J. Appl. Prob. **55,** 431–449 (2018)
*doi:*10.1017/jpr.2018.28
© *Applied Probability Trust* 2018

EXACT MODERATE AND LARGE DEVIATIONS FOR LINEAR RANDOM FIELDS

HAILIN SANG,* *The University of Mississippi*
YIMIN XIAO,** *Michigan State University*

Abstract

By extending the methods of Peligrad *et al.* (2014), we establish exact moderate and large deviation asymptotics for linear random fields with independent innovations. These results are useful for studying nonparametric regression with random field errors and strong limit theorems.

Keywords: Large deviation; moderate deviation; linear random field; nonparametric regression; Davis–Gut law of the iterated logarithm

2010 Mathematics Subject Classification: Primary 60F10; 60G60; 62E20

1. Introduction

Random fields play a central role in modeling and analyzing spatially correlated data and have a wide range of applications. As a consequence, there has been an increasing interest in their study in the probability and statistics literature.

Consider a linear random field $X = \{X_{j,k},\ (j,k) \in \mathbb{Z}^2\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ by

$$X_{j,k} = \sum_{r,s\in\mathbb{Z}} a_{r,s}\xi_{j-r,k-s}, \tag{1.1}$$

where $\{a_{r,s},\ (r,s) \in \mathbb{Z}^2\}$ is a square-summable sequence of constants, and the innovations $\{\xi_{r,s},\ (r,s) \in \mathbb{Z}^2\}$ and ξ_0 are independent and identically distributed (i.i.d.) random variables with $\mathbb{E}\xi_0 = 0$ and $\mathbb{E}\xi_0^2 = 1$. Under these conditions, $X_{j,k}$ in (1.1) is well defined since the series on the right-hand side of (1.1) converges in the $L^2(\Omega, \mathbb{P})$-sense and almost surely; see Lemma A.1. The existing literature contains extensive studies on limit theorems and estimation problems for linear random fields. For example, Marinucci and Poghosyan (2001) and Paulauskas (2010) studied the asymptotics for linear random fields, including the law of large numbers, central limit theorems, and invariance principles, by applying the Beveridge–Nelson decomposition method. Banys *et al.* (2010) applied ergodic theory to study the strong law of large numbers for linear random fields. Mallik and Woodroofe (2011) also established the central limit theorem for linear random fields, and their method does not rely on the Beveridge–Nelson decomposition. Under various settings, Tran (1990), Hallin *et al.* (2004a), (2004b), El Machkouri (2007), (2014), El Machkouri and Stoica (2010), and Wang and Woodroofe (2014) studied local linear regression, kernel density estimation, and their asymptotics for

Received 3 November 2016; revision received 31 August 2017.
* Postal address: Department of Mathematics, The University of Mississippi, University, MS 38677, USA.
Email address: sang@olemiss.edu
** Postal address: Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, USA.
Email address: xiao@stt.msu.edu

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

linear random fields. Gu and Tran (2009) developed a fixed design regression for negatively associated random fields.

However, few authors have studied moderate and large deviations for linear random fields. Davis and Hsing (1995), Mikosch and Samorodnitsky (2000), and Mikosch and Wintenberger (2013) established large deviation results for certain stationary sequences, including linear processes with short-range dependence. For linear processes which allow long-range dependence, we mention that Djellout and Guillin (2001) proved moderate and large deviation results for linear processes with i.i.d. and bounded innovations; Djellout *et al.* (2006) studied a moderate deviation estimate for the empirical periodogram of a linear process; Wu and Zhao (2008) obtained moderate deviations for stationary causal processes and their main theorem can be applied to functionals of linear processes; and, more recently, Peligrad *et al.* (2014) established exact moderate and large deviation asymptotics for linear processes with independent innovations.

The main purpose of this paper is to extend the method of Peligrad *et al.* (2014) to establish exact moderate and large deviations for linear random fields as in (1.1). Let $\{\Gamma_n\}$ be a sequence of finite subsets of $\mathbb{Z}^2$ and denote the cardinality of Γ_n by $|\Gamma_n|$. To be specific, we can take $\Gamma_n = [-n, n]^2 \cap \mathbb{Z}^2$, or $[1, n]^2 \cap \mathbb{Z}^2$ or more general rectangles. Define $S_n := S_{\Gamma_n} := \sum_{(j,k)\in\Gamma_n} X_{j,k}$. By Lemma A.1, this can be written as

$$S_n = \sum_{r,s\in\mathbb{Z}} b_{n,r,s}\xi_{-r,-s}, \tag{1.2}$$

where $b_{n,r,s} = \sum_{(j,k)\in\Gamma_n} a_{j+r,k+s}$. Let $\sigma_n^2 = \mathbb{E}S_n^2$. In the main results of this paper, Theorems 2.1–2.3, we quantify the roles of the moment and right-tail properties of ξ_0, the magnitude of the coefficients $\{a_{r,s}\}$, as well as the speed of convergence of $x_n \to \infty$, in the moderate and large deviation probabilities for $\mathbb{P}(S_n \geq x_n\sigma_n)$. These results are useful for studying asymptotic properties and statistical inference of linear random fields. As examples, we show that our moderate and large deviation results can be applied to the study of nonparametric regressions and for obtaining the convergence rate in the law of the iterated logarithm for linear random fields.

For simplicity of presentation, we focus on linear random fields indexed by $\mathbb{Z}^2$. The theorems presented in this paper can be easily extended to the linear random field $X_{\boldsymbol{j}} = \sum_{\boldsymbol{r}\in\mathbb{Z}^N} a_{\boldsymbol{r}}\xi_{\boldsymbol{j}-\boldsymbol{r}}$ on $\mathbb{Z}^N$ with $N \geq 3$.

In this paper we will use the following notation. For any constant $p \geq 1$, we define $\|a\|_p := [\sum_{r,s\in\mathbb{Z}} |a_{r,s}|^p]^{1/p}$. Then $\|a\|_2 < \infty$ by the assumption and $\|a\|_p$ may be finite for some values of $p < 2$. Similarly, for a random variable ξ, we use $\|\xi\|_p$ to denote its $L^p(\mathbb{P})$-norm for $p \geq 1$. Let $\Phi(x)$ be the distribution function of the standard normal random variable.

For two sequences $\{a_n\}$ and $\{b_n\}$ of real numbers, $a_n \sim b_n$ means $a_n/b_n \to 1$ as $n \to \infty$; $a_n \propto b_n$ means that $a_n/b_n \to C$ as $n \to \infty$ for some constant $C > 0$; for positive sequences, the notation $a_n \ll b_n$ or $b_n \gg a_n$ replaces the Vinogradov symbol O and they mean that a_n/b_n is bounded; $\lceil x \rceil$ means the smallest integer which is greater than or equal to x.

The rest of this paper has the following structure. In Section 2 we state the main results on moderate and large deviations for S_n in (1.2). In Section 3 we apply the main results to nonparametric regression estimates and prove a Davis–Gut law of iterated logarithm for linear random fields. In Appendix A we provide the existing results which are useful in proving the theorems in Section 2.

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

2. Main results

Although the double sum S_n in (1.2) can be written (in infinitely many ways) as a single weighted sum of infinitely many i.i.d. random variables indexed by nonnegative integers, for example, $S_n = \sum_{i=0,\,|r|=i,\,|s|\le i \text{ or } |s|=i,\,|r|<i}^{\infty} b_{n,r,s}\xi_{-r,-s}$ (where the number of terms for each index i is finite and the summation of the terms with the same i can be taken in any order), the important role of the configuration of Γ_n is usually hidden in such a representation, and a partial order in $\mathbb{Z}^2$, which may not be natural for the problem under investigation, has to be imposed. For these reasons it is difficult to solve the problems for random fields satisfactorily by applying directly the results on a weighted sum of random variables indexed by one variable. Quite often new methods have to be developed. We refer the reader to Klesov (2014, Chapter 1) for further illustrations on connections, as well as significant differences, in limit theorems of random fields and stochastic processes of one variable.

The objective of this section is to study moderate and large deviations for the partial sum S_n in (1.2) by extending the results for linear processes as in Peligrad *et al.* (2014) to linear random fields.

We will need the following notation. Define

$$D_{nt} := \sum_{r,s\in\mathbb{Z}} |b_{n,r,s}|^t, \qquad U_{nt} := (D_{n2})^{-t/2} D_{nt}.$$

Then $\sigma_n^2 := \mathbb{E}(S_n^2) = D_{n2}$. To avoid degeneracy, we assume that tacitly $\sigma_n > 0$ for every n. Let $\rho_n^2 := \max_{r,s\in\mathbb{Z}} b_{n,r,s}^2/\sigma_n^2$. We will assume that $\rho_n^2 \to 0$ which means that the contribution of any single coefficient $b_{n,r,s}$ is negligible compared with σ_n^2. We remark that the magnitudes of $b_{n,r,s}^2$ and D_{nt} depend on the coefficients $\{a_{r,s},\ (r,s) \in \mathbb{Z}^2\}$ and the configuration of Γ_n. An interesting case is when $\{a_{r,s}\}$ is isotropic and $\Gamma_n = [1,n]^2 \cap \mathbb{Z}^2$. See (3.4) and Lemma 3.1 below for details. More generally, the case when $\{a_{r,s}\}$ is anisotropic (i.e. $|a_{r,s}|$ depends r and s at different rates) and Γ_n is a rectangle in $\mathbb{Z}^2$ can also be considered.

Our first theorem is the following moderate deviation result.

Theorem 2.1. *Assume that the random variable ξ_0 satisfies $\|\xi_0\|_p < \infty$ for some $p > 2$ and $\rho_n^2 \to 0$ as $n \to \infty$. Then, for $x_n \ge 0$, $x_n^2 \le 2\ln(U_{np}^{-1})$, the moderate deviation result holds:*

$$\mathbb{P}(S_n \ge x_n\sigma_n) = (1-\Phi(x_n))(1+o(1)) \quad \text{as } n\to\infty, \tag{2.1}$$
$$\mathbb{P}(S_n \le -x_n\sigma_n) = (1-\Phi(x_n))(1+o(1)) \quad \text{as } n\to\infty.$$

Proof. We only need to prove (2.1). The proof is a modification of that of Peligrad *et al.* (2014, Corollary 3(iii), supplementary material). The main idea is to apply Theorem A.2 for triangular arrays. For this purpose, we decompose the partial sum S_n as $S_n = M_n + R_n$, where $M_n = \sum_{|r|\le k_n}\sum_{|s|\le k_n} b_{n,r,s}\xi_{-r,-s}$ for some integer k_n which will be chosen later, and R_n is the remainder

$$R_n = \Bigg(\sum_{|r|>k_n}\sum_{s\in\mathbb{Z}} + \sum_{|r|\le k_n}\sum_{|s|>k_n}\Bigg) b_{n,r,s}\xi_{-r,-s}.$$

By the Cauchy–Schwarz inequality, we have

$$b_{n,r,s}^2 \le |\Gamma_n| \sum_{(j,k)\in\Gamma_n} a_{j+r,k+s}^2$$

and

$$\sum_{r,s\in\mathbb{Z}} b_{n,r,s}^2 \le |\Gamma_n| \sum_{r,s\in\mathbb{Z}}\sum_{(j,k)\in\Gamma_n} a_{j+r,k+s}^2 = |\Gamma_n|^2 \sum_{r,s\in\mathbb{Z}} a_{r,s}^2. \tag{2.2}$$

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

In the above, we applied Fubini's theorem to change the order of summations. Therefore, for every integer $n \geq 1$, we have $\sum_{r,s\in\mathbb{Z}} b_{n,r,s}^2 < \infty$, which yields $\sum_{r,s\in\mathbb{Z}} |b_{n,r,s}|^p < \infty$ since $p > 2$.

By applying Rosenthal's inequality (see De la Peña and Giné (1999)) to R_n, we see that there is a constant C_p such that

$$\mathbb{E}(|R_n|^p) \leq C_p\Bigg[\Bigg(\Bigg(\sum_{|r|>k_n}\sum_{s\in\mathbb{Z}} + \sum_{|r|\leq k_n}\sum_{|s|>k_n}\Bigg) b_{n,r,s}^2\Bigg)^{p/2} + \mathbb{E}(|\xi_0|^p)\Bigg(\sum_{|r|>k_n}\sum_{s\in\mathbb{Z}} + \sum_{|r|\leq k_n}\sum_{|s|>k_n}\Bigg)|b_{n,r,s}|^p\Bigg].$$

Now, for each positive integer n, we select an integer k_n large enough such that

$$\Bigg(\sum_{|r|>k_n}\sum_{s\in\mathbb{Z}} + \sum_{|r|\leq k_n}\sum_{|s|>k_n}\Bigg) b_{n,r,s}^2 \leq \|\xi_0\|_p^2\Bigg(\sum_{r,s\in\mathbb{Z}} |b_{n,r,s}|^p\Bigg)^{2/p}.$$

This is possible due to (2.2) and the fact that $\sum_{r,s\in\mathbb{Z}} a_{r,s}^2 < \infty$.

With the above selection of k_n, we obtain

$$\mathbb{E}(|R_n|^p) \leq 2C_p\mathbb{E}(|\xi_0|^p)\sum_{r,s\in\mathbb{Z}} |b_{n,r,s}|^p. \tag{2.3}$$

Similarly, we can verify that M_n and, thus, S_n also have finite moments of order p.

Since M_n is the sum of $(2k_n+1)^2$ independent random variables, we can view $S_n = M_n + R_n$ as the sum of $(2k_n+1)^2+1$ independent random variables and apply Theorem A.2 to prove (2.1). As in the proof of Peligrad *et al.* (2014, Corollary 2.3(iii), supplementary material), we use (2.3) to derive

$$\begin{aligned} L_{np} &= \frac{1}{\sigma_n^p}\Bigg(\sum_{|r|\leq k_n}\sum_{|s|\leq k_n} \mathbb{E}((b_{n,r,s}\xi_{-r,-s})^p\,\mathbf{1}(b_{n,r,s}\xi_{-r,-s} \geq 0)) + \mathbb{E}(R_n^p\,\mathbf{1}(R_n \geq 0))\Bigg) \\ &\leq (2C_p+1)\mathbb{E}(|\xi_0|^p)U_{np}, \end{aligned}$$

where $\mathbf{1}(A)$ is the indicator function on the event A. Since $\rho_n^2 \to 0$ and $p > 2$ imply that $U_{np} \to 0$ as $n \to \infty$, we have $L_{np} \to 0$. Similarly, for all $x \geq 0$ such that $x^2 \leq 2\ln(U_{np}^{-1})$, we can verify that $\Lambda_n(x^4, x^5, \varepsilon) \to 0$ for any $\varepsilon > 0$, and $x^2 - 2\ln(L_{np}^{-1}) - (p-1)\ln\ln(L_{np}^{-1}) \to -\infty$ as $n \to \infty$. Hence, the conditions of Theorem A.2 are satisfied, and (2.1) follows. □

Remark 2.1. The condition that $\rho_n \to 0$ in Theorem 2.1 and the following Theorems 2.2 and 2.3 can be replaced by suitable conditions on $|\Gamma_n|$ and σ_n^2 that are easier to verify. By Hölder's inequality, we have $\rho_n \leq \|a\|_u|\Gamma_n|^{1/v}/\sigma_n$, where $1 \leq u \leq 2$ and v is the conjugate of u, $1/u + 1/v = 1$. Therefore, we can replace the condition $\rho_n \to 0$ by $\|a\|_u < \infty$ and $|\Gamma_n|^{1/v}/\sigma_n \to 0$. In particular, if $\|a\|_1 < \infty$, which is the short-range dependence case, then $\rho_n \leq \|a\|_1/\sigma_n$. In this case, we can replace the condition $\rho_n \to 0$ by $\sigma_n \to \infty$ (as a consequence, we also have $|\Gamma_n| \to \infty$). See Mallik and Woodroofe (2011) for more information on bounds for ρ_n. If Γ_n is a union of l finitely many discrete rectangles, by Proposition 2 of the same paper, $\rho_n \leq 20(\sqrt{l}\|a\|_2/\sigma_n)^{1/5} + 8\sqrt{l}\|a\|_2/\sigma_n$. Therefore, in this case, we can replace the condition $\rho_n \to 0$ by $\|a\|_2 < \infty$ and $\sigma_n \to \infty$.

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

Next, we study precise large deviations for the partial sums S_n defined in (1.2). We will focus only on the case when ξ_0 has a right regularly varying tail (see Remark 2.2 below for information on other interesting cases). More precisely, we assume that there is a constant $t > 2$ such that

$$\mathbb{P}(\xi_0 \geq x) = \frac{h(x)}{x^t} \quad \text{as } x \to \infty. \tag{2.4}$$

Here $h(x)$ is a slowly varying function at ∞. Namely, $h(x)$ is a measurable positive function satisfying $\lim_{x\to\infty} h(\lambda x)/h(x) = 1$ for all constants $\lambda > 0$. Bingham *et al.* (1987) or Seneta (1976) contain systematic accounts on regularly varying functions. For the convenience of the reader, we collect some useful properties of slowly varying functions in Lemma A.3.

Note that condition (2.4) is an assumption on the right tail of ξ_0. The left tail of ξ_0 can be arbitrary. In particular, it implies that ξ_0 does not have pth moments for $p > t$, and it may or may not have pth moments for $p < t$.

The notion of regular variation such as defined in (2.4) is closely related to large deviation results for sums of random variables or processes. Such results were proved by Nagaev (1969a)–(1969c) and Nagaev (1979) for partial sums of i.i.d. random variables, and have been extended to partial sums of certain stationary sequences by Davis and Hsing (1995), Mikosch and Samorodnitsky (2000), Mikosch and Wintenberger (2013), and Peligrad *et al.* (2014).

We now comment briefly on the connections and differences of the results and methods in the aforementioned references to those in this paper. The approach of Davis and Hsing (1995) is based on weak convergence of point processes and the link between the large deviation probability and the asymptotic behavior of extremes. As shown in Davis and Hsing (1995, Example 5.5), their results are applicable to a class of linear processes with short-range dependence. Moreover, as pointed out by Mikosch and Wintenberger (2013, p. 853), the method of Davis and Hsing (1995) could not be extended to the case of $t \geq 2$.

Mikosch and Samorodnitsky (2000) studied precise large deviation results for a class linear processes with a negative drift. More specifically, they considered

$$X_n = -\mu + \sum_{j\in\mathbb{Z}} \varphi_{n-j}\varepsilon_j, \qquad n \in \mathbb{Z},$$

where $\mu > 0$ is a constant, $\{\varepsilon_j\}$ are i.i.d. innovations that satisfy a two-sided version of (2.4), and the coefficients $\{\varphi_j\}$ satisfy $\sum_{j\in\mathbb{Z}} |j\varphi_j| < \infty$. In particular, the process $\{X_n,\ n \in \mathbb{Z}\}$ is short-range dependent.

Mikosch and Wintenberger (2013) established precise large deviation results for a stationary sequence $\{X_n,\ n \in \mathbb{Z}\}$ that satisfies the following (and some other technical) conditions:

(i) all finite-dimensional distributions of $\{X_n,\ n \in \mathbb{Z}\}$ are regularly varying with the same index α;

(ii) the anti-clustering conditions.

See Mikosch and Wintenberger (2013) for precise descriptions of these conditions. We remark that, even though their methods and results cover a wide class of stationary sequences, condition (i) is a lot stronger than (2.4) and is not easy to verify for a general linear process. Moreover, as pointed out by Mikosch and Wintenberger (2013, p. 856), condition (ii) excludes stationary sequences with 'long-range dependencies of extremes'.

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

We believe that it would be interesting from both theoretical and application viewpoints to extend the large deviation results of Davis and Hsing (1995), Mikosch and Samorodnitsky (2000), Mikosch and Wintenberger (2013), and Peligrad *et al.* (2014) to stationary random fields. The present paper is one step in this direction. More specifically, we follow the approach of Peligrad *et al.* (2014) and prove the following precise large deviation theorem, which is applicable to linear random fields with long-range dependence.

Theorem 2.2. *Assume that* $\{b_{n,r,s}, r, s \in \mathbb{Z}\}$ *is a sequence of positive numbers with* $\rho_n^2 \to 0$ *as* $n \to \infty$ *and* ξ_0 *satisfies condition (2.4) for a certain constant* $t > 2$. *For* $x = x_n \geq C_t[\ln(U_{nt}^{-1})]^{1/2}$, *where* $C_t > \mathrm{e}^{t/2}(t+2)/\sqrt{2}$ *is a constant, we have*

$$\begin{aligned}\mathbb{P}(S_n \geq x) &= (1+o(1)) \sum_{r,s\in\mathbb{Z}} \mathbb{P}(b_{n,r,s}\xi_{-r,-s} \geq x) \\ &= (1+o(1))x^{-t} \sum_{r,s\in\mathbb{Z}} b_{n,r,s}^t h\left(\frac{x}{b_{n,r,s}}\right) \quad \text{as } n \to \infty. \end{aligned} \tag{2.5}$$

Proof. Since the second equality in (2.5) follows directly from the first and (2.4), we only need to prove the first equality. The proof is essentially a modification of the proof of Peligrad *et al.* (2014, Theorem 2.2, supplementary material) by replacing the quantities c_{ni} there by $b_{n,r,s}$, and the sum $\sum_{i=1}^{k_n}$ by the double sum $\sum_{r,s\in\mathbb{Z}}$. A somewhat new ingredient for the proof is to use a new version of the Fuk–Nagaev inequality for the double sums of infinitely many random variables which is stated as Theorem A.1. Hence, we will only sketch the main steps of the proof.

Without loss of generality, we normalize the partial sum S_n by its variance and assume that

$$\sum_{r,s\in\mathbb{Z}} b_{n,r,s}^2 = 1 \quad \text{and} \quad \rho_n^2 = \max_{r,s\in\mathbb{Z}} b_{n,r,s}^2 \to 0 \quad \text{as } n \to \infty. \tag{2.6}$$

Then, for any constant $t > 2$, we have

$$U_{nt} = D_{nt} \quad \text{and} \quad D_{nt} = \sum_{r,s\in\mathbb{Z}} b_{n,r,s}^t \leq \max_{r,s\in\mathbb{Z}} b_{n,r,s}^{t-2} \to 0,$$

which implies that $D_{nt}^{-1} \to \infty$. Moreover, the sequence S_n is stochastically bounded (i.e. $\lim_{K\to\infty} \sup_n \mathbb{P}(|S_n| > K) = 0$) since $\mathbb{E}(S_n^2) = 1$. By following the proofs of Peligrad *et al.* (2014, Lemma 4.1 and Proposition 4.1, supplementary material), for any $0 < \eta < 1$, $\varepsilon > 0$ such that $1 - \eta > \varepsilon$, and any $x_n \to \infty$, we have

$$\begin{aligned}&\left|\mathbb{P}(S_n \geq x_n) - \mathbb{P}(S_n^{(\varepsilon x_n)} \geq x_n) - \sum_{r,s\in\mathbb{Z}} \mathbb{P}(b_{n,r,s}\xi_{-r,-s} \geq (1-\eta)x_n)\right| \\ &\quad \leq o(1) \sum_{r,s\in\mathbb{Z}} \mathbb{P}(b_{n,r,s}\xi_{-r,-s} \geq \varepsilon x_n) \\ &\qquad + \sum_{r,s\in\mathbb{Z}} \mathbb{P}((1-\eta)x_n \leq b_{n,r,s}\xi_{-r,-s} < (1+\eta)x_n), \end{aligned} \tag{2.7}$$

where $S_n^{(\varepsilon x_n)} = \sum_{r,s\in\mathbb{Z}} b_{n,r,s}\xi_{-r,-s}\,\mathbf{1}(b_{n,r,s}\xi_{-r,-s} < \varepsilon x_n)$, $o(1)$ depends on the sequence x_n, η, and ε and converges to 0 as $n \to \infty$. See also Peligrad *et al.* (2014, Lemma 4.2 and Remark 4.1, supplementary material) for sums of infinitely many random variables.

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

As in the proof of Peligrad *et al.* (2014, Theorem 2.2, supplementary material), by analyzing the two terms on the right-hand side and the last term on the left-hand side of (2.7), we derive, for any fixed $\varepsilon > 0$,

$$\mathbb{P}(S_n \geq x) = (1 + o(1)) \sum_{r,s\in\mathbb{Z}} \mathbb{P}(b_{n,r,s}\xi_{-r,-s} \geq x) + \mathbb{P}(S_n^{(\varepsilon x)} \geq x) \quad \text{as } n \to \infty. \tag{2.8}$$

It remains to show that the term $\mathbb{P}(S_n^{(\varepsilon x)} \geq x)$ is negligible compared with the first term in (2.8). To this end, we apply Theorem A.1 to the sequence $\{b_{n,r,s}\xi_{-r,-s},\ r, s \in \mathbb{Z}\}$ with $y = \varepsilon x$ to derive, for any constant $m > t$,

$$\mathbb{P}(S_n^{(\varepsilon x)} \geq x) \leq \exp\left(-\frac{\alpha^2 x^2}{2\mathrm{e}^m}\right) + \left(\frac{A_n(m;\ 0, \varepsilon x)}{\beta\varepsilon^{m-1}x^m}\right)^{\beta/\varepsilon}, \tag{2.9}$$

where $\beta = m/(m+2), \alpha = 1-\beta = 2/(m+2)$, and we have used the fact that $B_n^2(-\infty, \varepsilon x) \leq 1$, which follows from (2.6).

Then, following the proof of Peligrad *et al.* (2014, Theorem 2.2, supplementary material), we can show that, for all $x = x_n \geq C_t[\ln(U_{nt}^{-1})]^{1/2}$, where $C_t > \mathrm{e}^{t/2}(t+2)/\sqrt{2}$ is a constant, we have

$$\begin{aligned}
&\exp\left(-\frac{\alpha^2 x^2}{2\mathrm{e}^m}\right) + \left(\frac{A_n(m;\ 0, \varepsilon x)}{\beta\varepsilon^{m-1}x^m}\right)^{\beta/\varepsilon} \\
&\quad = o(1) \sum_{r,s\in\mathbb{Z}} \frac{b_{n,r,s}^t}{x^t} h\left(\frac{x}{b_{n,r,s}}\right) \\
&\quad = o(1) \sum_{r,s\in\mathbb{Z}} \sum_{r,s\in\mathbb{Z}} \mathbb{P}(b_{n,r,s}\xi_{-r,-s} \geq x) \quad \text{as } n \to \infty.
\end{aligned} \tag{2.10}$$

In particular, we use the observation

$$D_{nt} = \sum_{r,s\in\mathbb{Z}} b_{n,r,s}^{2\eta} b_{n,r,s}^{t-2\eta} \leq \left(\sum_{r,s\in\mathbb{Z}} b_{n,r,s}^2\right)^{\eta} \left(\sum_{r,s\in\mathbb{Z}} b_{n,r,s}^{(t-2\eta)/(1-\eta)}\right)^{1-\eta}. \tag{2.11}$$

We have omitted the details for the derivation of (2.10) as it is very similar to the proof of Peligrad *et al.* (2014, Theorem 2.2, supplementary material). Finally, by combining (2.8)–(2.10), we obtain the first equality in (2.5). This completes the proof of Theorem 2.2. □

Remark 2.2. In addition to the case of regularly varying tails such as (2.4), large deviation results for sums of independent random variables or linear processes have been studied by several authors under the following two conditions:

- ξ_0 satisfies the Cramér condition: there exists a constant $h_0 > 0$ such that $\mathbb{E}(\mathrm{e}^{h\xi_0}) < \infty$ for $|h| \leq h_0$;
- ξ_0 satisfies the Linnik condition: there is a constant $\gamma \in (0, 1)$ such that $\mathbb{E}(\mathrm{e}^{|\xi_0|^\gamma}) < \infty$.

See, for example, Nagaev (1979), Jiang *et al.* (1995), Saulis and Statulevičius (2000), Djellout and Guillin (2001), Ghosh and Samorodnitsky (2009), and Li *et al.* (2009).

In light of Theorem 2.2 and the above discussions, it may prove interesting to study the following problems:

- the precise large deviation problems for linear random fields under the Cramér and Linnik conditions;

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

- an extension of the methods of Davis and Hsing (1995), Mikosch and Samorodnitsky (2000), and Mikosch and Wintenberger (2013) to establish large deviation results for stationary random fields.

Note that the tail conditions of Theorems 2.1 and 2.2 are different since one involves the moment and the other just involves the right tail behavior. Combining these conditions, we have the following tail probabilities over all $x_n \geq c$ for some $c > 0$. This theorem is a natural extension of the uniform moderate and large deviations for sums of i.i.d. random variables (see Nagaev (1979, Theorem 1.9)).

Theorem 2.3. *Assume that ξ_0 satisfies $\|\xi_0\|_p < \infty$ for some $p > 2$ and the right tail condition (2.4) for some constant $t > 2$. Assume also that $b_{n,r,s} > 0$ and $\rho_n^2 \to 0$ as $n \to \infty$. Let $(x_n)_{n\geq 1}$ be any sequence such that, for some $c > 0$, we have $x_n \geq c$ for all n. Then, as $n \to \infty$,*

$$\mathbb{P}(S_n \geq x_n\sigma_n) = (1 + o(1))\left[x_n^{-t} \sum_{r,s\in\mathbb{Z}} b_{n,r,s}^t h\left(\frac{x_n}{b_{n,r,s}}\right) + 1 - \Phi(x_n)\right]. \tag{2.12}$$

Proof. The proof is a modification of the proof of Peligrad *et al.* (2014, Theorem 2.1 and Corollary 2.3(i), supplementary material). We sketch the proof here for completeness. Without loss of generality, we may assume that $2 < p < t$. Let $x = x_n \to \infty$. For simplicity, we assume that (2.6) holds. Under the condition in this theorem, as in the proof of Theorem 2.2, it follows that (2.8) holds. Denote

$$X'_{n,r,s} = b_{n,r,s}\xi_{-r,-s}\,\mathbf{1}(b_{n,r,s}\xi_{-r,-s} \leq \varepsilon x).$$

We now apply Lemma A.2 to the second term on the right-hand side of (2.8). To this end, we decompose the sum $S_n^{(\varepsilon x)}$ as a finite sum, i.e. the sum of $\sum_{|r|\leq k_n}\sum_{|s|\leq k_n} X'_{n,r,s}$ with $(2k_n+1)^2$ terms for some k_n and the remainder

$$R'_n = \left(\sum_{|r|>k_n}\sum_{s\in\mathbb{Z}} + \sum_{|r|\leq k_n}\sum_{|s|>k_n}\right) X'_{n,r,s}.$$

By Rosenthal's inequality (see De la Peña and Giné (1999)), it is easy to derive

$$\mathbb{E}|R'_n|^p \leq 2C'_p\mathbb{E}|\xi_0|^p \sum_{r,s} |b_{n,r,s}|^p \quad \text{for some constant } C'_p.$$

Consequently, the quantity L_{np} in Lemma A.2 is bounded by

$$L_{np} \leq (2C'_p + 1)\mathbb{E}|\xi_0|^p \sum_{r,s} |b_{n,r,s}|^p = (2C'_p + 1)D_{np}\mathbb{E}|\xi_0|^p.$$

See also the proof of Peligrad *et al.* (2014, Corollary 2.3(i), supplementary material). Then, by Lemma A.2, if $x^2 \leq c\ln((2C'_p + 1)D_{np}\mathbb{E}|\xi_0|^p)^{-1}$ for $c < 1/\varepsilon$, we have $x^2 \leq c\ln(L_{np}^{-1})$ for $c < 1/\varepsilon$ and

$$\mathbb{P}(S_n^{(\varepsilon x)} \geq x) = (1 - \Phi(x))(1 + o(1)). \tag{2.13}$$

Note that (2.13) also holds for $x^2 \leq c\ln(D_{np})^{-1}$ for any $c < 1/\varepsilon$ and large enough n since $D_{np} \to 0$. Recall that $2 < p < t$. By applying (2.11) with $\eta = (t - p)/(t - 2)$, we have

$$D_{nt} \ll D_{np} \ll (D_{nt})^{(p-2)/(t-2)}.$$

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

Then (2.12) holds for $0 < x \le C[\ln(D_{nt}^{-1})]^{1/2}$ with C an arbitrary positive number. On the other hand, there is a constant $c_1 > 0$ such that, for $x > c_1[\ln(D_{nt}^{-1})]^{1/2}$, we also have

$$\mathbb{P}(S_n \ge x) = (1 + o(1)) \sum_{r,s\in\mathbb{Z}} \mathbb{P}(b_{n,r,s}\xi_0 \ge x), \qquad 1 - \Phi(x) = o\bigg(\sum_{r,s\in\mathbb{Z}} \mathbb{P}(b_{n,r,s}\xi_0 \ge x) \bigg).$$

See also the proof of Peligrad *et al.* (2014, Theorem 2.1, supplementary material). By choosing $c_1 < C$, (2.12) holds for all $x = x_n \to \infty$. If the sequence x_n is bounded, by Theorem 2.1, we have the moderate deviation result. Since $x_n \ge c > 0$, the expression in square brackets on the right-hand side of (2.12) is dominating as $n \to \infty$. □

3. Applications

In this section we provide two applications of the main results from Section 2, one to nonparametric regression and the other to the Davis–Gut law for linear random fields.

3.1. Nonparametric regression

We first provide an application of the deviation results for the nonparametric regression estimate. Consider the following regression model:

$$Y_{n,j,k} = g(z_{n,j,k}) + X_{n,j,k}, \qquad (j,k) \in \Gamma_n,$$

where g is a bounded continuous function on $\mathbb{R}^d$, the $z_{n,j,k}$ are the fixed design points over $\Gamma_n \subseteq \mathbb{Z}^2$ with values in a compact subset of $\mathbb{R}^d$, and $X_{n,j,k} = \sum_{r,s\in\mathbb{Z}} a_{r,s}\xi_{n,j-r,k-s}$ is a linear random field over $\mathbb{Z}^2$ with mean-zero i.i.d. innovations $\xi_{n,r,s}$. Regression models with independent or weakly dependent random field errors have been studied by several authors; see, for example, El Machkouri (2007), El Machkouri and Stoica (2010), and Hallin *et al.* (2004b). For related papers that deal with density estimations for random fields; see, for example, Tran (1990) and Hallin *et al.* (2004a).

An estimator for the function g on the basis of sample pairs $(z_{n,j,k}, Y_{n,j,k})$, $(j,k) \in \Gamma_n$ is the following general linear smoother:

$$g_n(z) = \sum_{(j,k)\in\Gamma_n} w_{n,j,k}(z) Y_{n,j,k},$$

where the $w_{n,j,k}(\cdot)$ are weight functions on $\mathbb{R}^d$. In the particular case of kernel regression estimation, $w_{n,j,k}(z)$ has the form

$$w_{n,j,k}(z) = \frac{K((z - z_{n,j,k})/h_n)}{\sum_{(j',k')\in\Gamma_n} K((z - z_{n,j',k'})/h_n)},$$

where $K : \mathbb{R}^d \to \mathbb{R}^+$ is a kernel function and h_n is a sequence of bandwidths which goes to 0 as $|\Gamma_n| \to \infty$. Gu and Tran (2009) developed a central limit theorem and the bias for the fixed design regression estimate $g_n(z)$ in the case when $X_{n,j,k} = \xi_{n,j,k}$ for all $(j,k) \in \Gamma_n$.

Theorems 2.1–2.3 in Section 2 can be applied to study, for every $z \in \mathbb{R}^d$, the speed of the almost sure convergence of $g_n(z) - \mathbb{E}g_n(z) \to 0$, or $g_n(z) \to g(z)$ if the weight functions are chosen to satisfy the condition $\sum_{(j,k)\in\Gamma_n} w_{n,j,k}(z) = 1$, which is the case in kernel regression estimation.

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

Let $S_n := g_n(z) - \mathbb{E}g_n(z)$. This can be written as

$$S_n = \sum_{(j,k)\in\Gamma_n} w_{n,j,k}(z) X_{n,j,k} = \sum_{r,s\in\mathbb{Z}} b_{n,r,s}\xi_{n,-r,-s},$$

where $b_{n,r,s} = \sum_{(j,k)\in\Gamma_n} w_{n,j,k}(z) a_{j+r,k+s}$. We choose the weight functions $w_{n,j,k}(z)$, the coefficients $\{a_{r,s}\}$, and the random variable ξ_0 to satisfy the conditions of Theorem 2.1. Then, for $x_n = \sqrt{2\ln(U_{np}^{-1})}$ (see Section 2 for the definition of U_{np}), we have

$$\mathbb{P}(|S_n| \geq x_n\sigma_n) = 2(1-\Phi(x_n))(1+o(1)) \quad \text{as } n\to\infty,$$

where $\sigma_n^2 = \sum_{r,s\in\mathbb{Z}} b_{n,r,s}^2 = \operatorname{var}(g_n(z))$.

If $U_{np} \to 0$ fast enough such that $\sum_n (1-\Phi(x_n)) < \infty$ then we can derive, by using the Borel–Cantelli lemma, the following upper bound on the speed of convergence of $g_n(z) - \mathbb{E}g_n(z)$:

$$\limsup_{n\to\infty} \frac{|g_n(z) - \mathbb{E}g_n(z)|}{\sigma_n\sqrt{2\ln(U_{np}^{-1})}} \leq 1 \quad \text{almost surely.} \tag{3.1}$$

Under further conditions, we may write (3.1) in a more familiar form. Since $p > 2$, we have $U_{np} \leq |\rho_n|^{p-2}$. If we have information on the rate for $\rho_n \to 0$, say, $|\rho_n| \leq (\ln n)^{-1/(p-2)}$, then we obtain an upper bound which coincides with the law of the iterated logarithm:

$$\limsup_{n\to\infty} \frac{|g_n(z) - \mathbb{E}g_n(z)|}{\sigma_n\sqrt{2\ln\ln n}} \leq 1 \quad \text{almost surely.} \tag{3.2}$$

In the particular case of $X_{n,j,k} = \xi_{n,j,k}$, $b_{n,r,s} = w_{n,r,s}(z)$ if $(r,s) \in \Gamma_n$ and, otherwise $b_{n,r,s} = 0$, we have $D_{nt} = \sum_{(r,s)\in\Gamma_n} w_{n,r,s}^t(z)$, $U_{nt} = (D_{n2})^{-t/2} D_{nt}$, and

$$\sigma_n^2 = \mathbb{E}(S_n^2) = D_{n2} = \sum_{(r,s)\in\Gamma_n} w_{n,r,s}^2(z).$$

Hence, under certain conditions on the weight functions $w_{n,j,k}(z)$, we can obtain, from (3.1) or (3.2), the speed of convergence of $g_n(z) - \mathbb{E}g_n(z) \to 0$, which complements the results of Gu and Tran (2009).

3.2. A Davis–Gut law of the iterated logarithm

Now we apply the moderate deviation result (Theorem 2.1) to prove a Davis–Gut-type law for linear random fields; see Davis (1968), Gut (1980), Li (1991), and Li and Rosalsky (2007) for the Davis–Gut laws for partial sums of i.i.d. random variables. The Davis–Gut-type law for linear processes with short memory (short-range dependence) was developed by Chen and Wang (2008).

For a linear random field defined in (1.1), we consider the partial sum (1.2) with $\Gamma_n = [1,n]^2 \cap \mathbb{Z}^2$, and assume the following condition:

(DG) $\|\xi_0\|_p < \infty$ for some $p > 2$ and $\{a_{r,s}\}$ satisfies either

$$A := \sum_{r,s\in\mathbb{Z}} |a_{r,s}| < \infty, \qquad a := \sum_{r,s\in\mathbb{Z}} a_{r,s} \neq 0, \tag{3.3}$$

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

or

$$a_{r,s} = (|r| + |s|)^{-\beta} L(|r| + |s|) b\left(\frac{r}{\sqrt{r^2 + s^2}}, \frac{s}{\sqrt{r^2 + s^2}} \right) \quad \text{for } r \neq 0 \text{ or } s \neq 0, \tag{3.4}$$

where $\beta \in (1, 2)$, $L(\cdot)$ is a slowly varying function at ∞, and $b(\cdot, \cdot)$ is a bounded piecewise continuous function defined on the unit circle.

Under condition (3.4), $\sum_{r,s \in \mathbb{Z}} |a_{r,s}| = \infty$. In the literature, the random field (1.1) is said to have long memory or long-range dependence. In the following lemma we provide the order of the quantity D_{np} (see the definition in Section 2) under condition (3.4). Recall that $a_n \propto b_n$ means that $a_n / b_n \to C$ as $n \to \infty$ for some constant $C > 0$.

Lemma 3.1. *Assume that (3.4) holds. Then, for $p > 2$,*

$$D_{np} = \sum_{r,s \in \mathbb{Z}} |b_{n,r,s}|^p = O(n^{p(2-\beta)+2} L^p(n)).$$

Proof. We use the properties of slowly varying functions as stated in Lemma A.3, the condition $1 < \beta < 2$, and, thus, $1 - \beta > -1$ and $1 - p\beta < -1$ throughout the proof. We also use $C > 0$ as a generic constant in the proof. First we consider the $r > n$ case. Since $b(\cdot, \cdot)$ is bounded,

$$\begin{aligned} |b_{n,r,s}| &\leq C \sum_{j,k=1}^{n} (j + r + |k + s|)^{-\beta} L(j + r + |k + s|) \\ &\propto n \sum_{k=1}^{n} (r + |k + s|)^{-\beta} L(r + |k + s|) \\ &\propto n^2 (r + |s|)^{-\beta} L(r + |s|). \end{aligned}$$

Then

$$\begin{aligned} \sum_{s \in \mathbb{Z}, r > n} |b_{n,r,s}|^p &= \sum_{|s| \leq n,\, r > n} |b_{n,r,s}|^p + \sum_{|s| > n,\, r > n} |b_{n,r,s}|^p \\ &\leq Cn \sum_{r > n} n^{2p} r^{-p\beta} L^p(r) + 2C \sum_{s > n,\, r > n} n^{2p} (r + s)^{-p\beta} L^p(r + s) \\ &\propto n^{2p+1} n^{1-p\beta} L^p(n) + \sum_{r > n} n^{2p} (r + n)^{1-p\beta} L^p(r + n) \\ &\propto n^{p(2-\beta)+2} L^p(n) + n^{2p} n^{2-p\beta} L^p(n) \\ &= 2n^{p(2-\beta)+2} L^p(n). \end{aligned} \tag{3.5}$$

For the $r < -2n$ case, let $R = -r - n$, then $R > n$, and

$$\begin{aligned} |b_{n,r,s}| &\leq C \sum_{j,k=1}^{n} (-j - r + |k + s|)^{-\beta} L(-j - r + |k + s|) \\ &= C \sum_{j,k=1}^{n} (-j + n + R + |k + s|)^{-\beta} L(-j + n + R + |k + s|) \\ &\propto n^2 (|r| + |s|)^{-\beta} L(|r| + |s|). \end{aligned}$$

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

Hence, similarly to (3.5), we have

$$\sum_{s\in\mathbb{Z},\, r<-2n} |b_{n,r,s}|^p = O(n^{p(2-\beta)+2}L^p(n)). \tag{3.6}$$

By symmetry, we also have

$$\sum_{r\in\mathbb{Z},\, s>n} |b_{n,r,s}|^p = O(n^{p(2-\beta)+2}L^p(n)), \tag{3.7}$$

$$\sum_{r\in\mathbb{Z},\, s<-2n} |b_{n,r,s}|^p = O(n^{p(2-\beta)+2}L^p(n)). \tag{3.8}$$

In the $-2n \le r, s \le n$ case,

$$\begin{aligned}
|b_{n,r,s}| &\le C \sum_{j,k=1}^{n} (|j+r|+|k+s|)^{-\beta} L(|j+r|+|k+s|) \\
&\le 4C \sum_{j,k=1}^{2n} (j+k)^{-\beta} L(j+k) \\
&\ll \sum_{j=1}^{2n} j^{1-\beta} \max_{1\le k\le 2n} L(j+k) \\
&\propto n^{2-\beta} \max_{1\le k\le 2n} L(2n+k) \propto n^{2-\beta} L(n).
\end{aligned}$$

Hence,

$$\sum_{-2n\le s,\, r\le n} |b_{n,r,s}|^p \ll n^2 n^{p(2-\beta)} L^p(n) = n^{p(2-\beta)+2} L^p(n). \tag{3.9}$$

Combining (3.5)–(3.9), we complete the proof of the lemma. □

In the next theorem we provide a Davis–Gut-type law for linear random fields that satisfies condition (DG).

Theorem 3.1. *Assume that condition (DG) holds. Let $h(\cdot)$ be a positive nondecreasing function on $[c,\infty)$ for some constant $c \ge 1$, such that $\int_c^\infty (th(t))^{-1}\,\mathrm{d}t = \infty$. Let $\Psi(t) = \int_c^t (sh(s))^{-1}\,\mathrm{d}s$, $t \ge c$. Let $m = \arg\min_{t\ge c,\, t\in\mathbb{N}}\{\Psi(t) > 1\}$. Then, for real numbers ε and $n \ge m$, we have*

$$\mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n\sqrt{2\ln\Psi(n)}) \propto \frac{1}{\sqrt{\ln\Psi(n)}}\Psi(n)^{-(1+\varepsilon)^2}.$$

Define

$$S_\Psi := \sum_{n=m}^{\infty} \frac{1}{nh(n)} \mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n\sqrt{2\ln\Psi(n)}).$$

Then $S_\Psi < \infty$ if $\varepsilon > 0$ and $S_\Psi = \infty$ if $\varepsilon \le 0$.

Proof. First we consider the short memory case (3.3). Recall that $a = \sum_{r,s\in\mathbb{Z}} a_{r,s} \ne 0$. Under condition (3.3), since

$$\sigma_n^2 = \sum_{r,s\in\mathbb{Z}} b_{n,r,s}^2 = \sum_{r,s\in\mathbb{Z}} \left(\sum_{j=0}^{n}\sum_{k=0}^{n} a_{j+r,k+s}\right)^2,$$

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

it is easy to see that $\sigma_n^2/n^2 - a^2 \to 0$. Hence, $a^2n^2/\sigma_n^2 \to 1$ as $n \to \infty$. Also the numbers $b_{n,r,s}$ that satisfy $|b_{n,r,s}| \geq 1$ are at most $\lceil a^2n^2 \rceil$ asymptotically. Then

$$\begin{aligned} D_{np} &= \sum_{r,s\in\mathbb{Z}} |b_{n,r,s}|^p \\ &= \sum_{r,s\in\mathbb{Z},\ |b_{n,r,s}|\geq 1} |b_{n,r,s}|^p + \sum_{r,s\in\mathbb{Z},\ |b_{n,r,s}|<1} |b_{n,r,s}|^p \\ &\leq \sum_{r,s\in\mathbb{Z},\ |b_{n,r,s}|\geq 1} A^p + \sum_{r,s\in\mathbb{Z},\ |b_{n,r,s}|<1} b_{n,r,s}^2 \\ &\leq \lceil a^2n^2 \rceil A^p + \sigma_n^2 \end{aligned}$$

has order $O(n^2)$ for $p > 2$. Therefore, $\ln(U_{np}^{-1}) = \ln(\sigma_n^p/D_{np}) \geq (p-2)\ln n$.

Next we study the long memory case. Under condition (3.4), Lemma 3.1 yields $D_{np} = O(n^{p(2-\beta)+2}L^p(n))$. On the other hand, by Surgailis (1982, Theorem 2),

$$\sigma_n^2 = \sum_{r,s\in\mathbb{Z}} b_{n,r,s}^2 = c_\beta n^{6-2\beta}L^2(n)$$

for some constant c_β depending only on β. Hence, we also have $\ln(U_{np}^{-1}) = \ln(\sigma_n^p/D_{np}) \geq (p-2)\ln n$.

By the definition of $\Psi(t)$, $\Psi(n) \leq \int_c^n (sh(c))^{-1}\,\mathrm{d}s \leq \ln n/h(c)$. Let $x_n = (1+\varepsilon)\sqrt{2\ln\Psi(n)}$. Then $x_n^2 \ll 2(p-2)\ln n \leq 2\ln(U_{np}^{-1})$. By Remark 2.1, $\sigma_n \to \infty$ implies that $\rho_n \to 0$ in our case. Then, by Theorem 2.1,

$$\begin{aligned} &\mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n\sqrt{2\ln\Psi(n)}) \\ &\quad = 2(1-\Phi(x_n))(1+o(1)) \\ &\quad = 2(2\pi)^{-1/2}\frac{1}{(1+\varepsilon)\sqrt{2\ln\Psi(n)}}\exp(-(1+\varepsilon)^2\ln\Psi(n))(1+o(1)) \qquad (3.10) \\ &\quad \propto \frac{1}{\sqrt{\ln\Psi(n)}}(\Psi(n))^{-(1+\varepsilon)^2}. \end{aligned}$$

In (3.10) we have used the well-known inequality

$$\frac{1}{(2\pi)^{1/2}(1+x)}\mathrm{e}^{-x^2/2} \leq 1-\Phi(x) \leq \frac{1}{(2\pi)^{1/2}x}\mathrm{e}^{-x^2/2} \quad \text{for } x > 1.$$

Therefore,

$$\begin{aligned} S_\Psi &= \sum_{n=m}^{\infty}\frac{1}{nh(n)}\mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n\sqrt{2\ln\Psi(n)}) \\ &\propto \sum_{n=m}^{\infty}\frac{1}{nh(n)\sqrt{\ln\Psi(n)}\Psi(n)^{(1+\varepsilon)^2}} \\ &= \sum_{n=m}^{\infty}\frac{\Psi'(n)}{\sqrt{\ln\Psi(n)}\Psi(n)^{(1+\varepsilon)^2}}. \end{aligned}$$

It is clear that $S_\Psi < \infty$ if $\varepsilon > 0$ and $S_\Psi = \infty$ if $\varepsilon \leq 0$. □

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

Corollary 3.1. *Assume that condition (DG) holds. Let*

$$S = \sum_{n=3}^{\infty} \frac{1}{n(\ln \ln n)^b} \mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n \sqrt{2 \ln \ln n}).$$

Then, for any $b \in \mathbb{R}$, $S < \infty$ *if* $\varepsilon > 0$ *and* $S = \infty$ *if* $\varepsilon < 0$. *If* $\varepsilon = 0$, $S < \infty$ *if* $b > \frac{1}{2}$ *and* $S = \infty$ *if* $b \le \frac{1}{2}$.

Proof. Let $h(t) = 1$ and $c = 1$. Then $\Psi(n) = \ln n$. By Theorem 3.1, for $n \ge 3$,

$$\mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n \sqrt{2 \ln \ln n}) \propto \frac{1}{\sqrt{\ln \ln n}} (\ln n)^{-(1+\varepsilon)^2}.$$

For any $b \in \mathbb{R}$,

$$\begin{aligned} S &= \sum_{n=3}^{\infty} \frac{1}{n(\ln \ln n)^b} \mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n \sqrt{2 \ln \ln n}) \\ &\propto \sum_{n=3}^{\infty} \frac{1}{n(\ln \ln n)^{b+1/2}} (\ln n)^{-(1+\varepsilon)^2}. \end{aligned} \tag{3.11}$$

It is clear that $S < \infty$ if $\varepsilon > 0$ and $S = \infty$ if $\varepsilon < 0$. In the $\varepsilon = 0$ case, by (3.11), it is easy to see that $S < \infty$ if $b > \frac{1}{2}$ and $S = \infty$ if $b \le \frac{1}{2}$. □

Corollary 3.2. *Assume that condition (DG) holds. For* $0 \le r < 1$, *let*

$$S_r = \sum_{n=3}^{\infty} \frac{1}{n(\ln n)^r} \mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n \sqrt{2(1-r) \ln \ln n}).$$

Then $S_r < \infty$ *if* $\varepsilon > 0$ *and* $S_r = \infty$ *if* $\varepsilon \le 0$.

Proof. Let $h(t) = (\ln t)^r/(1-r)$, $c = 1$. Then $\Psi(n) = (\ln n)^{1-r}$. By Theorem 3.1, for $n \ge 3$,

$$\mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n \sqrt{2(1-r) \ln \ln n}) \propto \frac{1}{\sqrt{\ln \ln n}} (\ln n)^{-(1+\varepsilon)^2(1-r)}$$

and

$$\begin{aligned} S_r &= \sum_{n=3}^{\infty} \frac{1}{n(\ln n)^r} \mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n \sqrt{2(1-r) \ln \ln n}) \\ &\propto \sum_{n=3}^{\infty} \frac{1}{n\sqrt{\ln \ln n}(\ln n)^{(1+\varepsilon)^2(1-r)+r}} \\ &= \sum_{n=3}^{\infty} \frac{1}{n\sqrt{\ln \ln n}(\ln n)^{1+(2\varepsilon+\varepsilon^2)(1-r)}}. \end{aligned}$$

It is clear that $S_r < \infty$ if $\varepsilon > 0$ and $S_r = \infty$ if $\varepsilon \le 0$. □

Corollary 3.3. *Assume that condition (DG) holds. Let*

$$S = \sum_{n=16}^{\infty} \frac{1}{n \ln n} \mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n \sqrt{2 \ln \ln \ln n}).$$

Then $S < \infty$ *if* $\varepsilon > 0$ *and* $S = \infty$ *if* $\varepsilon \le 0$.

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

Proof. Let $h(t) = \ln t$ and $c = \mathrm{e}$. Then $\Psi(n) = \ln\ln n$. By Theorem 3.1, for $n \geq 16$,

$$\mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n\sqrt{2\ln\ln\ln n}) \propto \frac{1}{\sqrt{\ln\ln\ln n}}(\ln\ln n)^{-(1+\varepsilon)^2}$$

and

$$S = \sum_{n=16}^{\infty} \frac{1}{n\ln n}\mathbb{P}(|S_n| > (1+\varepsilon)\sigma_n\sqrt{2\ln\ln\ln n}) \propto \sum_{n=16}^{\infty} \frac{1}{n\ln n(\ln\ln n)^{(1+\varepsilon)^2}\sqrt{\ln\ln\ln n}}.$$

It is clear that $S < \infty$ if $\varepsilon > 0$ and $S = \infty$ if $\varepsilon \leq 0$. □

Remark 3.1. We can also prove the Davis–Gut law for linear processes by applying the moderate deviation results for linear processes from Peligrad *et al.* (2014). Let $S_n = \sum_{k=1}^{n} X_k$, where

$$X_k = \sum_{j=-\infty}^{\infty} a_{k-j}\xi_j$$

and the innovations ξ_j are i.i.d. random variables with $\mathbb{E}\xi_j = 0$ and $\mathbb{E}\xi_j^2 = 1$. Consider the short memory case $\sum_{i=-\infty}^{\infty} |a_i| < \infty$, $a = \sum_{i=-\infty}^{\infty} a_i \neq 0$. Observe that $S_n = \sum_{i=-\infty}^{\infty} b_{ni}\xi_i$, where $b_{ni} = a_{1-i} + \cdots + a_{n-i}$. Then $\sigma_n^2 = \mathrm{var}(S_n) = \sum_i b_{ni}^2$. For the short memory case, it is well known that σ_n^2 has order n. Furthermore, $a^2 n/\sigma_n^2 \to 1$ as $n \to \infty$; $\sum_i |b_{ni}|^p$ has order n for $p > 2$. Let

$$U_{np} = \left(\sum_i b_{ni}^2\right)^{-p/2} \sum_i |b_{ni}|^p.$$

Then $\ln(U_{np}^{-1}) = \ln[(\sum_i b_{ni}^2)^{p/2}/\sum_i |b_{ni}|^p] \sim \frac{1}{2}(p-2)\ln n$. Let $h(t)$ and $\Psi(t)$ be the functions defined as in Theorem 3.1. Hence,

$$x_n = (1+\varepsilon)\sqrt{2\ln\Psi(n)} \ll \sqrt{(p-2)\ln n} \sim \sqrt{2\ln(U_{np}^{-1})}. \tag{3.12}$$

Then using Peligrad *et al.* (2014, Corollary 3(iii)), the Davis–Gut-type laws, Theorem 3.1, and Corollaries 3.1–3.3 hold for short memory linear processes.

Remark 3.2. For the causal long memory linear process, $X_k = \sum_{j=0}^{\infty} a_{k-j}\xi_j$ with $\sum_{i=0}^{\infty}|a_i| = \infty$ and $\sum_{i=0}^{\infty} a_i^2 < \infty$, we assume that $a_n = (n+1)^{-\alpha}L(n+1)$, where $\frac{1}{2} < \alpha < 1$, and $L(n) > 0$ is a slowly varying function at ∞. Then $S_n = \sum_{i=1}^{\infty} b_{ni}\xi_{n-i}$, where $b_{ni} = \sum_{k=1}^{i} a_k$ for $i < n$ and $b_{ni} = \sum_{k=i-n+1}^{i} a_i$ for $i \geq n$. Also

$$\sigma_n^2 = \mathrm{var}(S_n) = \sum_{i=1}^{\infty} b_{ni}^2 \sim c_\alpha n^{3-2\alpha}L^2(n), \tag{3.13}$$

where

$$c_\alpha = \frac{1}{(1-\alpha)^2}\int_0^{\infty} [x^{1-\alpha} - \max(x-1,0)^{1-\alpha}]^2\,\mathrm{d}x.$$

The asymptotic equivalence in (3.13) is well known; see, for example, Wu and Min (2005, Theorem 2). On the other hand, there are constants C_1 and C_2 such that, for all $n \geq 1$,

$$b_{ni} \leq \begin{cases} C_1 i^{1-\alpha}L(i) & \text{for } i \leq 2n, \\ b_{ni} \leq C_2 n(i-n)^{-\alpha}L(i) & \text{for } i > 2n. \end{cases}$$

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

Hence, by the properties of slowly varying functions as stated in Lemma A.3,

$$\sum_i b_{ni}^p \ll \sum_{i<2n} i^{(1-\alpha)p} L^p(i) + \sum_{i\geq 2n} n^p (i-n)^{-\alpha p} L^p(i) \ll n^{1+p(1-\alpha)} L^p(n).$$

Therefore,

$$\ln(U_{np}^{-1}) = \ln\left[\frac{(\sum_i b_{ni}^2)^{p/2}}{\sum_i b_{ni}^p}\right] \gg \ln\left[\frac{n^{p(3-2\alpha)/2} L^p(n)}{n^{1+p(1-\alpha)} L^p(n)}\right] \sim \frac{1}{2}(p-2)\ln n.$$

Let $h(t)$ and $\Psi(t)$ be the functions defined as in Theorem 3.1. Hence, (3.12) still holds. Then, by Peligrad *et al.* (2014, Corollary 3(iii)), the Davis–Gut-type laws, Theorem 3.1, and Corollaries 3.1–3.3 hold for long memory linear processes.

Appendix A

In this appendix we first justify (1.1) and (1.2) in the Introduction, and then collect some results that are useful for the proofs in Section 2.

Lemma A.1. *Let* $\{\xi_{r,s},\ (r,s)\in\mathbb{Z}^2\}$ *and* ξ_0 *be i.i.d. random variables with* $\mathbb{E}\xi_0 = 0$ *and* $\mathbb{E}\xi_0^2 = 1$. *Let* $\{a_{r,s},\ (r,s)\in\mathbb{Z}^2\}$ *be a square-summable sequence of constants. Then the following statements hold:*

(i) *the series* $\sum_{r,s\in\mathbb{Z}} a_{r,s}\xi_{j-r,k-s}$ *converges in* $L^2(\Omega,\mathbb{P})$ *almost surely;*

(ii) *(1.2) holds in* $L^2(\Omega,\mathbb{P})$ *almost surely.*

Proof. We refer to the series $\sum_{r,s\in\mathbb{Z}} a_{r,s}\xi_{j-r,k-s}$ from (1.1). Let $\{\Upsilon_n,\ n\geq 1\}$ be an arbitrary sequence of finite subsets of $\mathbb{Z}^2$ that satisfies $\Upsilon_n\subset\Upsilon_{n+1}$ and $|\Upsilon_n|\to\infty$. Then, for any $m<n$,

$$\mathbb{E}\left[\left(\sum_{(r,s)\in\Upsilon_n\setminus\Upsilon_m} a_{r,s}\xi_{j-r,k-s}\right)^2\right] = \sum_{(r,s)\in\Upsilon_n\setminus\Upsilon_m} a_{r,s}^2,$$

which tends to 0 as $m\to\infty$. This implies that (1.1) converges in $L^2(\Omega,\mathbb{P})$.

Since the summands in (1.1) are independent random variables, the almost sure convergence of (1.1) follows from Kolmogorov's three-series theorem. Alternatively, it follows from Lévy's equivalence theorem, which states that the almost sure convergence is equivalent to convergence in probability or in law.

Next, let $n\geq 1$ be fixed. We write the partial sum S_n as

$$\begin{aligned} S_n &= \sum_{(j,k)\in\Gamma_n}\sum_{r,s\in\mathbb{Z}} a_{r,s}\xi_{j-r,k-s} \\ &= \sum_{(j,k)\in\Gamma_n}\sum_{r,s\in\mathbb{Z}} a_{j+r,k+s}\xi_{-r,-s} \\ &= \sum_{(j,k)\in\Gamma_n}\lim_{m\to\infty}\sum_{r,s\in[-m,m]} a_{j+r,k+s}\xi_{-r,-s} \\ &= \lim_{m\to\infty}\sum_{(j,k)\in\Gamma_n}\sum_{r,s\in[-m,m]} a_{j+r,k+s}\xi_{-r,-s}, \end{aligned}$$

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

where the limit is taken either in $L^2(\Omega, \mathbb{P})$ or in the almost sure sense. Since both index sets Γ_n and $[-m, m]^2$ are finite, we change the order of summation to obtain

$$\begin{aligned} S_n &= \lim_{m\to\infty} \sum_{r,s\in[-m,m]} \Bigg(\sum_{(j,k)\in\Gamma_n} a_{j+r,k+s} \Bigg) \xi_{-r,-s} \\ &= \sum_{r,s\in\mathbb{Z}} \Bigg(\sum_{(j,k)\in\Gamma_n} a_{j+r,k+s} \Bigg) \xi_{-r,-s} \\ &= \sum_{r,s\in\mathbb{Z}} b_{n,r,s} \xi_{-r,-s}. \end{aligned}$$

This verifies (ii). □

The following theorem is an extended version of the Fuk–Nagaev inequality (see Nagaev (1979, Corollary 1.7)) for a double sum of infinitely many random variables. See also Peligrad *et al.* (2014, Theorem 5.1 and Remark 5.1, supplementary material) for the extension of the Fuk–Nagaev inequality for the sum of infinitely many random variables.

Theorem A.1. *Let $(X_{ni})_{i\in\mathbb{N}^2}$ be a set of independent random variables with mean 0. For a constant $m \geq 2$, let $\beta = m/(m+2)$ and $\alpha = 1-\beta = 2/(m+2)$. For any $y > 0$, define $X_{ni}^{(y)} = X_{ni}\,\mathbf{1}(X_{ni} \leq y)$, $A_n(m;\, 0, y) := \sum_{i\in\mathbb{N}^2} \mathbb{E}(X_{ni}^m\,\mathbf{1}(0 < X_{ni} < y))$, and $B_n^2(-\infty, y) := \sum_{i\in\mathbb{N}^2} \mathbb{E}(X_{ni}^2\,\mathbf{1}(X_{ni} < y))$. Then, for any $x > 0$ and $y > 0$,*

$$\mathbb{P}\Bigg(\sum_{i\in\mathbb{N}^2} X_{ni}^{(y)} \geq x \Bigg) \leq \exp\Bigg(-\frac{\alpha^2 x^2}{2\mathrm{e}^m B_n^2(-\infty, y)} \Bigg) + \Bigg(\frac{A_n(m;\, 0, y)}{\beta x y^{m-1}} \Bigg)^{\beta x/y}.$$

The following result can be found in Peligrad *et al.* (2014, Theorem 5.2, supplementary material) and is an immediate consequence of Frolov (2005, Theorem 1.1).

Theorem A.2. *Let $(X_{nj})_{1\leq j\leq k_n}$ be an array of row-wise independent centered random variables. Let $S_n = \sum_{j=1}^{k_n} X_{nj}$ and $\sigma_n^2 = \sum_{j=1}^{k_n} \mathbb{E}X_{nj}^2$. For any positive numbers u, v, and ε, denote*

$$\Lambda_n(u, v, \varepsilon) = \frac{u}{\sigma_n^2} \sum_{j=1}^{k_n} \mathbb{E}(X_{nj}^2\,\mathbf{1}(X_{nj} \leq -\varepsilon\sigma_n/v)).$$

Assume that, for some constant $p > 2$, $M_{np} = \sum_{j=1}^{k_n} \mathbb{E}(X_{nj}^p\,\mathbf{1}(X_{nj} \geq 0)) < \infty$ and $L_{np} := \sigma_n^{-p} M_{np} \to 0$ as $n \to \infty$. If

$$\Lambda_n(x^4, x^5, \varepsilon) \to 0 \quad \textit{for any } \varepsilon > 0$$

and

$$x^2 - 2\ln(L_{np}^{-1}) - (p-1)\ln\ln(L_{np}^{-1}) \to -\infty \quad \textit{as } n \to \infty,$$

then

$$\mathbb{P}(S_n \geq x\sigma_n) = (1 - \Phi(x))(1 + o(1)).$$

The following lemma is useful in the proof of Theorem 2.3; see Peligrad *et al.* (2014, Proposition 5.1, supplementary material).

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

Lemma A.2. *Assume that the conditions in Theorem A.2 are satisfied. Fix $\varepsilon > 0$. Define*

$$X_{nj}^{(\varepsilon x\sigma_n)} = X_{nj}\,\mathbf{1}(X_{nj} \le \varepsilon x\sigma_n) \quad \text{and} \quad S_n^{(\varepsilon x\sigma_n)} = \sum_{j=1}^{k_n} X_{nj}^{(\varepsilon x\sigma_n)}.$$

If $x^2 \le c\ln(L_{np}^{-1})$ with $c < 1/\varepsilon$ then, as $n \to \infty$, we have

$$\mathbb{P}(S_n^{(\varepsilon x\sigma_n)} \ge x\sigma_n) = (1 - \Phi(x))(1 + o(1)).$$

In the following lemma we list some properties of the slowly varying function. Their proofs can be found in Seneta (1976) or Bingham *et al.* (1987).

Lemma A.3. *A slowly varying function $l(x)$ defined on $[A, \infty)$ has the following properties:*

(i) *for $A < c < C < \infty$, $\lim_{x\to\infty} l(tx)/l(x) = 1$ uniformly in $c \le t \le C$;*

(ii) *for any $\theta > -1$, $\int_A^x y^\theta l(y)\,\mathrm{d}y \sim x^{\theta+1}l(x)/(\theta+1)$ as $x \to \infty$;*

(iii) *for any $\theta < -1$, $\int_x^\infty y^\theta l(y)\,\mathrm{d}y \sim x^{\theta+1}l(x)/(-\theta-1)$ as $x \to \infty$;*

(iv) *for any $\eta > 0$, $\sup_{t\ge x}(t^\eta l(t)) \sim x^\eta l(x)$ as $x \to \infty$. Moreover, $\sup_{t\ge x}(t^\eta l(t)) = x^\eta \bar{l}(x)$, where $\bar{l}(x)$ is slowly varying and $\bar{l}(x) \sim l(x)$.*

Acknowledgements

The authors thank the anonymous referee and the associate editor for their careful reading of the manuscript and for their insightful comments, which have helped to improve the quality of this paper. The research of Yimin Xiao is partially supported by the National Science Foundation (grant numbers DMS-1612885 and DMS-1607089).

References

Banys, P., Davydov, Y. and Paulauskas, V. (2010). Remarks on the SLLN for linear random fields. *Statist. Prob. Lett.* **80,** 489–496.

Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation*. Cambridge University Press.

Chen, P. Y. and Wang, D. C. (2008). Convergence rates for probabilities of moderate deviations for moving average processes. *Acta Math. Sinica (English Ser.)* **24,** 611–622.

Davis, J. A. (1968). Convergence rates for the law of the iterated logarithm. *Ann. Math. Statist.* **39,** 1479–1485.

Davis, R. A. and Hsing, T. (1995). Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Prob.* **23,** 879–917.

De la Peña, V. H. and Giné, E. (1999). *Decoupling: From Dependence to Independence*. Springer, New York.

Djellout, H. and Guillin, A. (2001). Large and moderate deviations for moving average processes. *Ann. Fac. Sci. Toulouse Math. (6)* **10,** 23–31.

Djellout, H., Guillin, A. and Wu, L. (2006). Moderate deviations of empirical periodogram and non-linear functionals of moving average processes. *Ann. Inst. H. Poincaré Prob. Statist.* **42,** 393–416.

El Machkouri, M. (2007). Nonparametric regression estimation for random fields in a fixed-design. *Statist. Inference Stoch. Process.* **10,** 29–47.

El Machkouri, M. (2014). Kernel density estimation for stationary random fields. *ALEA Latin Amer. J. Prob. Math. Statist.* **11,** 259–279.

El Machkouri, M. and Stoica, R. (2010). Asymptotic normality of kernel estimates in a regression model for random fields. *J. Nonparametric Statist.* **22,** 955–971.

Frolov, A. N. (2005). On probabilities of moderate deviations of sums of independent random variables. *J. Math. Sci. (N.Y.)* **127,** 1787–1796.

Ghosh, S. and Samorodnitsky, G. (2009). The effect of memory on functional large deviations of infinite moving average processes. *Stoch. Process. Appl.* **119,** 534–561

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press

Gu, W. and Tran, L. T. (2009). Fixed design regression for negatively associated random fields. *J. Nonparametric Statist.* **21,** 345–363.
Gut, A. (1980). Convergence rates for probabilities of moderate deviations for sums of random variables with multidimensional indices. *Ann. Prob.* **8,** 298–313.
Hallin, M., Lu, Z. and Tran, L. T. (2004a). Kernel density estimation for spatial processes: the L_1 theory. *J. Multivariate Anal.* **88,** 61–75.
Hallin, M., Lu, Z. and Tran, L. T. (2004b). Local linear spatial regression. *Ann. Statist.* **32,** 2469–2500.
Jiang, T. F., Rao, M. B., and Wang, X. C. (1995). Large deviations for moving average processes. *Stoch. Process. Appl.* **59,** 309–320.
Klesov, O. (2014). *Limit Theorems for Multi-Indexed Sums of Random Variables* (Prob. Theory Stoch. Modelling **71**). Springer, Heidelberg.
Li, D. L. (1991). Convergence rates of law of iterated logarithm for B-valued random variables. *Sci. China Ser. A* **34,** 395–404.
Li, D. and Rosalsky, A. (2007). A supplement to the Davis–Gut law. *J. Math. Anal. Appl.* **330,** 1488–1493.
Li, L., Liu, J. and Xiao, Y. (2009). Wavelet regression with long memory infinite moving average errors. *J. Appl. Prob. Statist.* **4,** 183–211.
Mallik, A. and Woodroofe, M. (2011). A central limit theorem for linear random fields. *Statist. Prob. Lett.* **81,** 1623–1626.
Marinucci, D. and Poghosyan, S. (2001). Asymptotics for linear random fields. *Statist. Prob. Lett.* **51,** 131–141.
Mikosch, T. and Samorodnitsky, G. (2000). The supremum of a negative drift random walk with dependent heavy-tailed steps. *Ann. Appl. Prob.* **10,** 1025–1064.
Mikosch, T. and Wintenberger, O. (2013). Precise large deviations for dependent regularly varying sequences. *Prob. Theory Relat. Fields* **156,** 851–887.
Nagaev, A. V. (1969a). Integral limit theorems taking large deviations into account when Cramér's condition does not hold. I. *Theory Prob. Appl.* **14,** 51–64.
Nagaev, A. V. (1969b). Integral limit theorems taking large deviations into account when Cramér's condition does not hold. II. *Theory Prob. Appl.* **14,** 193–208.
Nagaev, A. V. (1969c). Limit theorems for large deviations when Cramér's condition is violated. *Izv. Akad. Nauk Uzbek. SSR Ser. Fiz.-Mat. Nauk* **6,** 17–22 (in Russian).
Nagaev, S. V. (1979). Large deviations of sums of independent random variables. *Ann. Prob.* **7,** 745–789.
Paulauskas, V. (2010). On Beveridge-Nelson decomposition and limit theorems for linear random fields. *J. Multivariate Anal.* **101,** 621–639.
Peligrad, M., Sang, H., Zhong, Y. and Wu, W. B. (2014). Exact moderate and large deviations for linear processes. *Statistica Sinica* **24,** 957–969.
Saulis, L. and Statulevičius, V. (2000). Limit theorems on large deviations. In *Limit Theorems of Probability Theory*, Springer, Berlin, pp. 185–266.
Seneta, E. (1976). *Regularly Varying Functions* (Lecture Notes Math. **508**). Springer, Berlin.
Surgailis, D. (1982). Zones of attraction of self-similar multiple integrals. *Lithuanian Math. J.* **22,** 327–340.
Tran, L. T. (1990). Kernel density estimation on random fields. *J. Multivariate Anal.* **34,** 37–53.
Wang, Y. and Woodroofe, M. (2014). On the asymptotic normality of kernel density estimators for causal linear random fields. *J. Multivariate Anal.* **123,** 201–213.
Wu, W. B. and Min, W. (2005). On linear processes with dependent innovations. *Stoch. Process. Appl.* **115,** 939–958.
Wu, W. B. and Zhao, Z. (2008). Moderate deviations for stationary processes. *Statistica Sinica* **18,** 769–782.

https://doi.org/10.1017/jpr.2018.28 Published online by Cambridge University Press