_2|, |y_1 - y_2|)$ . Without restriction of generality, let  $x_1 \leq x_2$ . If  $x_2 \leq y_1 \leq y_2$ , then  $\Delta = \|v_1 - v_2\|$  by definition. If  $x_2 < y_2 \leq y_1$ , then  $\Delta = |x_2 - x_1| + |y_2 - y_1| \geq \|v_1 - v_2\|$ . Finally, if  $y_1 \leq x_2$ , then  $\Delta = \max(y_1 - x_1, y_2 - x_2) \geq l_n^-/q_n \geq c_1 \log n$  by definition of  $\mathbb{V}_n$  and  $\|v_1 - v_2\| \leq w_n/q_n + l_n^+/q_n \leq c_2 \log n$ .  $\square$

**Lemma 4.16.** *Let  $S_n''(B)$  be defined as in (59). We have*

$$\lim_{B \rightarrow \infty} \limsup_{n \rightarrow \infty} n w_n^{-1} S_n''(B) = 0.$$

**Proof.** Take first some fixed  $B \in \mathbb{N}$ . Given a vector  $v = (x, y) \in \mathbb{R}^2$  we write  $v' = (x + 1, y)$ . Let  $(x, y) \in Bq_n\mathbb{Z}^2, x < y$ , be some interval. Then, we can represent the discrete square  $\mathbb{T}_{Bq_n}(x, y)$  as a disjoint union of  $B^2$  squares of the form  $\mathbb{T}_{q_n}(q_n v')$ , where

$$v \in v_0 + \{0, \dots, B - 1\}^2, \quad v_0 = \left( \frac{x - Bq_n}{q_n}, \frac{y}{q_n} \right) \in B\mathbb{Z}^2.$$

We can estimate the exceedance probability over  $\mathbb{T}_{Bq_n}(x, y)$  by the sum of the exceedance probabilities over the  $\mathbb{T}_{q_n}$ 's. It follows from (59) that

$$S_n''(B) \leq \sum_{v_1, v_2} \mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_{q_n}(q_n v_1')} Z_{i,j} > u_n, \max_{(i,j) \in \mathbb{T}_{q_n}(q_n v_2')} Z_{i,j} > u_n \right],$$

where the sum is taken over all  $v_1 \in \mathbb{V}_n$  and  $v_2 \in \mathbb{V}_n$  such that  $v_1 \approx v_2$ . Applying Lemma 4.12 and noting that  $\Delta(q_n v_1', q_n v_2') = q_n \Delta(v_1, v_2)$  we get

$$S_n''(B) \leq c_1 u_n^{-1} e^{-u_n^2/2} \sum_{v_1, v_2} e^{-c_2 \Delta(v_1, v_2)} \leq \frac{c_3}{n \log^{2-p} n} \sum_{v_1, v_2} e^{-c_4 \|v_1 - v_2\|}.$$

Here,  $\|\cdot\|$  denotes the sup-norm and the second inequality follows from (37) and Lemma 4.15. Applying to the right-hand side Lemma 4.14 and noting that  $|\mathbb{V}_n| \leq C w_n \log^{2-p} n$  we arrive at the required statement.  $\square$

#### 4.4. Global probability

Recall that for  $0 < A_1 < A_2$  we define  $l_n^- = A_1 \log^p n$  and  $l_n^+ = A_2 \log^p n$ . Denote by  $\mathbb{I}_n(A_1, A_2)$  the set of all intervals  $(i, j) \in \mathbb{I}_n$  with length  $l := j - i \in [l_n^-, l_n^+]$ . Our aim is to prove Theorem 1.2 which states that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \max_{(i,j) \in \mathbb{I}_n(A_1, A_2)} Z_{i,j} \leq u_n \right] = \exp \left\{ -e^{-\tau} \int_{A_1}^{A_2} \Lambda(a) da \right\}. \tag{71}$$

We will decompose the set  $\mathbb{I}_n(A_1, A_2)$  into sets of the form  $\mathbb{J}_n(z)$ ; see Lemma 4.8. Let  $w_n = \lceil 3A_2 \log^p n \rceil$ . To get rid of the boundary effects choose a sequence  $\eta_n > 0$  such that  $\eta_n = o(n)$  but  $\eta_n/w_n \rightarrow \infty$ . Consider the one-dimensional grids

$$\mathcal{R}'_n = [-\eta_n, n + \eta_n] \cap w_n\mathbb{Z}, \quad \mathcal{R}''_n = [\eta_n, n - \eta_n] \cap w_n\mathbb{Z}.$$

Note that the sets  $\mathbb{J}_n(z)$ , where  $z \in \mathcal{R}'_n$ , are disjoint and cover  $\mathbb{I}_n(A_1, A_2)$ . Similarly, the sets  $\mathbb{J}_n(z)$ , where  $z \in \mathcal{R}''_n$ , are disjoint and contained in  $\mathbb{I}_n(A_1, A_2)$ . The exceedance probability over each  $\mathbb{J}_n(z)$  satisfies, by Lemma 4.8,

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{J}_n(z)} \mathcal{Z}_{i,j} > u_n \right] \sim e^{-\tau} \frac{w_n}{n} \int_{A_1}^{A_2} \Lambda(a) da, \quad n \rightarrow \infty. \tag{72}$$

Also,  $|\mathcal{R}'_n| \sim |\mathcal{R}''_n| \sim n/w_n$  as  $n \rightarrow \infty$ . If the exceedance events over  $\mathbb{J}_n(z)$  were independent, the Poisson limit theorem would immediately yield (71). However, the events are dependent. In fact, the dependence is quite weak: the exceedance event over  $\mathbb{J}_n(z)$  depends only on the two neighboring exceedance events over  $\mathbb{J}_n(z \pm w_n)$ , if  $n$  is large. To justify the use of the Poisson limit theorem for such finite-range dependent events we need to check that (see, e.g., [2, Thm. 1])

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{J}_n(0)} \mathcal{Z}_{i,j} > u_n, \max_{(i,j) \in \mathbb{J}_n(w_n)} \mathcal{Z}_{i,j} > u_n \right] = o\left(\frac{w_n}{n}\right), \quad n \rightarrow \infty. \tag{73}$$

Since we can apply Lemma 4.8 to the set  $\mathbb{J}_n(0) \cup \mathbb{J}_n(w_n)$  (replacing  $w_n$  by  $2w_n$ ), we have

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{J}_n(0) \cup \mathbb{J}_n(w_n)} \mathcal{Z}_{i,j} > u_n \right] \sim e^{-\tau} \frac{2w_n}{n} \int_{A_1}^{A_2} \Lambda(a) da, \quad n \rightarrow \infty. \tag{74}$$

Combining (72) and (74) we obtain the required relation (73). Thus, the use of the Poisson limit theorem is justified. The proof of (71) is complete.

#### 4.5. Non-optimal lengths

We will now estimate the exceedance probability over all intervals  $(i, j) \in \mathbb{I}_n$  which are non-optimal in the sense that their length  $l := j - i$  is not between  $l_n^- = A^{-1} \log^p n$  and  $l_n^+ = A \log^p n$ , where  $A > 0$  is large. Denoting by  $\mathbb{B}_n(A)$  the set of all such intervals, we will show that

$$\lim_{A \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left[ \max_{(i,j) \in \mathbb{B}_n(A)} \mathcal{Z}_{i,j} > u_n \right] = 0.$$

This means that the contribution of  $\mathbb{B}_n(A)$  to  $\mathbf{M}_n$  becomes negligible as  $A \rightarrow \infty$ . Combining this with Theorem 1.2 proved above, we obtain Theorem 1.1. We will decompose the set  $\mathbb{B}_n(A)$  into three subsets  $\mathbb{B}'_n, \mathbb{B}''_n, \mathbb{B}'''_n$  and estimate the exceedance probabilities over these sets in the next three lemmas. We start by considering very small intervals.

**Lemma 4.17.** Fix an arbitrary  $a > 0$ . Let  $\mathbb{B}'_n$  be the set of all intervals  $(i, j) \in \mathbb{I}_n$  whose length  $l = j - i$  satisfies  $l \leq a \log n$ . Then,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \max_{(i,j) \in \mathbb{B}'_n} \mathcal{Z}_{i,j} > u_n \right] = 0.$$

**Proof.** Let  $l \in \mathbb{N}$  be such that  $l \leq a \log n$ . Recall from (37) that  $u_n \sim \sqrt{2 \log n}$ , as  $n \rightarrow \infty$ . Then,  $u_n/\sqrt{l} > 1/\sqrt{a}$ , for all large  $n$ . Consequently, by (35), there is  $\delta > 0$  such that for all large  $n$  and all  $l \leq a \log n$ ,

$$I\left(\frac{u_n}{\sqrt{l}}\right) > (1 + 2\delta)\frac{u_n^2}{2l} > (1 + \delta)\frac{\log n}{l}.$$

By Lemma 3.4 the exceedance probability for every individual interval from  $\mathbb{B}'_n$  satisfies

$$\mathbb{P}\left[\frac{S_l}{\sqrt{l}} > u_n\right] \leq \exp\left\{-lI\left(\frac{u_n}{\sqrt{l}}\right)\right\} \leq n^{-(1+\delta)}.$$

Since the number of intervals in  $\mathbb{B}'_n$  is at most  $an \log n$ , we obtain the statement of the lemma.  $\square$

**Lemma 4.18.** Fix any  $a > 1$ . Let  $\mathbb{B}''_n = \mathbb{B}''_n(A)$  be the set of all intervals  $(i, j) \in \mathbb{I}_n$  whose length  $l = j - i$  satisfies  $a \log n \leq l \leq A^{-1} \log^p n$ . Then,

$$\lim_{A \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}\left[\max_{(i,j) \in \mathbb{B}''_n(A)} \mathcal{Z}_{i,j} > u_n\right] = 0.$$

**Proof.** For  $k \in \mathbb{N}_0$  let  $\mathbb{B}''_{n,k}$  be the set of all intervals  $(i, j) \in \mathbb{I}_n$  whose length  $l$  satisfies  $2^{-(k+1)} \log^p n \leq l \leq 2^{-k} \log^p n$ . We can cover the set  $\mathbb{B}''_{n,k} \subset \mathbb{Z}^2$  by disjoint discrete squares  $\mathbb{T}_{r_{n,k}}(x, x + l)$ , see Section 4.2, with side length  $r_{n,k} := \lfloor 2^{-k} \log^{p-1} n \rfloor$ , at least as long as  $2^k < \log^{p-1} n$ . The number of squares we need is at most  $c_1 2^k (\log^{2-p} n)n$ . The exceedance probability over any such square  $\mathbb{T}_{r_{n,k}}(x, x + l)$  can be estimated by Lemma 4.4 with  $r = r_{n,k}$ ,  $u = u_n$  and is at most

$$c_2 u_n^{-1} \exp\left\{-\frac{u_n^2}{2} - c_3 u_n^q l^{-\frac{q-2}{2}}\right\} \leq c_2 u_n^{-1} \exp\left\{-\frac{u_n^2}{2} - c_4 2^{k\frac{q-2}{2}}\right\}.$$

Here,  $c_2, \dots, c_4$  do not depend on  $n, k$ . We can cover the set  $\mathbb{B}''_n(A)$  by the sets  $\mathbb{B}''_{n,k}$ , where  $k$  is such that  $A \leq 2^k \leq a^{-1} \log^{p-1} n$ . For the exceedance probability over the set  $\mathbb{B}''_n(A)$  we obtain the estimate

$$\begin{aligned} \mathbb{P}\left[\max_{(i,j) \in \mathbb{B}''_n(A)} \mathcal{Z}_{i,j} > u_n\right] &\leq c_5 (\log^{2-p} n)n \cdot u_n^{-1} e^{-u_n^2/2} \sum_{k=\lceil \log_2 A \rceil}^{\infty} 2^k e^{-c_4 2^{k\frac{q-2}{2}}} \\ &\leq c_6 \sum_{k=\lceil \log_2 A \rceil}^{\infty} 2^k e^{-c_4 2^{k\frac{q-2}{2}}}. \end{aligned}$$

In the last inequality we have used that  $u_n e^{u_n^2/2} \geq c_7 (\log^{2-p} n)n$  by (37). To complete the proof note that the right-hand side tends to 0 as  $A \rightarrow \infty$ .  $\square$

**Lemma 4.19.** Let  $\mathbb{B}'''_n = \mathbb{B}'''_n(A)$  be the set of all intervals  $(i, j) \in \mathbb{I}_n$  whose length  $l = j - i$  satisfies  $l \geq A \log^p n$ . Then,

$$\lim_{A \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}\left[\max_{(i,j) \in \mathbb{B}'''_n(A)} \mathcal{Z}_{i,j} > u_n\right] = 0.$$

**Proof.** For  $k \in \mathbb{N}_0$  consider the set  $\mathbb{B}_{n,k}'''$  of all intervals  $(i, j) \in \mathbb{I}_n$  with length  $l$  satisfying  $2^k \log^p n \leq l \leq 2^{k+1} \log^p n$ . We can cover the set  $\mathbb{B}_{n,k}''' \in \mathbb{Z}^2$  by disjoint discrete squares  $\mathbb{T}_{r_{n,k}}(x, x + l)$ , see Section 4.2, with side length  $r_{n,k} := \lfloor 2^k \log^{p-1} n \rfloor$ . We need at most  $c_1 2^{-k} (\log^{2-p} n)n$  squares. Exceedance probability over any single square can be estimated by Lemma 4.4 by  $c_2 u_n^{-1} e^{-u_n^2/2}$ . For the exceedance probability over the set  $\mathbb{B}_{n,k}'''(A) \subset \cup_{k=\lfloor \log_2 A \rfloor}^{\infty} \mathbb{B}_{n,k}'''$  we obtain the estimate

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{B}_{n,k}'''(A)} \mathcal{Z}_{i,j} > u_n \right] \leq c_3 u_n^{-1} e^{-u_n^2/2} \cdot n \log^{2-p} n \sum_{k=\lfloor \log_2 A \rfloor}^{\infty} 2^{-k} \leq c_4 \sum_{k=\lfloor \log_2 A \rfloor}^{\infty} \frac{1}{2^k}.$$

The right-hand side goes to 0 as  $A \rightarrow \infty$ . The proof is complete.  $\square$

#### 4.6. Proof of Proposition 1.8

We assume that we are in the setting of Section 1.5.2. Let  $\gamma = \sqrt{(1-p)/p}$ . Note that the  $X_k$ 's take values  $\gamma$  and  $-\gamma^{-1}$ . First we will show that

$$I(s) = \sum_{k=1}^{\infty} \frac{\sigma s^{2k}}{4k(2k-1)} (\gamma^{2k-1} + \gamma^{1-2k}) + \sum_{k=1}^{\infty} \frac{\sigma s^{2k+1}}{4k(2k+1)} (\gamma^{-2k} - \gamma^{2k}). \tag{75}$$

Taking into account (21) and solving  $\varphi'(\tilde{t}) = s$ , we have

$$\tilde{t} = \frac{\sigma}{2} \left[ \log(1 + \gamma s) - \log\left(1 - \frac{s}{\gamma}\right) \right],$$

for  $s \in (-\gamma^{-1}, \gamma)$ . For  $s$  in this range,

$$I(s) = s\tilde{t} - \varphi(\tilde{t}) = p(1 + \gamma s) \log(1 + \gamma s) + (1-p) \left(1 - \frac{s}{\gamma}\right) \log\left(1 - \frac{s}{\gamma}\right). \tag{76}$$

For  $s$  outside the interval  $(-\gamma^{-1}, \gamma)$ , we have  $I(s) = +\infty$ . By Taylor's expansion,

$$\log(1 + \gamma s) = \sum_{k=1}^{\infty} (-1)^{k+1} \gamma^k \frac{s^k}{k}, \quad \log\left(1 - \frac{s}{\gamma}\right) = - \sum_{k=1}^{\infty} \gamma^{-k} \frac{s^k}{k}.$$

Inserting this into (76), we obtain (75). If  $p \in (1/2, 1)$ , then  $0 < \gamma < 1$  and hence, it follows from (75) that all coefficients in the Taylor expansion of  $I(s) - (s^2/2)$  are non-negative (and in fact, the coefficient of  $s^3$  is strictly positive). It follows that  $I(s) > s^2/2$  for all  $s > 0$ . This implies that (35) holds. By Proposition 4.1, this implies (7). Together with the Taylor expansion in (21), this shows that we are in the superlogarithmic case with  $q = 3$ .

### 5. Proof in the logarithmic case

Our aim in this section is to prove Theorems 1.3–1.5. Assume that conditions (4), (13), (14) hold. Fix  $\tau \in \mathbb{R}$  and define the normalizing sequence  $u_n = u_n(\tau) > 0$  by

$$u_n^2 = 2m_*(\log n + \tau). \tag{77}$$

Our aim is to compute the limit of  $\mathbb{P}[\mathbf{M}_n \leq u_n]$ , as  $n \rightarrow \infty$ .

5.1. Dual conditions

First of all, we need to replace conditions (13) and (14) by their Legendre–Fenchel conjugates. We will assume that there is  $s_* > 0$  such that

$$\frac{I(s_*)}{s_*^2/2} = \frac{1}{m_*} < 1 \tag{78}$$

and, additionally, for every  $\varepsilon > 0$ ,

$$\inf_{0 < s < s_* - \varepsilon} \frac{I(s)}{s^2/2} > \frac{1}{m_*} \quad \text{and} \quad \inf_{s > s_* + \varepsilon} \frac{I(s)}{s^2/2} > \frac{1}{m_*}. \tag{79}$$

**Proposition 5.1.** *Suppose that (4) holds. Then, conditions (13) and (14) imply conditions (78) and (79) and vice versa. Furthermore, if these conditions hold, then we have  $\varphi(t_*) = I(s_*) = s_* t_*/2$  and*

$$s_* = \varphi'(t_*) = t_* m_*, \quad t_* = I'(s_*) = \frac{s_*}{m_*}, \quad I''(s_*)\varphi''(t_*) = 1. \tag{80}$$

**Proof.** Assume that (13) and (14) hold. Define  $s_* = \varphi'(t_*)$ . We will show that (78) and (79) hold. By Legendre–Fenchel duality,  $I'$  is the inverse function of  $\varphi'$  and vice versa. Hence,  $t_* = I'(s_*)$ . The point  $t_*$  is the unique maximum of the function  $\frac{\varphi(t)}{t^2/2}$  by (13) and (14). The derivative of this function vanishes at  $t_*$  and hence,  $\varphi'(t_*)t_* = 2\varphi(t_*)$ . In view of (13) this implies that  $s_* = t_* m_*$ . Since the maximum of  $s_* t - \varphi(t)$  is attained at  $t = t_*$ , we have, see (30),

$$I(s_*) = \varphi'(t_*)t_* - \varphi(t_*) = \varphi(t_*) = \frac{m_* t_*^2}{2} = \frac{s_* t_*}{2} = \frac{s_*^2}{2m_*}.$$

The inverse function of  $\varphi'$  is  $I'$ . Taking the derivative we obtain  $I''(s_*)\varphi''(t_*) = 1$ . This proves (78) and (80).

We will now show that condition (79) is fulfilled. Fix  $\varepsilon > 0$ . Denote by  $S_\varepsilon(u)$  the set  $\{s > 0 : |s - u| > \varepsilon\}$ . By (14), for every  $\delta > 0$  there exists  $c = c(\delta) < 1$  such that  $\varphi(t) \leq cm_* \frac{t^2}{2}$  for all  $t \in S_\delta(t_*)$ . Then, for every  $s \in S_\varepsilon(s_*)$  and every  $\delta > 0$ ,

$$I(s) = \sup_{t \geq 0} (st - \varphi(t)) \geq \sup_{t \in S_\delta(t_*)} (st - \varphi(t)) \geq \sup_{t \in S_\delta(t_*)} \left( st - c(\delta)m_* \frac{t^2}{2} \right). \tag{81}$$

The supremum of  $st - c(\delta)m_* \frac{t^2}{2}$  is attained at  $t = \frac{s}{c(\delta)m_*}$ . However, we have to check that  $t \in S_\delta(t_*)$ . Recall that  $|s - s_*| > \varepsilon$ . It follows that

$$|t - t_*| = \left| \frac{s}{c(\delta)m_*} - t_* \right| = \left| \frac{s - c(\delta)s_*}{c(\delta)m_*} \right| > \delta,$$

where the last inequality holds if  $\delta = \delta(\varepsilon) > 0$  is sufficiently small. (Note that  $\lim_{\delta \downarrow 0} c(\delta) = 1$ ). In this case,  $t \in S_\delta(t_*)$ . It follows from (81) that  $I(s) \geq \frac{s^2}{2c(\delta)m_*}$  for all  $s \in S_\varepsilon(s_*)$ . This proves (79). The proof that (78) and (79) imply (13) and (14) is analogous, by the Legendre–Fenchel duality.  $\square$

5.2. Individual probability

In the sequel, we assume that conditions (4), (78), (79) hold. In this section we compute asymptotically the exceedance probability for the value attained by the random field  $\mathcal{Z}_{i,j}$  at some

individual point  $(i, j) \in \mathbb{I}$ . We focus here on intervals whose length is close to the optimal length  $d_* \log n$ , where

$$d_* = \frac{1}{\varphi(t_*)} = \frac{1}{I(s_*)} = \frac{2m_*}{s_*^2}. \tag{82}$$

It turns out that the exceedance probability remains the same, up to a constant factor, if we allow fluctuations of the interval length of order  $O(\sqrt{\log n})$  and fluctuations of the threshold of order  $O(u_n^{-1})$ .

**Lemma 5.2.** *Assume that  $X_1$  is non-lattice. Let  $l_n \in \mathbb{N}$  be any sequence such that  $l_n = d_* \log n + a\sqrt{\log n} + o(\sqrt{\log n})$ , for some  $a \in \mathbb{R}$ , as  $n \rightarrow \infty$ . Fix  $s \in \mathbb{R}$ . Then, as  $n \rightarrow \infty$ ,*

$$P_n(s) := \mathbb{P} \left[ \frac{S_{l_n}}{\sqrt{l_n}} > u_n - \frac{s}{u_n} \right] \sim \frac{\sqrt{m_*}}{2\sqrt{\pi}\sigma_*} \cdot e^{\frac{s}{m_*} - \frac{\beta_*^2 a^2}{2}} \cdot \frac{e^{-\tau}}{n\sqrt{\log n}}.$$

Here,  $\sigma_* = \sqrt{\varphi''(t_*)}$  and  $\beta_*^2 = \frac{s_*^4}{8m_*} \left( \frac{1}{\sigma_*^2} - \frac{1}{m_*} \right) > 0$ .

**Proof.** We are going to apply [Theorem 3.2](#) with  $\alpha := \lim_{n \rightarrow \infty} \frac{u_n}{\sqrt{l_n}} = s_*$ . Note that  $I'(\alpha) = t_* = s_*/m_*$  by [Proposition 5.1](#), and  $\sigma^2(\alpha) = \varphi''(t_*) = \sigma_*^2$ . We obtain, by [Theorem 3.2](#),

$$P_n(s) \sim \frac{\sqrt{m_*}}{2\sigma_*\sqrt{\pi \log n}} \exp \left\{ -l_n I \left( \frac{u_n - \frac{s}{u_n}}{\sqrt{l_n}} \right) \right\}. \tag{83}$$

Next we develop the term under the sign of exponential in [\(83\)](#) into a Taylor series. Consider the function  $J(v) = vI(1/\sqrt{v})$ ,  $v > 0$ . By assumptions [\(78\)](#), [\(79\)](#) it has a unique minimum at  $v_* := s_*^{-2}$ . The first two derivatives of  $J$  are given by

$$J'(v) = I \left( \frac{1}{\sqrt{v}} \right) - \frac{1}{2\sqrt{v}} I' \left( \frac{1}{\sqrt{v}} \right), \quad J''(v) = \frac{1}{4v^2} I'' \left( \frac{1}{\sqrt{v}} \right) - \frac{1}{4v^{3/2}} I' \left( \frac{1}{\sqrt{v}} \right). \tag{84}$$

For the values of  $J$  and its derivatives at  $v = v_*$  we obtain

$$J(v_*) = \frac{1}{2m_*}, \quad J'(v_*) = 0, \quad J''(v_*) = \frac{s_*^4}{4} \left( \frac{1}{\sigma_*^2} - \frac{1}{m_*} \right) > 0. \tag{85}$$

Note in passing that since  $J$  attains a minimum at  $v_*$ , we have  $J''(v_*) > 0$ . This proves that  $\beta_*^2$  is indeed positive. Now consider

$$v_n := \frac{l_n}{(u_n - \frac{s}{u_n})^2} = \frac{d_* \log n + a\sqrt{\log n} + o(\sqrt{\log n})}{2m_* \log n + o(\sqrt{\log n})} = v_* + \frac{a + o(1)}{2m_*\sqrt{\log n}}.$$

Expanding  $J$  into a Taylor series at  $v = v_*$ , we obtain

$$\begin{aligned} l_n I \left( \frac{u_n - \frac{s}{u_n}}{\sqrt{l_n}} \right) &= J(v_n) \left( u_n - \frac{s}{u_n} \right)^2 \\ &= \left( \frac{1}{2m_*} + \frac{1}{2} \cdot \frac{a^2}{4m_*^2 \log n} \cdot J''(v_*) \right) (2m_*(\log n + \tau) - 2s) + o(1) \\ &= \log n + \tau + \frac{\beta_*^2 a^2}{2} - \frac{s}{m_*} + o(1). \end{aligned}$$

To complete the proof insert this into [\(83\)](#).  $\square$

### 5.3. Local probability

Next we compute the exceedance probability over a discrete square in the space of intervals. Recall from Section 4.2 that for an interval  $(x, y) \in \mathbb{I}$  of length  $l := y - x$  and  $B \in \mathbb{N}$  we define  $\mathbb{T}_B(x, y)$  to be the set of all intervals  $(i, j) \in \mathbb{I}$  such that

$$x - B < i \leq x \quad \text{and} \quad y \leq j < y + B.$$

The set  $\mathbb{T}_B(x, y)$  is a discrete square with side length  $B$  in  $\mathbb{I} \subset \mathbb{Z}^2$ . Its right bottom point is the “base interval”  $(x, y)$  which is contained in all other intervals belonging to  $\mathbb{T}_B(x, y)$ .

**Lemma 5.3.** *Assume that  $X_1$  is non-lattice. Fix  $B \in \mathbb{N}$  and let  $l_n \in \mathbb{N}$  be a sequence such that  $l_n = d_* \log n + a\sqrt{\log n} + o(\sqrt{\log n})$ , for some  $a \in \mathbb{R}$ . Write  $\mathbb{T}_n = \mathbb{T}_B(x, x + l_n)$ . Then, as  $n \rightarrow \infty$ ,*

$$Q_n := \mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_n} Z_{i,j} > u_n \right] \sim P_n(0) \cdot \left\{ 1 + H_*^2(B) \right\}, \tag{86}$$

where  $P_n(0)$  is as in Lemma 5.2 and the function  $H_* : \mathbb{N} \rightarrow (0, \infty)$  is defined by

$$H_*(B) = \mathbb{E} \left[ \max_{k=0, \dots, B-1} e^{t_* S_k - k\varphi(t_*)} \right]. \tag{87}$$

**Proof.** The proof follows the same idea as the proof of Lemma 4.3, but the incremental process will be approximated by a discrete-time random walk rather than by a Brownian motion. Let  $S_k^{(1)}$  and  $S_k^{(2)}$  be independent random walks defined as in the proof of Lemma 4.3. Define a random variable  $V_n$  by  $Z_{x, x+l_n} = u_n - u_n^{-1} V_n$ . Every interval from  $\mathbb{T}_n$  has the form  $(x - k_1, x + l_n + k_2)$  for some integers  $0 \leq k_1, k_2 < B$ , hence

$$Q_n = \mathbb{P} \left[ \max_{0 \leq k_1, k_2 < B} \frac{(u_n - \frac{V_n}{u_n})\sqrt{l_n} + S_{k_1}^{(1)} + S_{k_2}^{(2)}}{\sqrt{l_n + k_1 + k_2}} > u_n \right].$$

Conditioning on  $V_n = s$  and integrating over  $s$ , we obtain

$$Q_n = P_n(0) + \int_0^\infty G_n(s) d\mu_n(s), \tag{88}$$

where  $\mu_n$  is the probability distribution of  $V_n$  and  $G_n$  is a non-increasing function defined by

$$G_n(s) = \mathbb{P} \left[ \max_{0 \leq k_1, k_2 < B} \left( S_{k_1}^{(1)} + S_{k_2}^{(2)} - u_n(\sqrt{l_n + k_1 + k_2} - \sqrt{l_n}) - \frac{s\sqrt{l_n}}{u_n} \right) > 0 \right]. \tag{89}$$

By Lemma 5.2, for every  $s \geq 0$ ,

$$\lim_{n \rightarrow \infty} \frac{\mu_n(\mathbb{I}(0, s))}{P_n(0)} = \lim_{n \rightarrow \infty} \frac{P_n(s)}{P_n(0)} = e^{\frac{s}{m_*}}. \tag{90}$$

Let  $s_n$  be any sequence converging to  $s \geq 0$ . We compute  $\lim_{n \rightarrow \infty} G_n(s_n)$ . Let  $f_n(k_1, k_2; s_n)$  be a function given by

$$f_n(k_1, k_2; s_n) = u_n(\sqrt{l_n + k_1 + k_2} - \sqrt{l_n}) + \frac{s_n \sqrt{l_n}}{u_n}.$$

Recall that  $u_n \sim \sqrt{2m_* \log n}$  and  $l_n \sim d_* \log n$ , as  $n \rightarrow \infty$ . An elementary calculus shows that

$$\lim_{n \rightarrow \infty} f_n(k_1, k_2; s_n) = \frac{s_*}{2}(k_1 + k_2) + \frac{s}{s_*}.$$

Since the a.s. convergence implies the distributional convergence, we obtain that

$$\max_{0 \leq k_1, k_2 < B} (S_{k_1}^{(1)} + S_{k_2}^{(2)} - f_n(k_1, k_2; s_n)) \xrightarrow[n \rightarrow \infty]{d} \frac{U_1 + U_2}{t_*} - \frac{s}{s_*},$$

where  $U_1, U_2$  are two independent copies of the random variable

$$U := \max_{k=0, \dots, B-1} (t_* S_k - k\varphi(t_*)).$$

It follows that for all but countably many  $s \geq 0$ , and all sequences  $s_n \rightarrow s$ ,

$$\lim_{n \rightarrow \infty} G_n(s_n) = \mathbb{P} \left[ U_1 + U_2 > \frac{s}{m_*} \right]. \tag{91}$$

Assuming for a moment that interchanging the limit and the integral is justified, we obtain from (90) and (91) that

$$\lim_{n \rightarrow \infty} \int_0^\infty G_n(s) \frac{d\mu_n(s)}{P_n(0)} = \int_0^\infty \mathbb{P} \left[ U_1 + U_2 > \frac{s}{m_*} \right] e^{\frac{s}{m_*}} \frac{ds}{m_*} = (\mathbb{E}e^U)^2. \tag{92}$$

Inserting this into (88) completes the proof of Lemma 5.3.

The first equality in (92) will be justified using Lemma 4.5. To verify its last condition we have to obtain uniform estimates on  $G_n$  and  $\mu_n$ ; see Remark 4.6. Continuing (89) and recalling that  $\sqrt{l_n} \sim u_n/s_*$ , we obtain that for all large  $n$ , and all  $s \geq 0$ ,

$$G_n(s) \leq \mathbb{P} \left[ \max_{0 \leq k_1, k_2 < B} (S_{k_1}^{(1)} + S_{k_2}^{(2)}) > \frac{s\sqrt{l_n}}{u_n} \right] \leq 2\mathbb{P} \left[ \max_{0 \leq k < B} S_k^{(1)} > \frac{s}{3s_*} \right].$$

By Lemma 3.4 and (78) we obtain that for some constants  $c_1, c_2 > 0$ , all large  $n$  and all  $s \geq 0$ ,

$$G_n(s) \leq 2 \sum_{k=0}^{B-1} \exp \left\{ -kI \left( \frac{s}{3s_*k} \right) \right\} \leq c_1 e^{-c_2 s^2}. \tag{93}$$

Now we bound  $\mu_n([0, s])$ . Let first  $s \in [u_n^2/2, u_n^2]$ . Using Lemma 3.4 and (78) we obtain that

$$\mathbb{P} \left[ \frac{S_{l_n}}{\sqrt{l_n}} > u_n - \frac{s}{u_n} \right] \leq \exp \left\{ -l_n I \left( \frac{u_n - \frac{s}{u_n}}{\sqrt{l_n}} \right) \right\} \leq \exp \left\{ -\frac{1}{2m_*} \left( u_n - \frac{s}{u_n} \right)^2 \right\}.$$

Let now  $s \in [0, u_n^2/2]$ . Using Theorem 3.3 and (78) we obtain

$$\mathbb{P} \left[ \frac{S_{l_n}}{\sqrt{l_n}} > u_n - \frac{s}{u_n} \right] \leq C u_n^{-1} \exp \left\{ -\frac{1}{2m_*} \left( u_n - \frac{s}{u_n} \right)^2 \right\}.$$

For  $s > u_n^2$  we can estimate the probability by 1. Combining all cases we obtain

$$\mu_n([0, s]) = \mathbb{P} \left[ \frac{S_{l_n}}{\sqrt{l_n}} > u_n - \frac{s}{u_n} \right] \leq C u_n^{-1} \exp \left\{ -\frac{u_n^2}{2m_*} + \frac{s}{m_*} \right\}.$$

Together with Lemma 5.2 this implies that  $\mu_n([0, s]) \leq C e^{s/m_*} P_n(0)$  for all large  $n \in \mathbb{N}$  and  $s \geq 0$ . Conditions of Remark 4.6 are thus verified.  $\square$

5.4. Estimating the double sum

Given real numbers  $A_1 < A_2$  define  $l_n^- = d_* \log n + A_1 \sqrt{\log n}$  and  $l_n^+ = d_* \log n + A_2 \sqrt{\log n}$ . The aim of this section is to prove the following result.

**Lemma 5.4.** *Assume that  $X_1$  is non-lattice. Let  $w_n \rightarrow \infty$  be any integer sequence such that  $w_n = O(\log n)$ . For  $z \in \mathbb{Z}$  let  $\mathbb{J}_n(z)$  be the set of all intervals  $(i, j) \in \mathbb{I}$  such that  $z \leq i < z + w_n$  and  $j - i \in [l_n^-, l_n^+]$ . Then, as  $n \rightarrow \infty$ ,*

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{J}_n(z)} \mathcal{Z}_{i,j} > u_n \right] \sim e^{-\tau} \frac{w_n}{n} \int_{A_1}^{A_2} \Theta(a) da, \tag{94}$$

where  $\Theta(a) = \frac{\sqrt{m_*} H_*^2}{2\sqrt{\pi} \sigma_*} e^{-\frac{\beta_*^2 a^2}{2}}$ ,  $a \in \mathbb{R}$ , and the constant  $H_*$  is given by

$$H_* = \lim_{B \rightarrow \infty} \frac{H_*(B)}{B} = \lim_{B \rightarrow \infty} \frac{1}{B} \mathbb{E} \left[ \max_{k=0, \dots, B-1} e^{t_* S_k - k \varphi(t_*)} \right] \in (0, 1). \tag{95}$$

**Proof.** The existence of the limit in (95) follows by taking  $Y_k = t_* X_k - \varphi(t_*)$  in Lemma 5.16, below. In fact, (95) can be also obtained as a byproduct of the double sum argument presented below. We prove (94). Without restriction of generality, let  $z = 0$ . To get rid of the boundary effects we introduce two sequences  $\varepsilon_n \rightarrow \infty$  and  $\delta_n \rightarrow \infty$  such that  $\varepsilon_n = o(w_n)$  and  $\delta_n = o(\sqrt{\log n})$  as  $n \rightarrow \infty$ . Take some  $B \in \mathbb{N}$  and let  $q_n = 1$ . Then, with the same notation as in (54), (55), (57), (58), (59), we have the Bonferroni inequality

$$S'_n(B) - S''_n(B) \leq \mathbb{P} \left[ \max_{(i,j) \in \mathbb{J}_n(0)} \mathcal{Z}_{i,j} > u_n \right] \leq S_n(B). \tag{96}$$

The statement of Lemma 5.4 follows by letting  $n \rightarrow \infty$  and then  $B \rightarrow \infty$  in (96) and applying Lemmas 5.5, 5.6 and 5.9 which we will prove below.  $\square$

Recall that  $w_n \rightarrow \infty$  is an integer sequence such that  $w_n = O(\log n)$  as  $n \rightarrow \infty$ .

**Lemma 5.5.** *Let  $S_n(B)$  be defined as in (57) with  $q_n = 1$ . We have*

$$\lim_{B \rightarrow \infty} \limsup_{n \rightarrow \infty} n w_n^{-1} S_n(B) \leq e^{-\tau} \int_{A_1}^{A_2} \Theta(a) da. \tag{97}$$

**Proof.** Since the probability on the right-hand side of (57) does not depend on  $x$ , we have

$$S_n(B) = \frac{w_n + o(w_n)}{B} \sum_{l \in \mathcal{L}_n(B)} \mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_B(0,l)} \mathcal{Z}_{i,j} > u_n \right],$$

where  $\mathcal{L}_n(B) = B\mathbb{Z} \cap [l_n^- - \delta_n, l_n^+ + \delta_n]$ . The idea is to apply to each probability Lemma 5.3 and replace Riemann sums by Riemann integrals. Introduce the function

$$\theta_{n,B}(a) = n \sqrt{\log n} \cdot \mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_B(0, l_{n,B}(a))} \mathcal{Z}_{i,j} > u_n \right], \quad a \in \mathbb{R},$$

where  $l_{n,B}(a) = \max\{l \in B\mathbb{Z} : l \leq d_* \log n + a\sqrt{\log n}\}$ . The function  $\theta_{n,B}(a)$  is locally constant and its constancy intervals have length  $B/\sqrt{\log n}$ . It follows that

$$S_n(B) \leq \frac{w_n + o(w_n)}{B^2 n} \int_{A_1 - \frac{2\delta_n}{\sqrt{\log n}}}^{A_2 + \frac{2\delta_n}{\sqrt{\log n}}} \theta_{n,B}(a) da. \tag{98}$$

For every fixed  $a \in \mathbb{R}$ , the sequence  $l_n = l_{n,B}(a)$  satisfies the assumption of Lemma 5.3. By Lemmas 5.2 and 5.3, for every  $a \in \mathbb{R}$ ,

$$\lim_{n \rightarrow \infty} \theta_{n,B}(a) = e^{-\tau} \Theta_B(a), \quad \Theta_B(a) = \frac{\sqrt{m_*}}{2\sqrt{\pi}\sigma_*} e^{-\frac{\beta_*^2 a^2}{2}} (1 + H_*^2(B)).$$

We also need an estimate for  $\theta_{n,B}(a)$  which is uniform in  $a$ . Assume that  $a \in [-c, c]$ , for some  $c > 0$ . For every interval  $(i, j) \in \mathbb{T}_B(0, l_{n,B}(a))$  of length  $l$  we have, by Theorem 3.3 and (78),

$$\mathbb{P}[Z_{i,j} > u_n] \leq C u_n^{-1} \exp\left\{-lI\left(\frac{u_n}{\sqrt{l}}\right)\right\} \leq \frac{C}{\sqrt{\log n}} \exp\left\{-\frac{u_n^2}{2m_*}\right\} \leq \frac{C}{n\sqrt{\log n}}.$$

Since  $\mathbb{T}_B(0, l_{n,B}(a))$  consists of  $B^2$  intervals, we obtain that  $\theta_{n,B}(a) \leq C$  for all  $a \in [-c, c]$ , where  $C$  does not depend on  $n$  and  $a$ . Taking the limit as  $n \rightarrow \infty$  in (98) and applying the dominated convergence theorem, we obtain

$$\limsup_{n \rightarrow \infty} n w_n^{-1} S_n(B) \leq e^{-\tau} \int_{A_1}^{A_2} B^{-2} \Theta_B(a) da. \tag{99}$$

This holds for every  $B \in \mathbb{N}$ . We let  $B \rightarrow \infty$ . The limit  $H_* := \lim_{B \rightarrow \infty} B^{-1} H_*(B) \in (0, \infty)$  exists by (95). Hence,  $\lim_{B \rightarrow \infty} B^{-2} \Theta_B(a) = \Theta(a)$  uniformly in  $a \in \mathbb{R}$ , where  $\Theta(a)$  is defined as in Lemma 5.4. To complete the proof let  $B \rightarrow \infty$  in (99).  $\square$

**Lemma 5.6.** *Let  $S'_n(B)$  be defined as in (58) with  $q_n = 1$ . We have*

$$\lim_{B \rightarrow \infty} \liminf_{n \rightarrow \infty} n w_n^{-1} S'_n(B) \geq e^{-\tau} \int_{A_1}^{A_2} \Theta(a) da. \tag{100}$$

**Proof.** Analogous to the proof of Lemma 5.5.  $\square$

**Remark 5.7.** It follows from (97) and (100) that in both equations we can replace inequality by equality.

The next lemma is needed to estimate the “double sum”  $S''_n(B)$ . It provides an estimate for the correlation between exceedance events over different intervals. Consider two intervals  $K_1 = (i_1, j_1) \in \mathbb{I}$  and  $K_2 = (i_2, j_2) \in \mathbb{I}$  such that  $k_1 := j_1 - i_1 \in [l_n^-, l_n^+]$  and  $k_2 := j_2 - i_2 \in [l_n^-, l_n^+]$ . Let  $K$  be the intersection of  $K_1$  and  $K_2$ . Denote by  $k \in \mathbb{N}_0$  the cardinality of  $K$ . Assume that  $k_1 \leq k_2$  and let  $\Delta = \Delta(K_1, K_2) = k_2 - k$ .

**Lemma 5.8.** *There exist  $C_1, C_2 > 0$  not depending on  $n, \tau$  such that for all intervals  $K_1$  and  $K_2$  as above and all  $|\tau| < \sqrt{\log n}$ ,*

$$\mathbb{P}[Z_{i_1, j_1} > u_n, Z_{i_2, j_2} > u_n] \leq \frac{C_1 e^{-\tau}}{n\sqrt{\log n}} e^{-C_2 \Delta(K_1, K_2)}. \tag{101}$$

**Proof.** Fix  $\varepsilon > 0$ . The event  $\{Z_{i_1, j_1} > u_n, Z_{i_2, j_2} > u_n\}$  is contained in the event  $E_1 \cup (E_2 \cap E_3)$ , where

$$E_1 = \left\{ S_K > \sqrt{k}u_n + \varepsilon \frac{\sqrt{k}\Delta}{u_n} \right\},$$

$$E_2 = \{S_{K_1} > \sqrt{k_1}u_n\},$$

$$E_3 = \left\{ S_{K_2 \setminus K} > (\sqrt{k_2} - \sqrt{k})u_n - \varepsilon \frac{\sqrt{k}\Delta}{u_n} \right\}.$$

We estimate  $\mathbb{P}[E_1]$ . Let first  $u_n \geq 2s_*k$ . Using Lemma 3.4 and (79) we obtain that there is  $\delta > 0$  such that

$$\mathbb{P}[E_1] \leq \exp \left\{ -kI \left( \frac{u_n + \varepsilon u_n^{-1}\Delta}{\sqrt{k}} \right) \right\} \leq \exp \left\{ -\frac{1 + \delta}{2m_*} (u_n^2 + 2\varepsilon\Delta) \right\}.$$

Now let  $u_n \leq 2s_*k$ . We have, by Theorem 3.3 and (78),

$$\mathbb{P}[E_1] \leq C u_n^{-1} \exp \left\{ -kI \left( \frac{u_n + \varepsilon u_n^{-1}\Delta}{\sqrt{k}} \right) \right\} \leq \frac{C}{\sqrt{\log n}} \exp \left\{ -\frac{(u_n^2 + 2\varepsilon\Delta)}{2m_*} \right\}.$$

Combining both cases we obtain that

$$\mathbb{P}[E_1] \leq \frac{C e^{-\tau}}{n \sqrt{\log n}} e^{-\varepsilon\Delta/m_*}. \tag{102}$$

We estimate  $\mathbb{P}[E_2]$ . By Theorem 3.3 and (78),

$$\mathbb{P}[E_2] \leq C u_n^{-1} \exp \left\{ -kI \left( \frac{u_n}{\sqrt{k}} \right) \right\} \leq \frac{C}{\sqrt{\log n}} \exp \left\{ -\frac{u_n^2}{2m_*} \right\} \leq \frac{C e^{-\tau}}{n \sqrt{\log n}}. \tag{103}$$

We estimate  $\mathbb{P}[E_3]$ . We have, using that  $k_2, k \leq l_n^+$ ,

$$\mathbb{P}[E_3] = \mathbb{P} \left[ \frac{S_\Delta}{\Delta} > \frac{u_n}{\sqrt{k_2} + \sqrt{k}} - \varepsilon \frac{\sqrt{k}}{u_n} \right] \leq \mathbb{P} \left[ \frac{S_\Delta}{\Delta} > \frac{u_n}{2\sqrt{l_n^+}} - \varepsilon \frac{\sqrt{l_n^+}}{u_n} \right].$$

We can choose  $\varepsilon > 0$  so small that  $\frac{u_n}{2\sqrt{l_n^+}} - \varepsilon \frac{\sqrt{l_n^+}}{u_n} > \varepsilon$ . It follows by Lemma 3.4 that

$$\mathbb{P}[E_3] \leq \mathbb{P}[S_\Delta > \varepsilon\Delta] \leq e^{-\Delta I(\varepsilon)}. \tag{104}$$

Here,  $I(\varepsilon) > 0$ . The probability on the left-hand side of (101) is not larger than  $\mathbb{P}[E_1] + \mathbb{P}[E_2]\mathbb{P}[E_3]$  since the events  $E_2$  and  $E_3$  are independent. Combining (102), (103), (104) we obtain the required estimate.  $\square$

**Lemma 5.9.** Let  $S''_n(B)$  be defined as in (59) with  $q_n = 1$ . We have

$$\lim_{B \rightarrow \infty} \limsup_{n \rightarrow \infty} n w_n^{-1} S''_n(B) = 0.$$

**Proof.** Introduce the finite set  $\mathbb{V}_n = \mathbb{Z}^2 \cap ([0, w_n] \times [l_n^-, l_n^+])$ . Take some fixed  $B \in \mathbb{N}$ . Let  $(x, y) \in B\mathbb{Z}^2, x < y$ , be some interval. The discrete square  $\mathbb{T}_B(x, y)$  consists of  $B^2$  intervals.

We estimate the exceedance probability over  $\mathbb{T}_B(x, y)$  by the sum of the exceedance probabilities over these intervals. It follows from (59) that

$$S''_n(B) \leq \sum_{v_1, v_2} \mathbb{P} \left[ \mathcal{Z}_{v_1} > u_n, \mathcal{Z}_{v_2} > u_n \right],$$

where the sum is taken over all  $v_1 \in \mathbb{V}_n$  and  $v_2 \in \mathbb{V}_n$  such that  $v_1 \approx v_2$ . Applying Lemmas 5.8 and 4.15 we get

$$S''_n(B) \leq \frac{C}{n\sqrt{\log n}} \sum_{v_1, v_2} e^{-c_2 \Delta(v_1, v_2)} \leq \frac{C}{n\sqrt{\log n}} \sum_{v_1, v_2} e^{-c_4 \|v_1 - v_2\|}.$$

Here,  $\|\cdot\|$  is the sup-norm. Applying to the right-hand side Lemma 4.14 and noting that  $|\mathbb{V}_n| \leq C_1 w_n \sqrt{\log n}$  we arrive at the required statement.  $\square$

### 5.5. Global probability

Given  $A_1 < A_2$  recall that  $l_n^- = d_* \log n + A_1 \sqrt{\log n}$  and  $l_n^+ = d_* \log n + A_2 \sqrt{\log n}$ . Denote by  $\mathbb{I}_n(A_1, A_2)$  the set of all intervals  $(i, j) \in \mathbb{I}_n$  with length  $l := j - i \in [l_n^-, l_n^+]$ . Theorem 1.5 states that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \max_{(i, j) \in \mathbb{I}_n(A_1, A_2)} \mathcal{Z}_{i, j} \leq u_n \right] = \exp \left\{ -e^{-\tau} \int_{A_1}^{A_2} \Theta(a) da \right\}. \tag{105}$$

The proof of (105) goes as follows. Let  $w_n = [3d_* \log n]$ . We decompose the set  $\mathbb{I}_n(A_1, A_2)$  into  $\sim n/w_n$  sets of the form  $\mathbb{J}_n(z)$ ,  $z \in w_n \mathbb{Z}$ . The exceedance probability over any of these sets is asymptotically equivalent to  $e^{-\tau} (\int_{A_1}^{A_2} \Theta(a) da) w_n/n$  by Lemma 5.4. Also, the exceedance event over  $\mathbb{J}_n(z)$  is independent of all other exceedance events except for  $\mathbb{J}_n(z \pm w_n)$ . Justifying the use of the Poisson limit theorem we obtain (105). The proof, up to trivial changes, is the same as in Section 4.4.

### 5.6. Non-optimal lengths

In this section we complete the proof of Theorem 1.3. For  $A > 0$  write  $l_n^- = d_* \log n - A\sqrt{\log n}$  and  $l_n^+ = d_* \log n + A\sqrt{\log n}$ . Denote by  $\mathbb{B}_n(A) = \mathbb{I}_n \setminus \mathbb{I}_n(-A, A)$  the set of all intervals  $(i, j) \in \mathbb{I}_n$  whose length  $l := j - i$  satisfies  $l \notin [l_n^-, l_n^+]$ . The aim of this section is to show that the contribution of these non-optimal lengths to  $\mathbb{M}_n$  is negligible, if  $A$  is large. More precisely, we will show that

$$\lim_{A \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left[ \max_{(i, j) \in \mathbb{B}_n(A)} \mathcal{Z}_{i, j} > u_n \right] = 0. \tag{106}$$

Combined with Theorem 1.5 proved above this yields Theorem 1.3. We will cover the set  $\mathbb{B}_n(A)$  by three sets  $\mathbb{B}'_n, \mathbb{B}''_n, \mathbb{B}'''_n$  (depending on some further parameters) which will be considered separately in the next three lemmas. In this section we do not need the non-lattice assumption. First we consider intervals which are sufficiently small but not close to the optimal length  $d_* \log n$ .

**Lemma 5.10.** *Let  $0 < \varepsilon < d_*$  be fixed. For  $\delta > 0$  denote by  $\mathbb{B}'_n = \mathbb{B}'_n(\varepsilon, \delta)$  the set of all intervals  $(i, j) \in \mathbb{I}_n$  whose length  $l = j - i$  satisfies  $l \leq n^\delta$  and  $|l - d_* \log n| > \varepsilon \log n$ . Then, we can choose  $\delta > 0$  so small that*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \max_{(i, j) \in \mathbb{B}'_n} \mathcal{Z}_{i, j} > u_n \right] = 0.$$

**Proof.** Let  $l \in \mathbb{N}$  be such that  $l \leq n^\delta$  and  $|l - d_* \log n| > \varepsilon \log n$ . Then, we can find an  $\varepsilon_1 > 0$  (depending on  $\varepsilon$ , but not on  $\delta, l, n$ ) such that  $|\frac{u_n}{\sqrt{l}} - s_*| > \varepsilon_1$ . By (79) there are  $\delta_1, \delta_2 > 0$  such that for all large  $n$  and all  $l$  as above,

$$II \left( \frac{u_n}{\sqrt{l}} \right) = \frac{I(u_n/\sqrt{l})}{(u_n/\sqrt{l})^2/2} \cdot \frac{u_n^2}{2} > \left( \frac{1}{m_*} + \delta_1 \right) \cdot \frac{u_n^2}{2} > (1 + \delta_2) \log n.$$

Using Lemma 3.4 we obtain that for all  $(i, j) \in \mathbb{B}'_n$ , the individual exceedance probability can be estimated as follows:

$$\mathbb{P}[Z_{i,j} > u_n] \leq e^{-II(u_n/\sqrt{l})} < n^{-(1+\delta_2)}.$$

Choose any  $\delta \in (0, \delta_2)$ . Since the number of intervals in  $\mathbb{B}'_n$  is at most  $n^{1+\delta}$ , the statement of the lemma follows.  $\square$

Next we consider intervals whose length is close to being optimal.

**Lemma 5.11.** For  $\varepsilon > 0$  and  $A > 0$  let  $\mathbb{B}''_n = \mathbb{B}''_n(A)$  be the set of all intervals  $(i, j) \in \mathbb{I}_n$  whose length  $l = j - i$  satisfies  $(d_* - \varepsilon) \log n \leq l \leq l_n^-$  or  $l_n^+ \leq l \leq (d_* + \varepsilon) \log n$ . Then, we can choose  $\varepsilon > 0$  so small that

$$\lim_{A \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left[ \max_{(i,j) \in \mathbb{B}''_n(A)} Z_{i,j} > u_n \right] = 0.$$

**Proof.** It follows from (85) that we can choose  $\varepsilon, \delta > 0$  so small that for all  $v \in (v_* - \varepsilon, v_* + \varepsilon)$ ,

$$J(v) > \frac{1}{2m_*} + \delta(v - v_*)^2. \tag{107}$$

Here,  $v_* = s_*^{-2}$ . For  $k \in \mathbb{N}_0$  let  $\mathbb{B}''_{n,k}$  be the set of all intervals  $(i, j) \in \mathbb{I}_n$  with length  $l$  satisfying

$$d_* \log n + 2^k \sqrt{\log n} \leq l \leq d_* \log n + 2^{k+1} \sqrt{\log n}.$$

It follows from (107) that for every  $(i, j) \in \mathbb{B}''_{n,k}$ ,

$$II \left( \frac{u_n}{\sqrt{l}} \right) = u_n^2 J \left( \frac{l}{u_n^2} \right) > \frac{u_n^2}{2m_*} + \frac{\delta}{u_n^2} \left( d_* \log n + 2^k \sqrt{\log n} - \frac{u_n^2}{s_*^2} \right)^2 > \log n + c_1 2^{2k}.$$

By Theorem 3.3 we obtain

$$\mathbb{P} \left[ \frac{S_l}{\sqrt{l}} > u_n \right] \leq c_2 u_n^{-1} \exp \left\{ -II \left( \frac{u_n}{\sqrt{l}} \right) \right\} \leq \frac{c_3}{n \sqrt{\log n}} e^{-c_4 2^{2k}}.$$

Let  $\mathbb{B}''_{n,+}(A)$  be the set of all intervals  $(i, j) \in \mathbb{I}_n$  such that  $l_n^+ \leq l \leq (d_* + \varepsilon) \log n$ . The number of intervals in  $\mathbb{B}''_{n,k}$  is at most  $c_5 2^k n \sqrt{\log n}$ . For the exceedance probability over the set  $\mathbb{B}''_{n,+}(A)$  we obtain the estimate

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{B}''_{n,+}(A)} Z_{i,j} > u_n \right] \leq c_6 \sum_{k=\lceil \log_2 A \rceil}^{\infty} 2^k e^{-c_4 2^{2k}}.$$

The right-hand side goes to 0 as  $A \rightarrow \infty$ . Exceedance probability over the set  $\mathbb{B}''_{n,-}(A)$  consisting of all intervals  $(i, j) \in \mathbb{I}_n$  with length satisfying  $(d_* - \varepsilon) \log n \leq l \leq l_n^-$  can be estimated analogously.  $\square$

**Lemma 5.12.** *Let  $\delta > 0$  be arbitrary. Let  $\mathbb{B}'''_n$  be the set of all intervals  $(i, j) \in \mathbb{I}_n$  whose length  $l = j - i$  satisfies  $l \geq \log^{1+\delta} n$ . Then,*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \max_{(i,j) \in \mathbb{B}'''_n} \mathcal{Z}_{i,j} > u_n \right] = 0.$$

**Proof.** For  $k \in \mathbb{N}$  such that  $2^k > \log n$  let  $\mathbb{B}'''_{n,k}$  be the set of all intervals  $(i, j) \in \mathbb{I}_n$  with length  $l$  satisfying  $2^k \leq l \leq 2^{k+1}$ . We can cover this set by discrete squares of the form  $\mathbb{T}_{r_k}(x, x+l)$ , see Section 5.3, where  $r_k = \lfloor 2^k / \log n \rfloor$  and  $x, l \in r_k \mathbb{Z}$ . We need at most  $c2^{-k} n \log^2 n$  such squares. The exceedance probability over each such square can be estimated using the same method as in Lemma 4.4. For every  $\varepsilon > 0$  there is  $\eta > 0$  such that  $I(y) \geq (1 - \varepsilon)y^2/2$  for all  $y \in [0, \eta]$ . Recall that  $l \geq \log^{1+\delta} n$  and hence,  $u_n/\sqrt{l} \leq \eta$  for all sufficiently large  $n$ . By Lemma 3.4, for every  $s \in [0, u_n^2]$ ,

$$F_{l,u_n}(s) := \mathbb{P} \left[ \frac{S_l}{\sqrt{l}} > u_n - \frac{s}{u_n} \right] \leq \exp \left\{ -lI \left( \frac{u_n - \frac{s}{u_n}}{\sqrt{l}} \right) \right\} \leq e^s \exp \left\{ -\frac{(1 - \varepsilon)u_n^2}{2} \right\}.$$

This inequality continues to hold for  $s \geq u_n^2$ , since in this case the right-hand side is greater than 1. Arguing in the same way as in the proof of Lemma 4.4, but replacing (50) by the above inequality, we obtain

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{T}_{r_k}(x,x+l)} \mathcal{Z}_{i,j} > u_n \right] \leq C \exp \left\{ -\frac{(1 - \varepsilon)u_n^2}{2} \right\} \leq Cn^{-(1+\varepsilon)}.$$

For the exceedance probability over the set  $\mathbb{B}'''_{n,k}$  we obtain

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{B}'''_{n,k}} \mathcal{Z}_{i,j} > u_n \right] \leq C2^{-k} n^{-\frac{\varepsilon}{2}}.$$

To complete the proof, take the sum over all  $k \in \mathbb{N}$ . □

### 5.7. Tightness in the lattice case

Now we allow the distribution of  $X_1$  to be lattice and prove Theorem 1.4. The tightness of the sequence  $\mathbf{M}_n^2 - 2m_* \log n$  follows from Lemmas 5.13 and 5.15 below. We use the same notation as in Section 5.6. Namely, for  $A > 0$  we write  $l_n^- = d_* \log n - A\sqrt{\log n}$  and  $l_n^+ = d_* \log n + A\sqrt{\log n}$ . Let  $\mathbb{I}_n(A)$  be the set of all intervals  $(i, j) \in \mathbb{I}_n$  with length  $l = j - i$  satisfying  $l \in [l_n^-, l_n^+]$ . Write  $\mathbb{B}_n(A) = \mathbb{I}_n \setminus \mathbb{I}_n(A)$ . Let  $u_n = u_n(\tau) > 0$  be defined by

$$u_n^2(\tau) = 2m_*(\log n + \tau), \quad \tau \in \mathbb{R}. \tag{108}$$

**Lemma 5.13.** *For every  $\varepsilon > 0$  we can find  $\tau = \tau(\varepsilon)$  such that  $\mathbb{P}[\mathbf{M}_n > u_n(\tau)] < \varepsilon$  for all large  $n$ .*

**Proof.** Fix  $\varepsilon > 0$ . By (106) we can find  $A = A(\varepsilon)$  such that for large  $n$ ,

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{B}_n(A)} \mathcal{Z}_{i,j} > u_n(0) \right] < \frac{\varepsilon}{2}. \tag{109}$$

We estimate the exceedance probability over  $\mathbb{I}_n(A)$ . Let  $\tau > 0$ . For every interval  $(i, j) \in \mathbb{I}_n(A)$  with length  $l \in [l_n^-, l_n^+]$  we have, by Theorem 3.3 and (79),

$$\mathbb{P}[\mathcal{Z}_{i,j} > u_n(\tau)] \leq C u_n^{-1} \exp \left\{ -lI \left( \frac{u_n}{\sqrt{l}} \right) \right\} \leq \frac{C}{\sqrt{\log n}} \exp \left\{ -\frac{u_n^2}{2m_*} \right\} = \frac{C e^{-\tau}}{n\sqrt{\log n}}.$$

Since the number of elements in  $\mathbb{I}_n(A)$  is at most  $4An\sqrt{\log n}$ , we obtain that there is  $C_1$  depending only on  $A$  such that

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{I}_n(A)} \mathcal{Z}_{i,j} > u_n(\tau) \right] \leq C_1 e^{-\tau}, \tag{110}$$

We can choose  $\tau > 0$  so large that  $C_1 e^{-\tau} < \varepsilon/2$ . To complete the proof combine (109) and (110).  $\square$

We now give a lower estimate for the exceedance probability. Let  $w_n = [3d_* \log n]$ . Take  $A = 1$ . Define  $\mathbb{J}_n(z)$  as in Lemma 5.4.

**Lemma 5.14.** *There is a constant  $C$  such that for all  $z \in \mathbb{Z}$ , large  $n \in \mathbb{N}$ , and all  $|\tau| < \sqrt{\log n}$ ,*

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{J}_n(z)} \mathcal{Z}_{i,j} > u_n(\tau) \right] \geq \frac{C e^{-\tau}}{n/w_n}.$$

**Proof.** Without restriction of generality let  $z = 0$ . Take some  $B \in \mathbb{N}$ . Let  $\mathcal{J}_n(B)$  be a two-dimensional discrete grid with mesh size  $B$  defined by

$$\mathcal{J}_n(B) = B\mathbb{Z}^2 \cap \mathbb{J}_n(0) = \{(i, j) \in B\mathbb{Z}^2 : i \in [0, w_n], j - i \in [l_n^-, l_n^+]\}.$$

Then, by the Bonferroni inequality,

$$\mathbb{P} \left[ \max_{(i,j) \in \mathbb{J}_n(0)} \mathcal{Z}_{i,j} > u_n(\tau) \right] \geq S'_n(B) - S''_n(B), \tag{111}$$

where  $S'_n(B)$  and  $S''_n(B)$  are defined by

$$S'_n(B) = \sum_{(i,j) \in \mathcal{J}_n(B)} \mathbb{P}[\mathcal{Z}_{i,j} > u_n(\tau)], \quad S''_n(B) = \sum \mathbb{P}[\mathcal{Z}_{v_1} > u_n(\tau), \mathcal{Z}_{v_2} > u_n(\tau)],$$

and the second sum is taken over all  $v_1 \neq v_2$  with  $v_1 = (i_1, j_1) \in \mathcal{J}_n(B)$  and  $v_2 = (i_2, j_2) \in \mathcal{J}_n(B)$ . We estimate  $S''_n(B)$  first. By Lemmas 5.8 and 4.15,

$$\mathbb{P}[\mathcal{Z}_{v_1} > u_n(\tau), \mathcal{Z}_{v_2} > u_n(\tau)] \leq \frac{C e^{-\tau}}{n\sqrt{\log n}} e^{-c\|v_1 - v_2\|}.$$

It follows that

$$S''_n(B) \leq \frac{C_1 e^{-\tau}}{n\sqrt{\log n}} \sum e^{-c\|v_1 - v_2\|} \leq \frac{C_1 e^{-\tau}}{n\sqrt{\log n}} |\mathcal{J}_n(B)| \sum_{v \in B\mathbb{Z}^2} e^{-c\|v\|}. \tag{112}$$

Now we estimate  $S'_n(B)$ . By Theorem 6 of [32] (which is a converse inequality to Theorem 3.3),

$$\mathbb{P}[\mathcal{Z}_{i,j} > u_n(\tau)] \geq C_2 u_n^{-1} \exp \left\{ -lI \left( \frac{u_n}{\sqrt{l}} \right) \right\} \geq \frac{C_3 e^{-\tau}}{n\sqrt{\log n}}.$$

Hence,

$$S'_n(B) \geq |\mathcal{J}_n(B)| \frac{C_3 e^{-\tau}}{n\sqrt{\log n}}. \tag{113}$$

We can choose  $B$  so large that  $C_1 \sum_{v \in B\mathbb{Z}^2} e^{-c\|v\|} < C_3/2$ . Taking (111), (112), (113) together and noting that  $|\mathcal{J}_n(B)| > cw_n\sqrt{\log n}$  yields the statement of the lemma.  $\square$

**Lemma 5.15.** *For every  $\varepsilon > 0$  we can find  $\tau = \tau(\varepsilon)$  (sufficiently close to  $-\infty$ ) such that  $\mathbb{P}[\mathbf{M}_n < u_n(\tau)] < \varepsilon$  for all large  $n$ .*

**Proof.** Consider the sets  $\mathbb{J}_n(z)$ , where  $z \in 2w_n\mathbb{Z}$ . There are at least  $n/(3w_n)$  such sets contained in  $\mathbb{I}_n$ . The exceedance events over these sets are independent, hence,

$$\mathbb{P}[\mathbf{M}_n < u_n(\tau)] \leq \left( 1 - \mathbb{P} \left[ \max_{(i,j) \in \mathbb{J}_n(0)} \mathcal{Z}_{i,j} > u_n(\tau) \right] \right)^{\frac{n}{3w_n}} \leq \left( 1 - \frac{C e^{-\tau}}{n/w_n} \right)^{\frac{n}{3w_n}}.$$

The right-hand side converges to  $\exp\{-\frac{1}{3}C e^{-\tau}\}$ , as  $n \rightarrow \infty$ . It follows that we can choose  $\tau$  so close to  $-\infty$  that for all large  $n$ , the right-hand side is smaller than  $\varepsilon$ . The proof is complete.  $\square$

### 5.8. Pickands-type constant

In this section we provide two alternative expressions for the Pickands-type constant  $H_*$ ; see (17). Let  $Y_1, Y_2, \dots$  be non-degenerate i.i.d. random variables such that  $\mathbb{E}e^{Y_k} = 1, k \in \mathbb{N}$ . Independently, let also  $Y_{-1}, Y_{-2}, \dots$  be i.i.d. random variables such that

$$\mathbb{P}[Y_{-k} \in dy] = e^y \mathbb{P}[Y_k \in dy], \quad k \in \mathbb{N}. \tag{114}$$

Note that  $\mathbb{E}e^{-Y_{-k}} = 1, k \in \mathbb{N}$ . Define a stochastic process  $\{W_k, k \in \mathbb{Z}\}$  by  $W_0 = 0$  and

$$W_k = Y_1 + \dots + Y_k, \quad W_{-k} = Y_{-1} + \dots + Y_{-k}, \quad k \in \mathbb{N}. \tag{115}$$

**Lemma 5.16.** *Let  $L_N = \max_{k=0, \dots, N} W_k, k \in \mathbb{N}$ . Then,*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}e^{L_N} = \mathbb{P}[\forall k \in \mathbb{N} : W_k < 0, W_{-k} \leq 0]. \tag{116}$$

**Remark 5.17.** By [42] the probability on the right-hand side of (116) is equal to  $R_-R_+$ , where

$$R_+ = \exp \left\{ - \sum_{k=1}^{\infty} \frac{1}{k} \mathbb{P}[W_k > 0] \right\}, \quad R_- = \exp \left\{ - \sum_{k=1}^{\infty} \frac{1}{k} \mathbb{P}[W_{-k} \geq 0] \right\}. \tag{117}$$

**Remark 5.18.** By taking  $Y_k = t_* X_k - \varphi(t_*)$  we obtain alternative expressions for the constant  $H_*$  defined in (17). It follows from (116) that  $0 \leq H_* \leq 1$ . Since  $\mathbb{E}Y_k < 0, k \in \mathbb{Z}$ , we have  $\lim_{k \rightarrow \pm\infty} W_k = -\infty$  a.s. by the law of large numbers. Therefore, we even have strict inequalities  $0 < H_* < 1$ .

**Proof of Lemma 5.16.** Let  $g(w) = \sum_{N=0}^{\infty} w^N \mathbb{E}e^{L_N}$ ,  $|w| < 1$ . Then, by p. 207 of [42],

$$g(w) = \frac{1}{1-w} \exp \left\{ - \sum_{k=1}^{\infty} \frac{w^k}{k} \mathbb{E}[(1 - e^{W_k}) \mathbb{1}_{W_k > 0}] \right\}$$

$$= \frac{1}{(1-w)^2} \exp \left\{ - \sum_{k=1}^{\infty} \frac{w^k}{k} (\mathbb{P}[W_k > 0] + \mathbb{P}[W_{-k} \leq 0]) \right\}.$$

It follows that  $g(w) \sim R_+ R_- / (1-w)^2$ , as  $w \uparrow 1$ . The sequence  $\mathbb{E}e^{L_N}$  is non-decreasing. By the Hardy–Littlewood Tauberian theorem, see Corollary 1.7.3 on p. 40 in [5], it follows that  $\lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}e^{L_N} = R_+ R_-$ . In view of Remark 5.17 this completes the proof.  $\square$

### 6. Proof in the sublogarithmic case

#### 6.1. Proof of Theorem 1.6

Let  $U_n = \max\{X_1, \dots, X_n\}$ . Take any  $\frac{1}{2} < \beta < \min(\frac{1}{\alpha}, 1)$ . By the assumption of the theorem, we have

$$n \mathbb{P}[X_1 > \log^\beta n] > n e^{-\log^{\alpha\beta} n} \rightarrow \infty, \quad n \rightarrow \infty,$$

since  $\alpha\beta < 1$ . It follows that

$$\lim_{n \rightarrow \infty} \mathbb{P}[U_n \leq \log^\beta n] = 0.$$

The proof will be complete after we have shown that

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathbf{M}_n(a \log n, n) \leq \log^\beta n] = 1. \tag{118}$$

Recall the definition of  $I$  in (30). Since  $I$  is a convex function we can find  $s_0 > 0$  such that  $I(s) > 3/a$  for all  $s > s_0$ . For every interval  $(i, j) \in \mathbb{I}_n$  with length  $l = j - i$  such that  $a \log n \leq l \leq s_0^{-2} \log^{2\beta} n$  we have, by Lemma 3.4,

$$\mathbb{P}[\mathcal{Z}_{i,j} > \log^\beta n] \leq \exp \left\{ -l I \left( \frac{\log^\beta n}{\sqrt{l}} \right) \right\} \leq \exp \left\{ -\frac{3l}{a} \right\} \leq \frac{1}{n^3}.$$

Since  $I(s) \sim \frac{s^2}{2}$  as  $s \downarrow 0$ , we can find  $c > 0$  such that  $I(s) > cs^2$  for all  $s \in [0, s_0]$ . It follows that for every interval  $(i, j) \in \mathbb{I}_n$  of length  $l \geq s_0^{-2} \log^{2\beta} n$ ,

$$\mathbb{P}[\mathcal{Z}_{i,j} > \log^\beta n] \leq \exp \left\{ -l I \left( \frac{\log^\beta n}{\sqrt{l}} \right) \right\} \leq \exp\{-c \log^{2\beta} n\} \leq \frac{1}{n^3},$$

where we have used that  $\beta > \frac{1}{2}$ . Since the number of intervals in  $\mathbb{I}_n$  is at most  $n^2$  it follows that (118) holds.

#### 6.2. Proof of Theorem 1.7

Choose  $v \in (0, 1)$  such that  $2^{1-\frac{\alpha}{2}} v^{1+\alpha} > 1$  (recall that  $\alpha < 2$ ). By assumption (20) we have

$$n \mathbb{P} \left[ X_1 > v \left( \frac{\log n}{D} \right)^{1/\alpha} \right] = n \exp\{-(v^\alpha + o(1)) \log n\} \rightarrow \infty, \quad n \rightarrow \infty.$$

It follows that the maximum  $U_n = \max\{X_1, \dots, X_n\}$  satisfies

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ U_n \leq v \left( \frac{\log n}{D} \right)^{1/\alpha} \right] = 0.$$

In view of [Theorem 1.6](#) the proof of [Theorem 1.7](#) will be complete after we have shown that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ M_n(2, \log n) \geq v \left( \frac{\log n}{D} \right)^{1/\alpha} \right] = 0. \quad (119)$$

Assume first that  $\alpha = 1$ . Then, condition [\(20\)](#) implies that  $\varphi(t) = \log \mathbb{E}e^{tX_1}$  is finite for  $t \in [0, D)$ , and equal to  $+\infty$  for  $t > D$ . This implies that

$$I(s) \sim Ds \quad \text{as } s \rightarrow \infty.$$

Consider now the case  $\alpha \in (1, 2)$ . By [Kasahara's theorem](#) [[5](#), p. 253], condition [\(20\)](#) is equivalent to  $\varphi(t) \sim Gt^\beta$  as  $t \rightarrow \infty$  where  $\frac{1}{\alpha} + \frac{1}{\beta} = 1$  and  $G^{1-\alpha} \beta^{-\alpha} (1-\beta) = D$ . For the Legendre–Fenchel conjugate, one obtains [[5](#), p. 48]

$$I(s) \sim Ds^\alpha \quad \text{as } s \rightarrow \infty.$$

Hence, both for  $\alpha = 1$  and for  $\alpha \in (1, 2)$  we have  $I(s) > vDs^\alpha$  for large  $s$ . By [Lemma 3.4](#), for every interval  $(i, j) \in \mathbb{I}_n$  of length  $2 \leq l \leq \log n$ , we have, for large  $n$ ,

$$\mathbb{P} \left[ Z_{i,j} > v \left( \frac{\log n}{D} \right)^{1/\alpha} \right] \leq \exp \left\{ -lI \left( \frac{v}{\sqrt{l}} \left( \frac{\log n}{D} \right)^{1/\alpha} \right) \right\} \leq \exp \left\{ -2^{1-\frac{\alpha}{2}} v^{1+\alpha} \log n \right\}.$$

Recall that  $2^{1-\frac{\alpha}{2}} v^{1+\alpha} > 1$ . Since the number of intervals in  $\mathbb{I}_n$  with length not exceeding  $\log n$  is at most  $n \log n$ , we obtain [\(119\)](#).

## Acknowledgments

Zakhar Kabluchko is grateful to Axel Munk from whom he learned about multiscale scan statistics.

## References

- [1] D. Aldous, Probability approximations via the Poisson clumping heuristic, in: Applied Mathematical Sciences, vol. 77, Springer-Verlag, New York, 1989.
- [2] R. Arratia, L. Goldstein, L. Gordon, Two moments suffice for Poisson approximations: the Chen–Stein method, *Ann. Probab.* 17 (1) (1989) 9–25.
- [3] R.R. Bahadur, R. Ranga Rao, On deviations of the sample mean, *Ann. Math. Statist.* 31 (4) (1960) 1015–1027.
- [4] S.M. Berman, Sojourns and extremes of stochastic processes, in: Wadsworth & Brooks/Cole Statistics/Probability Series, 1992.
- [5] N.H. Bingham, C.M. Goldie, J.L. Teugels, Regular variation, in: Encyclopedia of Mathematics and its Applications, vol. 27, Cambridge University Press, Cambridge, 1987.
- [6] S.A. Book, An extension of the Erdős–Rényi new law of large numbers, *Proc. Amer. Math. Soc.* 48 (1975) 438–446.
- [7] V.V. Buldygin, Y.V. Kozachenko, Metric characterization of random variables and random processes, in: Translations of Mathematical Monographs, vol. 188, American Mathematical Society, Providence, RI, 2000.
- [8] H.P. Chan, Maxima of moving sums in a Poisson random field, *Adv. Appl. Probab.* 41 (3) (2009) 647–663.
- [9] H. Cramer, P. Levy, R. de Mises, Les sommes et les fonctions de variables aléatoires, Hermann, Paris, 1938.
- [10] M. Csörgö, P. Révész, Strong approximations in probability and statistics, in: Probability and Mathematical Statistics, Akadémiai Kiadó, Budapest, 1981.

- [11] S. Csörgő, Erdős–Rényi laws, *Ann. Statist.* 7 (4) (1979) 772–787.
- [12] P. Deheuvels, On the Erdős–Rényi theorem for random fields and sequences and its relationships with the theory of runs and spacings, *Z. Wahrsch. Verw. Gebiete* 70 (1) (1985) 91–115.
- [13] P. Deheuvels, L. Devroye, Limit laws of Erdős–Rényi–Shepp type, *Ann. Probab.* 15 (4) (1987) 1363–1386.
- [14] P. Deheuvels, L. Devroye, J. Lynch, Exact convergence rate in the limit theorems of Erdős–Rényi and Shepp, *Ann. Probab.* 14 (1) (1986) 209–223.
- [15] W. Feller, Generalization of a probability limit theorem of Cramér, *Trans. Amer. Math. Soc.* 54 (1943) 361–372.
- [16] J. Glaz, J. Naus, S. Wallenstein, *Scan Statistics*, Springer, New York, 2001.
- [17] J. Glaz, V. Pozdnyakov, S. Wallenstein, *Scan Statistics. Methods and Applications*, Birkhäuser, Boston, MA, 2009.
- [18] T. Höglund, A unified formulation of the central limit theorem for small and large deviations from the mean, *Z. Wahrsch. Verw. Gebiete* 49 (1) (1979) 105–117.
- [19] I.A. Ibragimov, Y.V. Linnik, *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff Publishing, Groningen, 1971.
- [20] Z. Kabluchko, Extreme-value analysis of self-normalized increments, Ph.D. Thesis, University of Göttingen, 2007. Available at <http://webdoc.sub.gwdg.de/diss/2007/kabluchko/>.
- [21] Z. Kabluchko, Extreme-value analysis of standardized Gaussian increments, 2007. Unpublished. Available at <http://www.arxiv.org/abs/0706.1849>.
- [22] Z. Kabluchko, Extremes of the standardized Gaussian noise, *Stochastic Process. Appl.* 121 (3) (2011) 515–533.
- [23] Z. Kabluchko, A. Munk, Shao’s theorem on the maximum of standardized random walk increments for multidimensional arrays, *ESAIM Probab. Stat.* 13 (2009) 409–416.
- [24] Z. Kabluchko, E. Spodarev, Scan statistics of Lévy noises and marked empirical processes, *Adv. Appl. Probab.* 41 (1) (2009) 013–037.
- [25] O. Kallenberg, Foundations of modern probability, in: *Probability and its Applications*, Springer–Verlag, New York, 1997.
- [26] J. Komlós, G. Tusnády, On sequences of pure heads, *Ann. Probab.* 3 (1975) 608–617.
- [27] H. Lanzinger, U. Stadtmüller, Maxima of increments of partial sums for certain subexponential distributions, *Stochastic Process. Appl.* 86 (2) (2000) 307–322.
- [28] M.R. Leadbetter, G. Lindgren, H. Rootzén, *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, New York, 1983.
- [29] T. Mikosch, M. Moser, The limit distribution of the maximum increment of a random walk with dependent regularly varying jump sizes, *Probab. Theory and Related Fields* (2012) 1–24.
- [30] T. Mikosch, A. Račkauskas, The limit distribution of the maximum increment of a random walk with regularly varying jump size distribution, *Bernoulli* 16 (4) (2010) 1016–1038.
- [31] V.V. Petrov, Generalization of Cramér’s limit theorem, *Uspekhi Mat. Nauk (N.S.)* 9 (4(62)) (1954) 195–202.
- [32] V.V. Petrov, On the probabilities of large deviations for sums of independent random variables, *Teor. Veroyatn. Primen.* 10 (1965) 310–322.
- [33] V.V. Petrov, Limit theorems of probability theory, in: *Sequences of Independent Random Variables*, in: *Oxford Studies in Probability*, vol. 4, Oxford University Press, New York, 1995.
- [34] J. Pickands, Upcrossing probabilities for stationary Gaussian processes, *Trans. Amer. Math. Soc.* 145 (1969) 51–73.
- [35] V.I. Piterbarg, On large jumps of a random walk, *Theory Probab. Appl.* 36 (1) (1991) 50–62.
- [36] V.I. Piterbarg, Asymptotic methods in the theory of Gaussian processes and fields, in: *Translations of Mathematical Monographs*, vol. 148, American Mathematical Society, Providence, RI, 1996.
- [37] V.I. Piterbarg, A.M. Kozlov, On large jumps of a random walk with the Cramér condition, *Theory Probab. Appl.* 47 (4) (2003) 719–729.
- [38] P. Révész, *Random Walk in Random and Nonrandom Environments*, World Scientific, Teaneck, NJ, 1990.
- [39] Q.-M. Shao, On a conjecture of Révész, *Proc. Amer. Math. Soc.* 123 (2) (1995) 575–582.
- [40] D. Siegmund, E.S. Venkatraman, Using the generalized likelihood ratio statistic for sequential detection of a change-point, *Ann. Statist.* 23 (1) (1995) 255–271.
- [41] D. Siegmund, B. Yakir, Tail probabilities for the null distribution of scanning statistics, *Bernoulli* 6 (2) (2000) 191–213.
- [42] F. Spitzer, Principles of random walk, in: *The University Series in Higher Mathematics*, D. Van Nostrand Co., Inc., Princeton, 1964.
- [43] J. Steinbach, On a conjecture of Révész and its analogue for renewal processes, in: B. Szyszkowicz (Ed.), *Asymptotic Methods in Probability and Statistics. A Volume in Honour of Miklós Csörgő. ICAMPS’97*, North-Holland/Elsevier, 1998.