

# LARGE DEVIATIONS FOR POINT PROCESSES BASED ON STATIONARY SEQUENCES WITH HEAVY TAILS

HENRIK HULT,\* *KTH*

GENNADY SAMORODNITSKY,\*\* *Cornell University*

## Abstract

In this paper we propose a framework that facilitates the study of large deviations for point processes based on stationary sequences with regularly varying tails. This framework allows us to keep track both of the magnitude of the extreme values of a process and the order in which these extreme values appear. Particular emphasis is put on (infinite) linear processes with random coefficients. The proposed framework provides a fairly complete description of the joint asymptotic behavior of the large values of the stationary sequence. We apply the general result on large deviations for point processes to derive the asymptotic decay of certain probabilities related to partial sum processes as well as ruin probabilities.

*Keywords:* Stationary sequence; regular variation; large deviations; point process

2000 Mathematics Subject Classification: Primary 60F10; 60G10; 60G55; 60B12

## 1. Introduction

In some applications of stochastic modeling, including network traffic and finance, time series are encountered where the marginal distributions are heavy tailed and clustering of extreme values is observed. More precisely, the marginal distributions have a power-like decay and large values tend to occur at nearby points in time, forming clusters. When studying the probability of rare events, it is usually important not only to determine the size and frequency of clusters of extreme values but also to capture the internal structure of the clusters. Unfortunately, in many ‘standard’ limit theorems dealing with heavy-tailed processes the fine structure of a cluster is lost in the limit, including the ordering of the points in a cluster. This point is discussed in some detail in Section 3. To overcome this problem, we propose a new framework for investigating large deviations for stochastic processes with heavy tails. Specifically, large deviations are studied at the level of point processes associated to the underlying stochastic process. In this way it is possible to preserve the fine structure of the clusters of large values for a fairly general class of multivariate time series.

---

Received 15 June 2009; revision received 4 December 2009.

\* Postal address: Department of Mathematics, KTH, SE-100 44 Stockholm, Sweden. Email address: hult@kth.se  
Research partially supported by the Swedish Research Council.

\*\* Postal address: School of Operations Research and Industrial Engineering, Cornell University, 220 Rhodes Hall, Ithaca, NY 14853, USA. Email address: gennady@orie.cornell.edu  
Research partially supported by NSA grant H98230-06-1-0069 and ARO grant W911NF-07-1-0078 at Cornell University.

The processes studied here constitute the class of random coefficient linear processes. It consists of  $d$ -dimensional time series  $(X_k)_{k \in \mathbb{Z}}$  with the stochastic representation

$$X_k = \sum_{j \in \mathbb{Z}} A_{k,j} Z_{k-j}, \quad (1.1)$$

where the ‘noise’ sequence  $(Z_j)_{j \in \mathbb{Z}}$  consists of independent and identically distributed (i.i.d.) random vectors with values in  $\mathbb{R}^p$ . A generic element of this sequence is denoted by  $Z$ . Each  $A_{k,j}$  is a random  $d \times p$  matrix. It is assumed that the sequence  $(\mathbb{A}_k)_{k \in \mathbb{Z}}$  is stationary and that each  $\mathbb{A}_k$  is itself a sequence of matrices,  $\mathbb{A}_k = (A_{k,j})_{j \in \mathbb{Z}}$ . It is assumed that the sequence  $(\mathbb{A}_k)_{k \in \mathbb{Z}}$  is independent of the sequence  $(Z_k)_{k \in \mathbb{Z}}$ .

The probability of large values of the process  $(X_k)$  depends, of course, on the distributional assumptions on  $Z$  and  $A_{k,j}$ . In this paper we consider a heavy-tailed case: we assume that the distribution of  $Z$  is regularly varying. We also impose certain moment conditions on the random matrices  $A_{k,j}$  (see Section 2).

Probability distributions with regularly varying tails have become important building blocks in a wide variety of stochastic models. Evidence for power-tail distributions is well documented in a large number of applications, including computer networks, telecommunications, finance, insurance, hydrology, atmospheric sciences, geology, ecology, etc. (see, e.g. Embrechts *et al.* (1997), Adlet *et al.* (1998), Rachev (2003), and Tsonis and Elsner (2007)).

For the multi-dimensional version of (1.1), we use the following notion of multivariate regular variation. A  $d$ -dimensional random vector  $Z$  has a regularly varying distribution if there exists a nonnull Radon measure  $\mu$  on  $\mathbb{R}^d \setminus \{0\}$  such that

$$\frac{\mathbb{P}(u^{-1}Z \in \cdot)}{\mathbb{P}(|Z| > u)} \rightarrow \mu(\cdot) \quad \text{in } \mathbb{M}_0(\mathbb{R}^d). \quad (1.2)$$

Here  $\mathbb{M}_0(\mathbb{R}^d)$  denotes the space of Radon measures on  $\mathbb{R}^d$  whose restriction to  $\{|x| \geq r\}$  is finite for each  $r > 0$ ,  $|\cdot|$  denotes the Euclidean norm, and convergence

$$m_n \rightarrow m \quad \text{in } \mathbb{M}_0(\mathbb{R}^d)$$

is defined as the convergence  $m_n(f) \rightarrow m(f)$  for each bounded continuous function  $f$  vanishing on some neighborhood of the origin. See Hult and Lindskog (2006) for more details on the space  $\mathbb{M}_0(\mathbb{R}^d)$ .

The limiting measure  $\mu$  necessarily obeys a homogeneity property: there is an  $\alpha > 0$  such that  $\mu(uB) = u^{-\alpha} \mu(B)$  for all Borel sets  $B \subset \mathbb{R}^d \setminus \{0\}$ . This follows from standard regular variation arguments (see, e.g. Hult and Lindskog (2006, Theorem 3.1)). The notation  $Z \in \text{RV}(\mu, \alpha)$  will be used for a random vector satisfying (1.2). See Basrak (2000), Resnick (1987), Resnick (2006), and Hult and Lindskog (2006) for more on multivariate regular variation.

The class of stochastic models with the representation (1.1) is quite flexible and contains a wide range of useful time series. Here are some examples.

**Example 1.1.** (*Linear process.*) Let  $(A_j)$  be a sequence of deterministic real-valued  $d \times p$  matrices. Then, assuming convergence,  $X_k = \sum_{j \in \mathbb{Z}} A_j Z_{k-j}$  is a linear process. The representation (1.1) is trivial, and  $(X_k)$  is stationary because  $(Z_k)$  is a sequence of i.i.d.  $\mathbb{R}^d$ -valued random vectors.

**Example 1.2.** (*SRE.*) An important particular case of the random coefficient linear process is the stationary solution of a stochastic recurrence equation (SRE).

Assume that  $p = d$ , and let  $(Y_k, Z_k)_{k \in \mathbb{Z}}$  be a sequence of i.i.d. pairs of  $d \times d$  matrices and  $d$ -dimensional random vectors. Set

$$\Pi_{n,m} = \begin{cases} Y_n \cdots Y_m, & n \leq m, \\ I_d, & n > m, \end{cases}$$

where  $I_d$  is the  $d \times d$  identity matrix. Under certain assumptions ensuring the existence of a stationary solution of the SRE

$$X_k = Y_k X_{k-1} + Z_k, \quad k \in \mathbb{Z}, \tag{1.3}$$

this stationary solution can be represented by a random coefficient linear process with  $A_{k,j} = \Pi_{k-j+1,k}$ ,  $j \geq 0$ , and  $A_{k,j} = 0$ ,  $j < 0$ ; see, e.g. Kesten (1973). Then the marginal distribution of the stationary solution to the SRE is of the form (1.1).

**Example 1.3. (Stochastic volatility.)** Let  $(X_k)$  be the solution of the SRE in Example 1.2, where we assume that  $X_k \in (0, \infty)^d$  almost surely (a.s.). Let  $(V_k)$  be a sequence of i.i.d. random diagonal matrices independent of  $(X_k)$ . Then  $U_k = V_k X_k$  has the representation

$$U_k = \sum_{j \in \mathbb{Z}} \tilde{A}_{k,j} Z_{k-j},$$

where  $\tilde{A}_{k,j} = V_k A_{k,j}$  and  $A_{k,j}$  is as in Example 1.2. The sequence  $U_k$  can be interpreted as a stochastic volatility model where  $X_k$  is the volatility; see, e.g. Davis and Mikosch (2008).

## 2. Convergence and tail behavior

Consider a time series  $(X_k)$  with stochastic representation (1.1). Throughout this paper, it is assumed that

$$Z \in \text{RV}(\mu, \alpha) \quad \text{and, if } \alpha > 1 \text{ also, } \quad \mathbb{E} Z = 0. \tag{2.1}$$

To begin the study of extreme values for the time series (1.1), a first requirement is to establish conditions under which the infinite series converge a.s. and to determine the tail behavior of the distribution of  $X_k$ . Hult and Samorodnitsky (2008) recently obtained results on the tail behavior under conditions that include a ‘predictability’ assumption on the matrices  $(A_{k,j})$ . Here we summarize the results and remind the reader that in this paper we assume that  $(A_{k,j})$  and  $(Z_j)$  are independent. Theorem 2.1, below, describes the marginal tails; for simplicity, we drop the time subscript  $k$  from both  $X_k$  and  $A_{k,j}$ .

Throughout the paper,  $\|A\|$  denotes the operator norm of a matrix  $A$ . The summation index will be omitted when it is clear what it is.

**Theorem 2.1.** *Suppose that (2.1) holds and that there is  $\varepsilon \in (0, \alpha)$  such that*

$$\sum \mathbb{E} \|A_j\|^{\alpha-\varepsilon} < \infty \quad \text{and} \quad \sum \mathbb{E} \|A_j\|^{\alpha+\varepsilon} < \infty, \quad \alpha \in (0, 1) \cup (1, 2), \tag{2.2}$$

$$\mathbb{E} \left( \sum \|A_j\|^{\alpha-\varepsilon} \right)^{(\alpha+\varepsilon)/(\alpha-\varepsilon)} < \infty, \quad \alpha \in \{1, 2\}, \tag{2.3}$$

$$\mathbb{E} \left( \sum \|A_j\|^2 \right)^{(\alpha+\varepsilon)/2} < \infty, \quad \alpha \in (2, \infty). \tag{2.4}$$

*Then the series (1.1) converges a.s. and*

$$\frac{\mathbb{P}(u^{-1} X \in \cdot)}{\mathbb{P}(|Z| > u)} \rightarrow \mathbb{E} \left( \sum \mu \circ A_j^{-1}(\cdot) \right) \quad \text{in } \mathbb{M}_0(\mathbb{R}^d). \tag{2.5}$$

The right-hand side of (2.5) is interpreted as

$$E\left(\sum \mu \circ A_j^{-1}(B)\right) = E\left(\sum \mu\{z: A_j z \in B\}\right)$$

for any Borel set  $B \subset \mathbb{R}^d$ .

When both  $Z$  and  $A_{k,j}$  are univariate ( $d = p = 1$ ), the limiting measure  $\mu$  of  $Z$  has the representation

$$\mu(dz) = (w\alpha z^{-\alpha-1} \mathbf{1}\{z > 0\} + (1-w)\alpha(-z)^{-\alpha-1} \mathbf{1}\{z < 0\}) dz \tag{2.6}$$

for some  $w \in [0, 1]$ . Then (2.5) becomes

$$\frac{P(X > ux)}{P(|Z| > u)} \rightarrow \sum E(|A_j|^\alpha (w \mathbf{1}\{A_j > 0\} + (1-w) \mathbf{1}\{A_j < 0\})) x^{-\alpha}$$

for each  $x > 0$ , with a similar expression for the negative tail; see Hult and Samorodnitsky (2008, Remark 3.3).

**Example 2.1.** (*Linear process.*) If  $(X_k)$  is a linear process ( $A_{k,j} = A_j$  deterministic) and  $d = p = 1$ , then, for  $u \rightarrow \infty$ ,

$$\frac{P(X > ux)}{P(|Z| > u)} \rightarrow \sum (|A_j|^\alpha (w \mathbf{1}\{A_j > 0\} + (1-w) \mathbf{1}\{A_j < 0\})) x^{-\alpha}.$$

**Example 2.2.** (*SRE.*) Suppose that  $(X_k)$  is the solution to the SRE in Example 1.2 with  $Y$  satisfying  $E \|Y\|^{\alpha+\varepsilon} < 1$  for some  $\varepsilon > 0$ . Then, in the case  $d = p = 1$ ,

$$\frac{P(X > ux)}{P(|Z| > u)} \rightarrow \frac{w(1 - E(Y^+)^{\alpha}) + (1-w)E(Y^-)^{\alpha}}{(1 - E(Y^+)^{\alpha})^2 + (E(Y^-)^{\alpha})^2} x^{-\alpha}; \tag{2.7}$$

see Hult and Samorodnitsky (2008, Example 3.3). Here, and throughout,  $x^+ = \max\{x, 0\}$  denotes the positive part of  $x$ , and  $x^- = \max\{-x, 0\}$  denotes its negative part. In particular, if  $Y$  is nonnegative then  $w = 1$ ,  $E Y^- = 0$ , and (2.7) reduces to

$$\frac{P(X > ux)}{P(|Z| > u)} \rightarrow (1 - E Y^{\alpha})^{-1} x^{-\alpha} \quad (u \rightarrow \infty).$$

**Example 2.3.** (*Stochastic volatility.*) Let  $(X_k)$  be as in Example 2.2, where  $d = p = 1$  and  $Y$  and  $Z$  are nonnegative. Let  $(V_k)$  be a sequence of i.i.d. random variables, independent of  $(X_k)$ . Suppose that  $E V^{\alpha+\varepsilon} < \infty$  for some  $\varepsilon > 0$ . Then  $U_k = V_k X_k$  satisfies

$$\frac{P(U > ux)}{P(Z > u)} \rightarrow \frac{E V^{\alpha}}{1 - E Y^{\alpha}} x^{-\alpha} \quad (u \rightarrow \infty).$$

**Remark 2.1.** The following two observations will be useful for later reference. It follows from Remark 4.1 of Hult and Samorodnitsky (2008) that, for any increasing truncation  $n(x) \uparrow \infty$ ,

$$\lim_{x \rightarrow \infty} \frac{P(|\sum_{|j|>n(x)} A_j Z_j| > x)}{P(|Z| > x)} = 0. \tag{2.8}$$

Furthermore, only values of  $Z_j$  comparable to the level  $x$  matter in the sense that

$$\lim_{\tau \rightarrow 0} \limsup_{x \rightarrow \infty} \frac{P(|\sum_{j \in \mathbb{Z}} A_j Z_j \mathbf{1}\{|Z_j| \leq \tau x\}| > x)}{P(|Z| > x)} = 0. \tag{2.9}$$

### 3. Why are large deviations of point processes needed?

In this section we discuss, somewhat informally, the joint asymptotic behavior of large values of the sequence  $(X_k)$  in (1.1). The goal is to set up the necessary background and intuition for the general result in Section 4. We consider two special cases: sequences of i.i.d. random variables and moving average processes. There is no clustering of the extremes in the former case, but such clustering does occur in the latter case. Our goal is to devise a limiting procedure that preserves the relevant fine structure of clusters of the extremes. The theme of clustering is important in other areas of probability as well; see Aldous (1989).

**Example 3.1.** (*Independent and identically distributed random variables.*) Consider a sequence  $(Z_k)$  of i.i.d. real-valued random variables with  $Z \in \text{RV}(\alpha, \mu)$  and  $\mu$  as in (2.6), and, as in (2.1), we assume that  $E Z_k = 0$  when  $\alpha > 1$ . It is well known (see, e.g. Resnick (1987)) that, for each  $n \geq 1$ , the vector  $(Z_1, \dots, Z_n)$  is regularly varying with limit measure  $\mu^{(n)}$  concentrated on the coordinate axes:

$$\mu^{(n)}(dz_1, \dots, dz_n) = \sum_{i=1}^n \mu(dz_i) \prod_{j \neq i} \delta_0(dz_j),$$

where  $\delta_x$  is a unit mass at  $x$ . The interpretation is that, asymptotically, only one of the variables  $Z_1, \dots, Z_n$  is of large absolute value and each variable is equally likely to be large.

The same intuition holds true when considering variables  $Z_1, \dots, Z_n$  in a time window of length  $n$  and letting  $n \rightarrow \infty$ , if the threshold increases with  $n$  at an appropriate rate. Let  $\gamma_n$  be a sequence with  $\gamma_n \rightarrow \infty$  and such that  $n P(|Z| > \gamma_n) \rightarrow 0$  as  $n \rightarrow \infty$ . Then, the probability that two different  $Z$ s among the variables  $Z_1, \dots, Z_n$  are of size of the order  $\gamma_n$  is small compared to the probability that just one of the  $Z_i$ s is of size of the order  $\gamma_n$ . Indeed, for any  $\varepsilon > 0$ ,

$$\begin{aligned} & \frac{P(\text{there exist } 1 \leq i < j \leq n \text{ such that } |Z_i| > \gamma_n \varepsilon \text{ and } |Z_j| > \gamma_n \varepsilon)}{P(|Z_i| > \gamma_n \text{ for some } 1 \leq i \leq n)} \\ & \sim \frac{(n(n-1)/2) P(|Z| > \gamma_n \varepsilon)^2}{n P(|Z| > \gamma_n)} \\ & \rightarrow 0. \end{aligned}$$

Here  $a_n \sim b_n$  is shorthand for  $\lim_{n \rightarrow \infty} a_n/b_n = 1$ .

A convenient description of the large values for the sequence  $Z_1, Z_2, \dots$  can be obtained by considering the convergence of the point measures

$$N_n = \sum_{k=1}^n \delta_{(k/n, \gamma_n^{-1} Z_k)}, \quad n = 1, 2, \dots,$$

on the state space  $[0, 1] \times (\mathbb{R}^d \setminus \{0\})$ . The assumption that  $n P(|Z| > \gamma_n) \rightarrow 0$  as  $n \rightarrow \infty$  implies that  $\gamma_n \rightarrow \infty$  too fast for a nontrivial weak convergence of  $N_n$  (described, for example, by Proposition 3.21 of Resnick (1987)). When  $\gamma_n$  grows so fast, the second coordinates of all points of the point measure  $N_n$  will tend to 0 with probability 1. Since points with the zero second coordinate are defined not to be in the state space on which the point measures live (see, once again, Resnick (1987)), it turns out that the point measure  $N_n$  converges a.s. to the null measure, denoted by  $\emptyset$  (we use the notation introduced in Daley and Vere-Jones (2003)). Intuitively, this is exactly the situation where large deviations in the space of point measures

might help: the hope is to find a sequence  $r_n \rightarrow \infty$  such that  $r_n P(N_n \in \cdot)$  converges to some limiting measure  $m$  on the space of point measures.

The above discussion makes it reasonable to expect that this limiting measure,  $m$ , is concentrated on point measures with one point, corresponding, for each  $n = 1, 2, \dots$ , to a single large value of  $Z_{k^*}$ ,  $k^* = 1, \dots, n$  (at the scale  $\gamma_n = n$ ). In fact, the limiting measure is expected to be

$$m(B) = (\text{Leb} \times \mu)\{(t, z) : \delta_{(t,z)} \in B\}, \quad B \text{ a measurable set of measures.}$$

Here and in the sequel,  $\text{Leb}$  denotes the appropriate Lebesgue measure (on the unit interval, in this case). The value of the coordinate  $t$  is interpreted as the time  $k^*$  of the large  $Z_{k^*}$  value rescaled within the set  $\{1, \dots, n\}$ , and since all  $Z_k$ s have equal probability of being large,  $t$  is ‘uniformly distributed’ on  $[0, 1]$ . The corresponding value  $z$  is governed by the limiting measure  $\mu$  which describes the large values of the  $Z$  variables. The suggested convergence is established rigorously (in a significantly more general setting) in Theorem 4.1, below.

It is possible to look at this convergence as the partial sum convergence of the underlying sequence  $(\delta_{(k/n, \gamma_n^{-1} Z_k)})$  in the space of point measures. This is similar to Sanov’s theorem in the light-tailed case (see, e.g. Dembo and Zeitouni (1998, Section 6.2)).

**Example 3.2.** (*A finite moving average.*) Suppose that, in (1.1),  $p = d = 1$  and  $A_{k,j} = A_j$  are deterministic coefficients with  $A_j = 0$  if  $j < 0$  or  $j > q$ . Then  $(X_k)$  is a sequence defined by

$$X_k = A_0 Z_k + A_1 Z_{k-1} + \dots + A_q Z_{k-q}.$$

Consider a time window of length  $n$  where, for now,  $n$  is fixed. That is, we consider the vector  $(X_1, \dots, X_n)$ . Then,  $(X_1, \dots, X_n)^\top = A'(Z_{1-q}, \dots, Z_n)^\top$ , where  $A'$  is the  $n \times (n + 1 + q)$  matrix

$$A' = \begin{pmatrix} A_q & A_{q-1} & \dots & A_0 & 0 & \dots & \dots & 0 \\ 0 & A_q & A_{q-1} & \dots & A_0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & A_q & A_{q-1} & \dots & \dots & A_0 \end{pmatrix}.$$

Since  $(Z_k)$  is an i.i.d. sequence, the vector  $(Z_{1-q}, \dots, Z_n)^\top$  is regularly varying with limit measure concentrated on the coordinate axes, as in Example 3.1. That is, asymptotically, only one variable among  $Z_{1-q}, \dots, Z_n$  will be large on the large deviations scale, and they all have equal probability of being large. Suppose that  $Z_{k^*}$  is large for some  $1 - q \leq k^* \leq n$ . Then, since all the other  $Z_k$ s are small in comparison to  $Z_{k^*}$ , we expect that  $X_k$  is small if  $k < k^*$  or  $k > k^* + q$ , while, for  $k^* \leq k \leq k^* + q$ , we have

$$X_k \approx A_{k-k^*} Z_{k^*}.$$

If we can find a sequence  $r_n \rightarrow \infty$  such that the sequence of measures  $(r_n P(N_n \in \cdot))$  converges to a nondegenerate limit, where

$$N_n = \sum_{k=1}^n \delta_{(k/n, \gamma_n^{-1} X_k)}, \quad n = 1, 2, \dots,$$

is defined on the state space  $[0, 1] \times (\mathbb{R} \setminus \{0\})$ , we would expect the limiting measure to be concentrated on point measures with  $q + 1$  points of the form  $(t, x_i)$ , with the same time coordinate  $t$  and space coordinates of the form  $x_i = A_i z$  for some  $z$ . In other words, we expect the limiting measure to be

$$m(B) = (\text{Leb} \times \mu) \left\{ (t, z) : \sum_{i=0}^q \delta_{(t, A_i z)} \in B \right\}, \quad B \text{ a measurable set of measures.}$$

The clustering of extreme values is captured in the limiting measure as there are  $q + 1$  points corresponding to large values of the  $X_k$ s. However, in the limit, all these points have the same time coordinate  $t$ , which means that the limiting measure does not keep track of the order in which the large values arrived. That is, the complete internal structure of the cluster of extreme values is not captured. The order at which the large values arrive is, however, of crucial importance when studying, for instance, the ruin probabilities, or the long strange segments corresponding to the process  $(X_k)$ ; see, e.g. Asmussen (2000), Dembo and Zeitouni (1998), Mikosch and Samorodnitsky (2000), and Hult *et al.* (2005). Therefore, information is lost in the limit.

The strategy we adopt in order to retain information on the internal cluster structure of extreme values is to consider point measures similar to the measures  $N_n$  above, but enlarging the dimension of the state space so that each point of the point measure describes more than one value of the process  $(X_k)$ . It is intuitive that, for a finite moving average of this example, it is enough to keep track of  $q + 1$  consecutive observations of the stationary process, and this tells us how large the state space of the point measures should be. Specifically, we consider the point measures

$$\sum_{k=1}^n \delta_{(k/n, \gamma_n^{-1}(X_k, X_{k-1}, \dots, X_{k-q}))}, \quad n = 1, 2, \dots$$

The above discussion makes plausible the assertion that, for such point processes, the limiting measure in a large deviations procedure is concentrated on point measures with  $2q + 1$  points of the form

$$(t, (A_0 z, 0, \dots, 0)), (t, (A_1 z, A_0 z, 0, \dots, 0)), \dots, (t, (0, \dots, 0, A_q z)).$$

Note that the information about the order in which the extreme values arrived can be obtained because the space coordinates are simply shifts of each other.

In general, all the information on the extreme values of the process will be preserved only if we keep track of an infinite (or increasing with  $n$ ) number of observations of the process  $(X_k)$ . This is possible to do, but we have chosen not to pursue this last possibility because it significantly complicates the technical details of the construction of the point measures and working with these measures. Instead, we have chosen to construct point measures based on finitely many consecutive observations of the stationary process, as if it were a finite moving average. In the applications we consider, this turns out to be sufficient when coupled with a truncation argument.

### 4. Large deviations for point processes: the main result

We start by specifying the precise assumptions on the normalizing sequence  $(\gamma_n)_{n \geq 1}$  that we need to obtain a large deviation scaling. Assume that, as  $n \rightarrow \infty$ ,

$$\frac{Z_1 + \dots + Z_n}{\gamma_n} \rightarrow 0 \quad \text{in probability} \tag{4.1}$$

and  $\frac{\gamma_n}{\sqrt{n^{1+\varepsilon}}} \rightarrow \infty$  for some  $\varepsilon > 0$  if  $\alpha = 2$        $\frac{\gamma_n}{\sqrt{n \log n}} \rightarrow \infty$  if  $\alpha > 2$ .

Note that these conditions are exactly the same as those that were used in Theorem 2.1 of Hult *et al.* (2005) to obtain a functional level large deviation result for the partial sums of i.i.d. random vectors. If we set  $r_n = 1/(n P(|Z| > \gamma_n))$  then  $r_n \rightarrow \infty$  as  $n \rightarrow \infty$  and it turns out that normalizing the probability measures of the point processes by  $(r_n)$  is the correct normalization to obtain a large deviation result.

For  $q \geq 0$ , define a point measure  $N_n^q$  on the space  $\mathbb{E}^q = [0, 1] \times (\mathbb{R}^{d(q+1)} \setminus \{0\})$  by

$$N_n^q = \sum_{k=1}^n \delta_{(k/n, \gamma_n^{-1} X_k, \gamma_n^{-1} X_{k-1}, \dots, \gamma_n^{-1} X_{k-q})}. \tag{4.2}$$

We will show that the sequence of measures on the space of point measures,

$$m_n^q(\cdot) = r_n P(N_n^q \in \cdot), \quad n \geq 1,$$

converges in the appropriate sense and compute the limiting measure, called  $m^q$ , for any  $q \geq 0$ . The limiting measure gives us a partial description of the extremal behavior of the sequence  $(X_k)$ . This description becomes more and more detailed as the number  $q$  is taken larger and larger.

A technical framework suitable for studying this problem is provided in Appendix A, and we use the notation introduced there. Let  $\mathbb{N}_p^q = \mathbb{N}_p(\mathbb{E}^q)$  be the space of point measures on  $\mathbb{E}^q$  equipped with the vague topology. The convergence  $m_n^q \rightarrow m^q$  takes place in the space  $\mathbb{M}_0(\mathbb{N}_p^q)$ , the space of Radon measures on  $\mathbb{N}_p^q$  that are finite on sets of the form  $\{\xi : d(\xi, \emptyset) > r\}$  for each  $r > 0$  (see Appendix A). Here  $\emptyset$  denotes the null measure and  $d(\cdot, \cdot)$  denotes the metric on  $\mathbb{N}_p^q$  given by (A.1). With this metric,  $(\mathbb{N}_p^q, d)$  is a complete separable metric space.

For a sequence  $\mathbb{A}$  of  $d \times p$  matrices  $(A_{k,j})_{j,k \in \mathbb{Z}}$  and  $(t, z) \in [0, 1] \times (\mathbb{R}^p \setminus \{0\})$ , we write

$$T_{\mathbb{A},q}(t, z) = \sum_{j \in \mathbb{Z}} \delta_{(t, A_{j,j}z, A_{j-1,j-1}z, \dots, A_{j-q,j-q}z)}.$$

Under certain conditions on the matrices in  $A_{k,j}$ ,  $T_{\mathbb{A},q}$  is a map from  $[0, 1] \times (\mathbb{R}^p \setminus \{0\})$  into the space  $\mathbb{N}_p^q$ .

We can now state the main result of this paper.

**Theorem 4.1.** *Suppose that (2.1)–(2.4) and (4.1) hold. Then, for any  $q \geq 0$ , the stationary process  $(X_k)_{k \in \mathbb{Z}}$  in (1.1) satisfies*

$$m_n^q(\cdot) = r_n P(N_n^q \in \cdot) \rightarrow E((\text{Leb} \times \mu) \circ T_{\mathbb{A},q}^{-1}(\cdot)) =: m^q(\cdot) \quad \text{in } \mathbb{M}_0(\mathbb{N}_p^q). \tag{4.3}$$

*In particular, with probability 1,  $T_{\mathbb{A},q}$  is a map from  $[0, 1] \times (\mathbb{R}^p \setminus \{0\})$  into the space  $\mathbb{N}_p^q$ .*

**Remark 4.1.** For any  $a > 0$ , the measure  $m^q$  on  $\mathbb{N}_p^q$  defined in (4.3) satisfies

$$\begin{aligned} & m^q\{\xi: \xi([0, 1] \times \{(x_0, \dots, x_q): |x_i| = a, \text{ some } i \in \{0, \dots, q\}\}) > 0\} \\ &= \mathbb{E}\left(\mu\left\{z: \sum_{j \in \mathbb{Z}} \delta_{(A_{j,jz}, \dots, A_{j-q,j-qz})}((x_0, \dots, x_q): |x_i| = a, \text{ some } i \in \{0, \dots, q\}) > 0\right\}\right) \\ &\leq \sum_{j \in \mathbb{Z}} \mathbb{E} \mu\{z: |A_{j,jz}| = a\} \\ &= 0, \end{aligned}$$

by the scaling property of the measure  $\mu$ . This fact is useful for establishing continuity almost everywhere with respect to the measure  $m^q$  of various mappings.

**Example 4.1.** (*Independent and identically distributed random vectors.*) For a sequence of i.i.d. random vectors, we have  $A_{k,j} = A\delta_{0j}$ , where  $A$  is a fixed  $d \times p$  matrix, and, hence, for  $q = 0$ , the limiting measure  $m^0$  is given by  $m^0(\cdot) = (\text{Leb} \times \mu) \circ T_{A,0}^{-1}(\cdot)$ , where  $T_{A,0}$  is the mapping

$$T_{A,0}(t, z) = \delta_{(t, Az)}.$$

**Example 4.2.** (*Linear process.*) For a linear process, the matrices  $A_{k,j} = A_j$ ,  $j \in \mathbb{Z}$ , are deterministic. The limiting measure  $m^q$  is given by  $m^q(\cdot) = (\text{Leb} \times \mu) \circ T_{A,q}^{-1}(\cdot)$ , where the mapping  $T_{A,q}$  is given by

$$T_{A,q}(t, z) = \sum_{j \in \mathbb{Z}} \delta_{(t, A_j z, A_{j-1} z, \dots, A_{j-q} z)}.$$

*Proof of Theorem 4.1.* By Theorem A.2 we need to prove that the measure  $m^q$  in (4.3) belongs to  $\mathbb{M}_0(\mathbb{N}_p^q)$  and that, as  $n \rightarrow \infty$ ,

$$m_n(F_{g_1, g_2, \varepsilon_1, \varepsilon_2}) \rightarrow m^q(F_{g_1, g_2, \varepsilon_1, \varepsilon_2}) \tag{4.4}$$

for all Lipschitz functions  $g_1, g_2 \in C_K^+(\mathbb{E}^q)$  and  $\varepsilon_1, \varepsilon_2 > 0$ , where the functions  $F_{g_1, g_2, \varepsilon_1, \varepsilon_2}$  are given in (A.2) in Appendix A. For the first statement, it is enough to prove that, for each  $\delta > 0$ ,

$$\mathbb{E} S_\delta =: \mathbb{E}\left(\sum_{j \in \mathbb{Z}} \mathbf{1}\{\|A_j\| > \delta\}\right) < \infty.$$

This is an easy consequence of conditions (2.2)–(2.4). For example, if  $0 < \alpha \leq 2$  then, for  $0 < \varepsilon < \alpha$ ,

$$\mathbb{E} S_\delta \leq \delta^{-(\alpha-\varepsilon)} \sum_{j \in \mathbb{Z}} \mathbb{E} \|A_j\|^{\alpha-\varepsilon} < \infty,$$

and the case  $\alpha > 2$  is similar.

We now prove (4.4). Note that

$$\begin{aligned} m_n(F_{g_1, g_2, \varepsilon_1, \varepsilon_2}) &= r_n \mathbb{E}\left(\left(1 - \exp\left\{-\left[\sum_{k=1}^n g_1\left(\frac{k}{n}, \frac{X_k}{\gamma_n}, \dots, \frac{X_{k-q}}{\gamma_n}\right) - \varepsilon_1\right]_+\right\}\right)\right. \\ &\quad \left.\times \left(1 - \exp\left\{-\left[\sum_{k=1}^n g_2\left(\frac{k}{n}, \frac{X_k}{\gamma_n}, \dots, \frac{X_{k-q}}{\gamma_n}\right) - \varepsilon_2\right]_+\right\}\right)\right). \end{aligned} \tag{4.5}$$

The first step is to truncate the infinite sum in the definition of  $X_k$ , replacing  $X_k$  by  $\sum_{|j| \leq J_n} A_{k,j} \times Z_{k-j}$ , as follows. Let  $(J_n)$  be a sequence of positive numbers such that  $J_n \rightarrow \infty$  and

$$J_n = \begin{cases} o(n) & \text{if } 0 < \alpha < 1, \\ o(\min(n, \gamma_n/l(\gamma_n))) & \text{if } \alpha = 1, \\ o(\min(n, \gamma_n)) & \text{if } \alpha > 1, \end{cases} \tag{4.6}$$

where, for  $x > 0$ ,  $l(x) = E(|Z| \mathbf{1}\{|Z| \leq x\})$ . The conditions on the asymptotic growth of  $J_n$  are used below.

By Lemma 4.1, there is a sequence  $\beta_n \downarrow 0$  such that

$$r_n P\left(\max_{1 \leq k \leq n} \frac{1}{\gamma_n} \left| \sum_{|j| > J_n} A_{k,j} Z_{k-j} \right| > \beta_n\right) \rightarrow 0$$

as  $n \rightarrow \infty$ . Therefore, the expression on the right-hand side of (4.5) is within  $o(1)$  of

$$\begin{aligned} & r_n E \left[ \left( 1 - \exp \left\{ - \left[ \sum_{k=1}^n g_1 \left( \frac{k}{n}, R_{k,n} + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k,j} Z_{k-j}, \dots, R_{k-q,n} \right. \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k-q,j} Z_{k-q-j} \right) - \varepsilon_1 \right]_+ \right\} \right) \right. \\ & \left. \times \left( 1 - \exp \left\{ - \left[ \sum_{k=1}^n g_2 \left( \frac{k}{n}, R_{k,n} + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k,j} Z_{k-j}, \dots, R_{k-q,n} \right. \right. \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k-q,j} Z_{k-q-j} \right) - \varepsilon_2 \right]_+ \right\} \right) \right] \\ & := r_n E(\Theta_n), \end{aligned}$$

where  $(R_{k,n})$  are random variables satisfying  $|R_{k,n}| \leq \beta_n$  for all  $k, n$ .

To proceed, we use the intuitive idea that only one of the  $Z$ s is likely to be large. Take  $\tau > 0$ . The above expression can be decomposed as

$$\begin{aligned} & r_n E(\Theta_n \mathbf{1}\{\text{all } |Z_{-J_n-q+1}|, \dots, |Z_{n+J_n}| \text{ are less than } \tau \gamma_n\}) \\ & + r_n E(\Theta_n \mathbf{1}\{\text{exactly one of } |Z_{-J_n-q+1}|, \dots, |Z_{n+J_n}| \text{ exceeds } \tau \gamma_n\}) \\ & + r_n E(\Theta_n \mathbf{1}\{\text{at least two of } |Z_{-J_n-q+1}|, \dots, |Z_{n+J_n}| \text{ exceed } \tau \gamma_n\}) \\ & = r_n E \left( \Theta_n \mathbf{1} \left\{ \bigcap_{t=-J_n-q+1}^{n+J_n} |Z_t| \leq \tau \gamma_n \right\} \right) \tag{4.7} \end{aligned}$$

$$+ r_n E \left( \Theta_n \mathbf{1} \left\{ \bigcup_{t=-J_n-q+1}^{n+J_n} \bigcap_{\substack{s=-J_n-q+1, \dots, n+J_n \\ s \neq t}} \{|Z_t| > \tau \gamma_n, |Z_s| \leq \tau \gamma_n\} \right\} \right) \tag{4.8}$$

$$+ r_n E \left( \Theta_n \mathbf{1} \left\{ \bigcup_{t=-J_n-q+1}^{n+J_n} \bigcup_{\substack{s=-J_n-q+1, \dots, n+J_n \\ s \neq t}} \{|Z_t| > \tau \gamma_n, |Z_s| > \tau \gamma_n\} \right\} \right). \tag{4.9}$$

We claim that the main contribution comes from (4.8) and that the contributions from the other terms vanish as  $n \rightarrow \infty$  and then  $\tau \rightarrow 0$ . Let us start with (4.7). Recall that  $g_1$  and  $g_2$  have



As  $|Z_t|$  is large and  $|Z_s|$  is small,  $s \neq t$ , we can practically ignore the contribution from the latter terms. To be precise, we claim that the above expression is asymptotically equal (written  $a_n \sim b_n$ ) to

$$\begin{aligned} & \sum_{t=-J_n-q+1}^{n+J_n} r_n \mathbb{E} \left[ \left( 1 - \exp \left\{ - \left[ \sum_{k=1}^n g_1 \left( \frac{k}{n}, R_{k,n} + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k,j} Z_{k-j} \mathbf{1}\{t = k - j\}, \dots, \right. \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. R_{k-q,n} + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k-q,j} Z_{k-q-j} \mathbf{1}\{t = k - q - j\} \right) - \varepsilon_1 \right]_+ \right\} \right) \\ & \times \left( 1 - \exp \left\{ - \left[ \sum_{k=1}^n g_2 \left( \frac{k}{n}, R_{k,n} + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k,j} Z_{k-j} \mathbf{1}\{t = k - j\}, \dots, \right. \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. R_{k-q,n} + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k-q,j} Z_{k-q-j} \mathbf{1}\{t = k - q - j\} \right) - \varepsilon_1 \right]_+ \right\} \right) \\ & \times \mathbf{1}\{|Z_t| > \tau \gamma_n\} \Big] \\ =: & \sum_{t=-J_n-q+1}^{n+J_n} r_n \mathbb{E}[\Theta'_n \mathbf{1}\{|Z_t| > \tau \gamma_n\}]. \tag{4.11} \end{aligned}$$

For now, we postpone the proof that (4.10)  $\sim$  (4.11) and proceed, instead, with analyzing (4.11). We can rewrite (4.11) as

$$\begin{aligned} & \sum_{t=-J_n-q+1}^{n+J_n} r_n \mathbb{E} \left[ \left( 1 - \exp \left\{ - \left[ \sum_{k=1}^n g_1 \left( \frac{k}{n}, R_{k,n} + \frac{1}{\gamma_n} A_{k,k-t} Z_t \mathbf{1}\{|k - t| \leq J_n\}, \dots, R_{k-q,n} \right. \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. + \frac{1}{\gamma_n} A_{k-q,k-q-t} Z_t \mathbf{1}\{|k - q - t| \leq J_n\} \right) - \varepsilon_1 \right]_+ \right\} \right) \\ & \times \left( 1 - \exp \left\{ - \left[ \sum_{k=1}^n g_2 \left( \frac{k}{n}, R_{k,n} + \frac{1}{\gamma_n} A_{k,k-t} Z_t \mathbf{1}\{|k - t| \leq J_n\}, \dots, R_{k-q,n} \right. \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. + \frac{1}{\gamma_n} A_{k-q,k-q-t} Z_t \mathbf{1}\{|k - q - t| \leq J_n\} \right) - \varepsilon_2 \right]_+ \right\} \right) \\ & \times \mathbf{1}\{|Z_t| > \tau \gamma_n\} \Big]. \end{aligned}$$

In the sequel, as the subscripts change, we shorten  $R_{k,n}$  to  $R_n$ . We do not impose any assumptions on these random variables apart from the fact that  $|R_n| \leq \beta_n$  for all  $n$ . With  $l = k - t$  we can rewrite the above expression as

$$\begin{aligned} & \sum_{t=-J_n-q+1}^{n+J_n} r_n \mathbb{E} \left[ \left( 1 - \exp \left\{ - \left[ \sum_{l=1-t}^{n-t} g_1 \left( \frac{t+l}{n}, R_n + \frac{1}{\gamma_n} A_{t+l,l} Z_t \mathbf{1}\{|l| \leq J_n\}, \dots, R_n \right. \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. + \frac{1}{\gamma_n} A_{t+l-q,l-q} Z_t \mathbf{1}\{|l - q| \leq J_n\} \right) - \varepsilon_1 \right]_+ \right\} \right) \end{aligned}$$

$$\begin{aligned} & \times \left( 1 - \exp \left\{ - \left[ \sum_{l=1-t}^{n-t} g_2 \left( \frac{t+l}{n}, R_n + \frac{1}{\gamma_n} A_{t+l,l} Z_t \mathbf{1}\{|l| \leq J_n\}, \dots, R_n \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. + \frac{1}{\gamma_n} A_{t+l-q,l-q} Z_t \mathbf{1}\{|l-q| \leq J_n\} \right) - \varepsilon_2 \right]_+ \right\} \right) \\ & \times \mathbf{1}\{|Z_t| > \tau \gamma_n\}. \end{aligned}$$

By stationarity we may replace  $A_{t+l-i,l}$  by  $A_{l-i,l}$ ,  $i = 0, \dots, q$ , and, conditioning on  $Z_t$ , the above equals

$$\begin{aligned} & \int_{|z|>\tau} \sum_{t=-J_n-q+1}^{n+J_n} r_n \mathbb{E} \left[ \left( 1 - \exp \left\{ - \left[ \sum_{l=1-t}^{n-t} g_1 \left( \frac{t+l}{n}, R_n + A_{l,l} z \mathbf{1}\{|l| \leq J_n\}, \dots, R_n \right. \right. \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. + A_{l-q,l-q} z \mathbf{1}\{|l-q| \leq J_n\} \right) - \varepsilon_1 \right]_+ \right\} \right) \\ & \qquad \qquad \qquad \times \left( 1 - \exp \left\{ - \left[ \sum_{l=1-t}^{n-t} g_2 \left( \frac{t+l}{n}, R_n + A_{l,l} z \mathbf{1}\{|l| \leq J_n\}, \dots, R_n \right. \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. + A_{l-q,l-q} z \mathbf{1}\{|l-q| \leq J_n\} \right) - \varepsilon_2 \right]_+ \right\} \right) \Bigg] \\ & \times \mathbb{P}(\gamma_n^{-1} Z_t \in dz) \\ & =: \int_{|z|>\tau} \kappa_n(z) r_n n \mathbb{P}(\gamma_n^{-1} Z \in dz) \\ & =: \int_{|z|>\tau} \kappa_n(z) \mu_n(dz) \\ & \sim \int_{|z|>\tau} \tilde{\kappa}_n(z) \mu_n(dz), \end{aligned}$$

where

$$\begin{aligned} \tilde{\kappa}_n(z) = & \sum_{t=-J_n+1}^{n+J_n} \mathbb{E} \left[ \left( 1 - \exp \left\{ - \left[ \sum_{l=1-t}^{n-t} g_1 \left( \frac{t+l}{n}, A_{l,l} z \mathbf{1}\{|l| \leq J_n\}, \dots, \right. \right. \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. A_{l-q,l-q} z \mathbf{1}\{|l-q| \leq J_n\} \right) - \varepsilon_1 \right]_+ \right\} \right) \\ & \times \left( 1 - \exp \left\{ - \left[ \sum_{l=1-t}^{n-t} g_2 \left( \frac{t+l}{n}, A_{l,l} z \mathbf{1}\{|l| \leq J_n\}, \dots, \right. \right. \right. \right. \\ & \qquad \qquad \qquad \left. \left. \left. A_{l-q,l-q} z \mathbf{1}\{|l-q| \leq J_n\} \right) - \varepsilon_2 \right]_+ \right\} \right) \Bigg] \frac{1}{n}, \end{aligned}$$

and we have used the uniform continuity of the functions  $g_1$  and  $g_2$  and the fact that  $|R_n| \leq \beta_n \downarrow 0$ . We claim that, as  $n \rightarrow \infty$ ,

$$\int_{|z|>\tau} \tilde{\kappa}_n(z) \mu_n(dz) \rightarrow \int_{|z|>\tau} \kappa(z) \mu(dz), \tag{4.12}$$

where

$$\begin{aligned} \kappa(z) = \int_0^1 \mathbb{E} & \left[ \left( 1 - \exp \left\{ - \left[ \sum_{l \in \mathbb{Z}} g_1(t, A_{l,lz}, \dots, A_{l-q,l-qz}) - \varepsilon_1 \right]_+ \right\} \right) \right. \\ & \left. \times \left( 1 - \exp \left\{ - \left[ \sum_{l \in \mathbb{Z}} g_2(t, A_{l,lz}, \dots, A_{l-q,l-qz}) - \varepsilon_2 \right]_+ \right\} \right) \right] dt. \end{aligned}$$

Note, first of all, that  $\mu_n \rightarrow \mu$  in  $\mathbb{M}_0(\mathbb{R}^d)$ . Since the functions  $(\tilde{\kappa}_n)$  and  $\kappa$  are uniformly bounded, it is enough to prove the convergence in (4.12) when integrating over the set  $\{\tau < |z| < M\}$  for any finite  $M > \tau$ . Using the fact that  $J_n/n \rightarrow 0$ , we need to check that, for any  $K$ ,

$$\int_{\tau < |z| < M} \kappa_n^{(K)}(z) \mu_n(dz) \rightarrow \int_{\tau < |z| < M} \kappa(z) \mu(dz), \tag{4.13}$$

with

$$\begin{aligned} \kappa_n^{(K)}(z) = \sum_{t=-K+1}^{n+K} \mathbb{E} & \left[ \left( 1 - \exp \left\{ - \left[ \sum_{l=1-t}^{n-t} g_1 \left( \frac{t+l}{n}, A_{l,lz} \mathbf{1}\{|l| \leq J_n\}, \dots, \right. \right. \right. \right. \\ & \left. \left. \left. \left. A_{l-q,l-qz} \mathbf{1}\{|l-q| \leq J_n\} \right) - \varepsilon_1 \right]_+ \right\} \right) \right. \\ & \left. \times \left( 1 - \exp \left\{ - \left[ \sum_{l=1-t}^{n-t} g_2 \left( \frac{t+l}{n}, A_{l,lz} \mathbf{1}\{|l| \leq J_n\}, \dots, \right. \right. \right. \right. \right. \\ & \left. \left. \left. \left. A_{l-q,l-qz} \mathbf{1}\{|l-q| \leq J_n\} \right) - \varepsilon_2 \right]_+ \right\} \right) \right] \frac{1}{n}. \end{aligned}$$

Recall that the supports of  $g_1$  and  $g_2$  do not intersect the set

$$[0, 1] \times \{(x_0, \dots, x_q) : \max\{|x_0|, \dots, |x_q|\} < \delta\}$$

for some  $\delta > 0$ . Assumptions (2.2)–(2.4) imply that

$$\mathbb{P} \left( \|A_{l,l}\| \geq \frac{\delta}{M} \text{ for some } |l| \geq K \right) \rightarrow 0 \text{ as } K \rightarrow \infty.$$

Since the limit in (4.13) does not depend on  $K$ , we may replace  $\kappa_n^{(K)}$  in it (but still using the same notation) with

$$\begin{aligned} \kappa_n^{(K)}(z) = \sum_{t=-K+1}^{n+K} \mathbb{E} & \left[ \left( 1 - \exp \left\{ - \left[ \sum_{l \in \mathbb{Z}} g_1 \left( \frac{t+l}{n}, A_{l,lz}, \dots, A_{l-q,l-qz} \right) - \varepsilon_1 \right]_+ \right\} \right) \right. \\ & \left. \times \left( 1 - \exp \left\{ - \left[ \sum_{l \in \mathbb{Z}} g_2 \left( \frac{t+l}{n}, A_{l,lz}, \dots, A_{l-q,l-qz} \right) - \varepsilon_2 \right]_+ \right\} \right) \right] \frac{1}{n}. \end{aligned}$$

However,  $\kappa_n^{(K)} \rightarrow \kappa$  uniformly (in  $z$ ). Therefore, (4.13) follows, e.g. by Billingsley (1968, Theorem 5.5). Having now established (4.12), we let  $\tau \rightarrow 0$  to obtain

$$\int_{|z| > \tau} \kappa(z) \mu(dz) \rightarrow \mathbb{E}[(\text{Leb} \times \mu) \circ T_{\mathbb{A}}^{-1}(F_{g_1, g_2, \varepsilon_1, \varepsilon_2})],$$

as required.

It remains only to prove the asymptotic equivalence (4.10)  $\sim$  (4.11). Define  $C_n = \{-J_n - q + 1, \dots, n + J_n\}$ . Subtracting (4.11) from (4.10) yields

$$\begin{aligned} & \sum_{t=-J_n-q+1}^{n+J_n} r_n (\mathbb{E}[\Theta_n \mathbf{1}\{|Z_t| > \tau\gamma_n, |Z_s| \leq \tau\gamma_n, \text{ all } s \in C_n, s \neq t\}] \\ & \quad - \mathbb{E}[\Theta'_n \mathbf{1}\{|Z_t| > \tau\gamma_n\}]) \\ &= \sum_{t=-J_n-q+1}^{n+J_n} r_n \mathbb{E}[(\Theta_n - \Theta'_n) \mathbf{1}\{|Z_t| > \tau\gamma_n, |Z_s| \leq \tau\gamma_n, \text{ all } s \in C_n, s \neq t\}] \quad (4.14) \\ & \quad + \sum_{t=-J_n-q+1}^{n+J_n} r_n \mathbb{E}[\Theta'_n \mathbf{1}\{|Z_t| > \tau\gamma_n\} (1 - \mathbf{1}\{|Z_s| \leq \tau\gamma_n, \text{ all } s \in C_n, s \neq t\})]. \end{aligned} \quad (4.15)$$

Since  $\Theta'_n \leq 1$ , we can bound (4.15) by

$$\sum_{t=-J_n-q+1}^{n+J_n} r_n \mathbb{P}(|Z| > \tau\gamma_n) (1 - (1 - \mathbb{P}(|Z| > \tau\gamma_n))^{n+q+2J_n}) \rightarrow 0$$

as  $n \rightarrow \infty$ , by the choice of  $r_n$  and the fact that  $J_n/n \rightarrow 0$ . Lemma 4.2, below, shows that (4.14) tends to 0 as  $n \rightarrow \infty$ .

**Lemma 4.1.** *For the stationary process  $(X_k)_{k \in \mathbb{Z}}$  in (1.1), we have, under assumptions (2.2)–(2.4) and (4.6),*

$$\lim_{n \rightarrow \infty} r_n \mathbb{P}\left(\max_{1 \leq k \leq n} \left| \sum_{|j| > J_n} A_{k,j} Z_{k-j} \right| > \gamma_n \varepsilon\right) = 0$$

for any  $\varepsilon > 0$ .

*Proof.* By stationarity we have

$$r_n \mathbb{P}\left(\max_{1 \leq k \leq n} \left| \sum_{|j| > J_n} A_{k,j} Z_{k-j} \right| > \gamma_n \varepsilon\right) \leq r_n n \mathbb{P}\left(\left| \sum_{|j| > J_n} A_{k,j} Z_{k-j} \right| > \gamma_n \varepsilon\right).$$

Using Remark 2.1 and the definition of  $r_n$ , we see that the above expression is bounded above by  $o(1)nr_n \mathbb{P}(|Z| > \gamma_n \varepsilon)$ , and this tends to 0 as  $n \rightarrow \infty$ .

**Lemma 4.2.** *Denote the sum in (4.14) by  $\tilde{\Delta}_n$ . Then*

$$\lim_{\tau \rightarrow 0} \limsup_{n \rightarrow \infty} \tilde{\Delta}_n = 0.$$

*Proof.* Note that, by taking norms, it is enough to consider the one-dimensional case  $d = p = 1$ . Furthermore, it is clearly enough to consider a single function  $g$  and  $\varepsilon > 0$ , and prove that

$$\lim_{\tau \rightarrow 0} \limsup_{n \rightarrow \infty} r_n \Delta_n = 0, \quad (4.16)$$

where

$$\begin{aligned} \Delta_n = \sum_{t \in C_n} \mathbb{E} & \left[ \left( \exp \left\{ - \left[ \sum_{k=1}^n g \left( \frac{k}{n}, R_{k,n} + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k,j} Z_{k-j}, \dots, R_{k-q,n} \right. \right. \right. \right. \\ & \left. \left. \left. + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k-q,j} Z_{k-q-j} \right) - \varepsilon \right] \right\} \right. \\ & \left. - \exp \left\{ - \left[ \sum_{k=1}^n g \left( \frac{k}{n}, R_{k,n} + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k,j} Z_{k-j} \mathbf{1}\{t = k - j\}, \dots, R_{k-q,n} \right. \right. \right. \right. \\ & \left. \left. \left. + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k-q,j} Z_{k-q-j} \mathbf{1}\{t = k - q - j\} \right) - \varepsilon \right] \right\} \right) \\ & \times \mathbf{1}\{|Z_t| > \tau \gamma_n, |Z_s| \leq \tau \gamma_n, \text{ all } s \in C_n, s \neq t\} \Big], \end{aligned}$$

where, as above,  $C_n = \{-J_n - q + 1, \dots, n + J_n\}$ . Let  $L$  be the Lipschitz constant of  $g$  with respect to the metric on  $\mathbb{E}^q$  given by

$$d((s, x_0, \dots, x_q), (t, y_0, \dots, y_q)) = |s - t| + \min\{1, |x_0 - y_0| + \dots + |x_q - y_q|\}.$$

Note that, in the obvious notation,

$$\begin{aligned} |\Delta_n| & \leq L \mathbb{E} \sum_{k=1}^n \min \left\{ 1, \sum_{i=0}^q \frac{1}{\gamma_n} \left| \sum_{|j| \leq J_n} A_{k-i,j} Z_{k-i-j} \mathbf{1}\{|Z_{k-i-j}| \leq \tau \gamma_n\} \right| \right\} \\ & \leq Ln(q + 1) \mathbb{E} \min \left[ 1, \frac{1}{\gamma_n} \left| \sum_{|j| \leq J_n} A_{0,j} Z_{-j} \mathbf{1}\{|Z_{-j}| \leq \tau \gamma_n\} \right| \right] \\ & = Ln(q + 1) \int_0^1 \mathbb{P} \left( \left| \sum_{|j| \leq J_n} A_{0,j} Z_{-j} \mathbf{1}\{|Z_{-j}| \leq \tau \gamma_n\} \right| > x \gamma_n \right) dx. \end{aligned}$$

Suppose first that  $0 < \alpha < 1$ . We have, by (2.9), as  $\tau \downarrow 0$ ,

$$r_n \Delta_n \leq o(1) \frac{nr_n}{\gamma_n} \int_0^{\gamma_n} \mathbb{P}(|Z| > x) dx = o(1) nr_n \mathbb{P}(|Z| > \gamma_n) \rightarrow 0,$$

by Karamata’s theorem, and (4.16) follows.

Now consider the case in which  $\alpha \geq 1$ . We abbreviate

$$\Delta_n := \mathbb{E}(D_n) = \sum_{t=-J_n-q+1}^{n+J_n} \mathbb{E}(D_n \mathbf{1}\{B_t\}),$$

where

$$B_t = \{|Z_t| > \tau \gamma_n, |Z_s| \leq \tau \gamma_n, \text{ all } s \in C_n, s \neq t\}.$$

Since  $g$  has compact support, there is a  $\delta > 0$  such that  $g(s, x_0, \dots, x_q) = 0$  for all  $s \in [0, 1]$

and  $\{(x_0, \dots, x_q) : |x_0| + \dots + |x_q| < \delta\}$ . Let  $t \in C_n$ . We have, on the event  $B_t$ ,

$$|D_n| \mathbf{1}\{B_t\} \leq \left| \sum_{k=1}^n g\left(\frac{k}{n}, R_{k,n} + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k,j} Z_{k-j}, \dots, R_{k-q,n} + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k-q,j} Z_{k-q-j}\right) - \sum_{k=1}^n g\left(\frac{k}{n}, R_{k,n} + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k,j} Z_{k-j} \mathbf{1}\{t = k - j\}, \dots, R_{k-q,n} + \frac{1}{\gamma_n} \sum_{|j| \leq J_n} A_{k-q,j} Z_{k-q-j} \mathbf{1}\{t = k - q - j\}\right) \right| \mathbf{1}\{B_t\}.$$

Let  $K_t = \{k : t - J_n \leq k \leq t + q + J_n\}$ , and decompose the last expression into the sum over  $K_t$  and  $\{1, \dots, n\} \setminus K_t$ . Then, on the event  $B_t$ ,  $|D_n|$  is bounded above by

$$\begin{aligned} & L \sum_{k \in K_t} \min \left\{ 1, \sum_{i=0}^q \frac{1}{\gamma_n} \left| \sum_{|j| \leq J_n} A_{k-i,j} Z_{k-i-j} \mathbf{1}\{|Z_{k-i-j}| \leq \tau \gamma_n\} \mathbf{1}\{j \neq k - i - t\} \right| \right\} \\ & + \|g\|_\infty \sum_{k \notin K_t} \mathbf{1} \left\{ \sum_{i=0}^q \left( |R_{k-i,n}| + \frac{1}{\gamma_n} \left| \sum_{|j| \leq J_n} A_{k-i,j} Z_{k-i-j} \mathbf{1}\{|Z_{k-i-j}| \leq \tau \gamma_n\} \right| \right) > \delta \right\} \\ & := D_{n,1} + D_{n,2}. \end{aligned} \tag{4.17}$$

We start with  $D_{n,2}$ . Recall that, for all  $i$ ,  $|R_{k-i,n}|$  is bounded by  $\beta_n \downarrow 0$ . In the sequel,  $C$  denotes a finite positive constant that may change from time to time. We see that, for large  $n$ ,

$$\begin{aligned} r_n \sum_{t=-J_n-q+1}^{n+J_n} \mathbb{E}(D_{n,2} \mathbf{1}\{B_t\}) & \leq Cr_n(n + q + 2J_n)n \mathbb{P}(|Z| > \tau \gamma_n) \\ & \quad \times \mathbb{P}\left(\sum_{i=0}^q \left| \sum_{|j| \leq J_n} A_{k-i,j} Z_{k-i-j} \mathbf{1}\{|Z_{k-i-j}| \leq \tau \gamma_n\} \right| > \frac{\delta \gamma_n}{2}\right) \\ & \leq C\tau^{-\alpha}n(q + 1) \\ & \quad \times \mathbb{P}\left(\left| \sum_{|j| \leq J_n} A_{k,j} Z_{k,j} \mathbf{1}\{|Z_{k,j}| \leq \tau \gamma_n\} \right| > \frac{\delta \gamma_n}{2(q + 1)}\right). \end{aligned}$$

Using (2.9) shows that, for small  $\tau > 0$ , this is further bounded by  $Cn \mathbb{P}(|Z| > \gamma_n)$ , and this tends to 0 as  $n \rightarrow \infty$  by the choice of  $\gamma_n$ .

It remains to consider the term  $D_{n,1}$  in (4.17). Note in  $D_{n,1}$  that, for each  $t$ ,  $k$  is restricted to at most  $2J_n + q + 1$  possible values. We have

$$\begin{aligned} r_n \sum_{t=-J_n-q+1}^{n+J_n} \mathbb{E}(D_{n,1} \mathbf{1}\{B_t\}) & \leq Cr_n(n + q + 2J_n) \mathbb{P}(|Z| > \tau \gamma_n) \\ & \quad \times J_n \mathbb{E} \min \left[ 1, \sum_{i=0}^q \frac{1}{\gamma_n} \left| \sum_{|j| \leq J_n} A_{k-i,j} Z_{k-i-j} \mathbf{1}\{|Z_{k-i-j}| \leq \tau \gamma_n\} \right| \right] \\ & \leq C(q + 1) \frac{J_n}{\gamma_n} \frac{\mathbb{P}(|Z| > \tau \gamma_n)}{\mathbb{P}(|Z| > \gamma_n)} \mathbb{E} \left| \sum_{|j| \leq J_n} A_{0,j} Z_{-j} \mathbf{1}\{|Z_{-j}| \leq \tau \gamma_n\} \right|. \end{aligned} \tag{4.18}$$

Suppose first that  $\alpha = 1$ . For large  $n$ , the last expression can be bounded by

$$C\tau^{-1} \frac{J_n}{\gamma_n} \mathbb{E}(|Z| \mathbf{1}\{|Z| \leq \tau\gamma_n\}) \mathbb{E} \sum_{|j| \leq J_n} |A_j|.$$

Note that  $\mathbb{E} \sum_{|j| \leq J_n} |A_j|$  stays bounded by (2.2). Furthermore, the function

$$l(x) = \mathbb{E}(|Z| \mathbf{1}\{|Z| \leq x\})$$

is slowly varying. Therefore, the above expression vanishes as  $n \rightarrow \infty$  by (4.6).

Next consider the case in which  $\alpha > 1$ . Let  $\mu_n = \mathbb{E}(Z \mathbf{1}\{|Z| \leq \tau\gamma_n\})$ . Note that

$$\begin{aligned} \mathbb{E} \left| \sum_{|j| \leq J_n} A_j Z_j \mathbf{1}\{|Z_j| \leq \tau\gamma_n\} \right| &\leq \mathbb{E} \left| \sum_{|j| \leq J_n} A_j (Z_j \mathbf{1}\{|Z_j| \leq \tau\gamma_n\} - \mu_n) \right| + |\mu_n| \mathbb{E} \left| \sum_{|j| \leq J_n} A_j \right| \\ &=: I + II. \end{aligned}$$

Let us start with *II*. Since  $\mathbb{E} Z = 0$ , we see that, as  $n \rightarrow \infty$ ,

$$|\mu_n| \leq \mathbb{E}(|Z| \mathbf{1}\{|Z| > \tau\gamma_n\}) \sim C\tau\gamma_n \mathbb{P}(|Z| > \tau\gamma_n).$$

Furthermore, to deal with  $\sum_{|j| \leq J_n} A_j$ , we use assumptions (2.2)–(2.4). Suppose, for example, that  $1 < \alpha \leq 2$ . Choose  $\varepsilon$  small enough so that  $\alpha - \varepsilon > 1$ , and note that

$$\mathbb{E} \left| \sum_{|j| \leq J_n} A_j \right| \leq C J_n^{1-(\alpha-\varepsilon)^{-1}} \mathbb{E} \left( \sum_{|j| \leq J_n} |A_j|^{\alpha-\varepsilon} \right)^{1/(\alpha-\varepsilon)} \leq C J_n^{1-(\alpha-\varepsilon)^{-1}},$$

so the corresponding term in (4.18) is bounded, for large  $n$ , by

$$C\tau^{1-2\alpha} J_n^{2-(\alpha-\varepsilon)^{-1}} \mathbb{P}(|Z| > \gamma_n).$$

Note that, for small enough  $\varepsilon$ ,  $\theta := 2 - (\alpha - \varepsilon)^{-1} < \alpha$ . In that case the above expression is  $o(J_n/\gamma_n) \rightarrow 0$  as  $n \rightarrow \infty$  by (4.6). Similarly, in the case in which  $\alpha > 2$  this term goes to 0 as well.

For *I*, by the Burkholder–Davis–Gundy inequality,

$$\begin{aligned} \mathbb{E} \left| \sum_{|j| \leq J_n} A_j (Z_j \mathbf{1}\{|Z_j| \leq \tau\gamma_n\} - \mu_n) \right| \\ \leq C \mathbb{E}_A \left[ \mathbb{E}_Z \left( \sum_{|j| \leq J_n} A_j^2 (Z_j \mathbf{1}\{|Z_j| \leq \tau\gamma_n\} - \mu_n)^2 \right)^{1/2} \right]. \end{aligned}$$

The notation  $\mathbb{E}_Z, \mathbb{E}_A$ , etc. specifies the random objects with respect to which the expectation is taken. We use, once again, assumptions (2.2)–(2.4). Assuming again that  $1 < \alpha \leq 2$ , and choosing  $\varepsilon$  as above, we see that the above expression is bounded by

$$\begin{aligned} C \mathbb{E}_A \left[ \mathbb{E}_Z \left( \sum_{|j| \leq J_n} |A_j|^{\alpha-\varepsilon} |Z_j \mathbf{1}\{|Z_j| \leq \tau\gamma_n\} - \mu_n|^{\alpha-\varepsilon} \right)^{1/(\alpha-\varepsilon)} \right] \\ \leq C (\mathbb{E} |Z \mathbf{1}\{|Z| \leq \tau\gamma_n\} - \mu_n|^{\alpha-\varepsilon})^{1/(\alpha-\varepsilon)} \mathbb{E} \left( \sum_{|j| \leq J_n} |A_j|^{\alpha-\varepsilon} \right)^{1/(\alpha-\varepsilon)}, \end{aligned}$$

which is bounded, and so the corresponding term in (4.18) converges to 0 because  $J_n/\gamma_n \rightarrow 0$  as  $n \rightarrow \infty$ . The case in which  $\alpha > 2$  is entirely analogous (and simpler).

### 5. First applications

Theorem 4.1 provides a fairly complete description of the asymptotics of the probability of rare events for the sequence  $(X_k)$ . In this section we provide some immediate applications of this theorem. For the sake of simplicity and to avoid complicated formulae, we restrict attention to the case where both  $A_{k,j}$  and  $Z_j$  are univariate and  $A_{k,j} \geq 0$  a.s. Then  $Z$  has a univariate regularly varying distribution and its limiting measure can be written as in (2.6).

**Example 5.1.** (*Order statistics.*) The first application is to order statistics. Let  $X_{i:n}$  be the  $i$ th order statistic of  $X_1, \dots, X_n$  in descending order. That is,

$$X_{1:n} \geq X_{2:n} \geq \dots \geq X_{n:n}.$$

Fix an integer  $q \geq 1$ , and consider the  $q$ -dimensional vector  $(X_{1:n}, \dots, X_{q:n})$  consisting of the  $q$  largest values. We denote by  $A_i^*$  the  $i$ th order statistic of the sequence  $\{A_{j,j}, j \in \mathbb{Z}\}$  in descending order; under assumptions (2.2)–(2.4), this is a well-defined random variable. Note that, for  $\infty > u_1 > u_2 > \dots > u_q > 0$ , we can write

$$P(X_{1:n} > \gamma_n u_1, \dots, X_{q:n} > \gamma_n u_q) = P(N_n^0 \in B(u_1, \dots, u_q))$$

with

$$B = B(u_1, \dots, u_q) = \bigcap_{i=1}^q \{\xi : \xi([0, 1] \times (u_i, \infty)) \geq i\}.$$

Then we have the following implication of Theorem 4.1.

**Corollary 5.1.** *Let  $d = p = 1$ . Assume that  $A_{k,j} \geq 0$  for all  $k, j$  and that the hypotheses of Theorem 4.1 hold. Then, as  $n \rightarrow \infty$ ,*

$$\frac{P((X_{1:n} > \gamma_n u_1, \dots, X_{q:n} > \gamma_n u_q))}{n P(|Z| > \gamma_n)} \rightarrow w E \min_{i=1, \dots, q} (A_i^* u_i^{-1})^\alpha.$$

*Proof.* First note that

$$\begin{aligned} m^0(B(u_1, \dots, u_q)) &= E[(\text{Leb} \times \mu) \circ T_{\mathbb{A}}^{-1}(B(u_1, \dots, u_q))] \\ &= E\left[\mu\left\{z : \sum \delta_{A_{j,j}z}(u_1, \infty) \geq 1, \dots, \sum \delta_{A_{j,j}z}(u_q, \infty) \geq q\right\}\right] \\ &= E[\mu\{z : A_1^* z \in (u_1, \infty), \dots, A_q^* z \in (u_q, \infty)\}] \\ &= w E \min_{i=1, \dots, q} (A_i^* u_i^{-1})^\alpha. \end{aligned}$$

The claim, therefore, is a direct application of Theorem 4.1 once we show that the set

$$B(u_1, \dots, u_q)$$

is bounded away from the null measure and  $m(\partial B(u_1, \dots, u_q)) = 0$ .

The set  $B(u_1, \dots, u_q)$  is open. To see this, write  $B(u_1, \dots, u_q) = \bigcap_{i=1}^q B_i$  with  $B_i = \{\xi : \xi([0, 1] \times (u_i, \infty)) \geq i\}$ . Then,  $B_i^c = \{\xi : \xi([0, 1] \times (u_i, \infty)) < i\}$  and, for a sequence of measures  $(\xi_n) \subset B_i^c$  with  $\xi_n \xrightarrow{V} \xi$ , we have, by the Portmanteau theorem,

$$i > \liminf_{n \rightarrow \infty} \xi_n([0, 1] \times (u_i, \infty)) \geq \xi([0, 1] \times (u_i, \infty)).$$

Hence,  $\xi \in B_i^c$  so  $B_i^c$  is closed. This shows that  $B_i$  is open and, consequently,  $B(u_1, \dots, u_q)$  is open. Similarly, the set  $C_i = \{\xi : \xi([0, 1] \times [u_i, \infty)) \geq i\}$  is closed. Since  $B_i \subset C_i$  and  $C_i$  does not contain the null measure, we see that each  $B_i$  is bounded away from the null measure and, hence, so is  $B$ .

Furthermore, it follows, by the above calculation, that

$$\begin{aligned} m(\partial B(u_1, \dots, u_q)) &= m(\overline{B(u_1, \dots, u_q)}) - m(B(u_1, \dots, u_q)) \\ &\leq m\left(\bigcap_{i=1}^q C_i\right) - m(B(u_1, \dots, u_q)) \\ &= \mathbb{E}\left[\int_0^\infty \mathbf{1}_{[u_1/A_1^*, \infty)}(z) \dots \mathbf{1}_{[u_q/A_q^*, \infty)}(z) w \alpha z^{-\alpha-1} dz\right] \\ &\quad - \mathbb{E}\left[\int_0^\infty \mathbf{1}_{(u_1/A_1^*, \infty)}(z) \dots \mathbf{1}_{(u_q/A_q^*, \infty)}(z) w \alpha z^{-\alpha-1} dz\right] \\ &= 0. \end{aligned}$$

This proves the claim.

**Example 5.2.** (*Hitting times.*) Next we consider the large deviations of first hitting times. Take  $a > 0$ , and consider the first hitting time of  $(a\gamma_n, \infty)$  by  $\{X_k\}$ , namely

$$\tau_n = \inf\{k : X_k > a\gamma_n\}.$$

**Corollary 5.2.** *Let  $d = p = 1$ . Assume that  $A_{k,j} \geq 0$  for all  $k, j$  and that the hypotheses of Theorem 4.1 hold. Then, for any  $\lambda > 0$  as  $n \rightarrow \infty$ ,*

$$\frac{\mathbb{P}(\tau_n \leq \lambda n)}{n \mathbb{P}(|Z| > \gamma_n)} \rightarrow \lambda \mathbb{E}[(A_1^*)^\alpha] w a^{-\alpha}.$$

*Proof.* It is enough to prove the statement for  $\lambda = 1$ ; the proof for a general  $\lambda > 0$  will then follow via denoting  $m = [\lambda n]$  and redefining appropriately the sequence  $(\gamma_n)$ . We have

$$r_n \mathbb{P}(\tau_n \leq n) = r_n \mathbb{P}\left(\sup_{0 \leq k \leq n} \frac{X_k}{\gamma_n} > a\right) = r_n \mathbb{P}(X_{1:n} > \gamma_n a),$$

and the statement follows from Corollary 5.1.

### 6. Large deviations of the partial sums

In this section large deviation results for the partial sums  $S_n = X_1 + \dots + X_n$ ,  $n = 1, 2, \dots$ , are considered. The main idea is to start from Theorem 4.1 and derive results for the partial sum by summing up the points in the point measure  $N_n$ , while applying the continuous mapping argument.

It turns out that for success of this program additional assumptions are needed. The first assumption is designed to control the contribution of ‘relatively small’ values of the  $X_k$ s. To this end, we introduce the following condition: for each  $\delta > 0$ ,

$$\lim_{\varepsilon \downarrow 0} \limsup_n r_n \mathbb{P}\left(\left|\sum_{k=1}^n X_k \mathbf{1}_{\{|X_k| \leq \gamma_n \varepsilon\}}\right| > \gamma_n \delta\right) = 0. \tag{6.1}$$

The second assumption we need is

$$\text{the sum } \sum A_{j,j} \text{ converges a.s. and } \mathbb{E} \sup_{J \subset \mathbb{Z}} \left\| \sum_{j \in J} A_{j,j} \right\|^\alpha < \infty. \tag{6.2}$$

**Theorem 6.1.** Assume that the hypotheses of Theorem 4.1 hold and, in addition, that (6.1) and (6.2) hold. Then

$$r_n \mathbb{P}(\gamma_n^{-1} S_n \in \cdot) \rightarrow \mathbb{E} \left[ \mu \left( z : \sum_{j \in \mathbb{Z}} A_{j,j} z \in \cdot \right) \right] \text{ in } \mathbb{M}_0(\mathbb{R}^d). \tag{6.3}$$

**Remark 6.1.** Note that the large deviation result is uniform in the sense that the normalization  $r_n$  is the same for all sets. In particular, the univariate result ( $p = d = 1$ ) can be stated as

$$\lim_{n \rightarrow \infty} \frac{\mathbb{P}(S_n > \rho \gamma_n)}{n \mathbb{P}(|Z| > \rho \gamma_n)} = w \mathbb{E} \left( \left[ \left( \sum A_{j,j} \right)^+ \right]^\alpha \right) + (1 - w) \mathbb{E} \left( \left[ \left( \sum A_{j,j} \right)^- \right]^\alpha \right)$$

for every  $\rho > 0$ , where the limiting measure associated with  $Z$  is given by (2.6).

**Remark 6.2.** In some cases, replacing conditions (6.1) and (6.2) by somewhat stronger conditions, we can modify the proof of Theorem 6.1 to obtain large deviations of the partial sum of the absolute values of the process. It is sufficient to change condition (6.2) to

$$\mathbb{E} \left( \sum \|A_{j,j}\| \right)^\alpha < \infty. \tag{6.4}$$

If  $0 < \alpha \leq 1$ , or  $\alpha > 1$  and  $n/\gamma_n \rightarrow 0$ , then it is sufficient to change condition (6.1) to, for each  $\delta > 0$ ,

$$\lim_{\varepsilon \downarrow 0} \limsup_n r_n \mathbb{P} \left( \sum_{k=1}^n |X_k| \mathbf{1}\{|X_k| \leq \gamma_n \varepsilon\} > \gamma_n \delta \right) = 0. \tag{6.5}$$

In this case we conclude that  $S_n^{\text{abs}} = \sum_{k=1}^n |X_k|$  satisfies

$$r_n \mathbb{P}(\gamma_n^{-1} S_n^{\text{abs}} \in \cdot) \rightarrow \mathbb{E} \left[ \mu \left( z : \sum_{j \in \mathbb{Z}} |A_{j,j} z| \in \cdot \right) \right] \text{ in } \mathbb{M}_0(\mathbb{R}^d).$$

If, on the other hand,  $\alpha > 1$  and  $\gamma_n \equiv n$ , then it is sufficient to change condition (6.1) to, for each  $\delta > 0$ ,

$$\lim_{\varepsilon \downarrow 0} \limsup_n r_n \mathbb{P} \left( \left| \sum_{k=1}^n (|X_k| - \mathbb{E} |X_0|) \mathbf{1}\{|X_k| \leq n \varepsilon\} \right| > n \delta \right) = 0, \tag{6.6}$$

and then  $(S_n^{\text{abs}} - n \mathbb{E} |X_0|)$  satisfies

$$r_n \mathbb{P}(n^{-1} (S_n^{\text{abs}} - n \mathbb{E} |X_0|) \in \cdot) \rightarrow \mathbb{E} \left[ \mu \left( z : \sum_{j \in \mathbb{Z}} |A_{j,j} z| \in \cdot \right) \right] \text{ in } \mathbb{M}_0(\mathbb{R}^d). \tag{6.7}$$

*Proof of Theorem 6.1.* The idea is to divide  $S_n$  into three parts. One part contains the terms where  $\varepsilon < |X_k| \leq 1/\varepsilon$  for a small positive  $\varepsilon$ , and the other two parts contain terms with  $|X_k| \leq \varepsilon$  and  $|X_k| > 1/\varepsilon$ . The contributions from the latter two parts turn out to be negligible.

For  $0 < \varepsilon < 1$ , let  $g_\varepsilon$  be a function  $[0, 1] \times (\mathbb{R}^d \setminus \{0\}) \mapsto \mathbb{R}^d$  such that  $g_\varepsilon(t, x) = g_\varepsilon(x) = x$  on  $\varepsilon < |x| \leq 1/\varepsilon$ ,  $g_\varepsilon(x) = 0$  for all other values of  $x$ . First we consider the convergence of  $r_n \mathbb{P}(N_n^0(g_\varepsilon) \in \cdot)$  with  $N_n^0$  as in Theorem 4.1. Let  $m^0$  be the limiting measure in (4.3) with  $q = 0$ . Note that  $g_\varepsilon$  is continuous except at the points  $|x| = \varepsilon$  and  $1/\varepsilon$ . By Remark 4.1,

$$m^0\{\xi : \xi([0, 1] \times \{|x| = \varepsilon \text{ or } 1/\varepsilon\}) > 0\} = 0.$$

Hence, the map  $\xi \mapsto \xi(g_\varepsilon)$  from  $\mathbb{N}_p$  to  $\mathbb{R}^d$  satisfies the continuity assumption in the mapping theorem (Lemma A.2). Therefore, Theorem 4.1, with  $q = 0$ , together with the mapping theorem, implies that

$$r_n \mathbb{P}(N_n^0(g_\varepsilon) \in \cdot) \rightarrow \mathbb{E}\left[(\text{Leb} \times \mu)\left((t, z) : \sum_{j \in \mathbb{Z}} g_\varepsilon(t, A_{j,j}z) \in \cdot\right)\right] := \tilde{m}_\varepsilon(\cdot) \quad \text{in } \mathbb{M}_0(\mathbb{R}^d). \tag{6.8}$$

Set  $\tilde{m}_n(\cdot) = r_n \mathbb{P}(\gamma_n^{-1} S_n \in \cdot)$  and  $\tilde{m}$  as on the right-hand side of (6.3). We need to show that  $\tilde{m}_n(f) \rightarrow \tilde{m}(f)$  for any  $f \in C_0(\mathbb{R}^d)$ ; in fact, it is sufficient to consider uniformly continuous  $f$  (see Appendix A). For any such  $f$ , there is an  $\eta > 0$  such that  $x \in \text{support}(f)$  implies that  $|x| > \eta$ . For any  $\delta > 0$ ,

$$\begin{aligned} \tilde{m}_n(f) &= r_n \mathbb{E}[f(\gamma_n^{-1} S_n)] \\ &= r_n \mathbb{E}[f(\gamma_n^{-1} S_n) \mathbf{1}\{|\gamma_n^{-1} S_n - N_n^0(g_\varepsilon)| > \delta\}] \\ &\quad + r_n \mathbb{E}[f(\gamma_n^{-1} S_n) \mathbf{1}\{|\gamma_n^{-1} S_n - N_n^0(g_\varepsilon)| \leq \delta\}]. \end{aligned}$$

The first term is bounded above by

$$\begin{aligned} |f|_\infty r_n \mathbb{P}(|\gamma_n^{-1} S_n - N_n^0(g_\varepsilon)| > \delta) &\leq |f|_\infty r_n \mathbb{P}\left(\left|\sum_{k=1}^n X_k \mathbf{1}\{|X_k| \leq \gamma_n \varepsilon\}\right| > \frac{\gamma_n \delta}{2}\right) \\ &\quad + |f|_\infty r_n \mathbb{P}\left(\left|\sum_{k=1}^n X_k \mathbf{1}\{|X_k| > \frac{\gamma_n}{\varepsilon}\}\right| > \frac{\gamma_n \delta}{2}\right). \end{aligned} \tag{6.9}$$

Assumption (6.1) guarantees that the first member on the right-hand side of (6.9) is asymptotically negligible. The second member on the right-hand side of (6.9) is, up to a constant, bounded above by

$$r_n \mathbb{P}\left(\max_{k=1, \dots, n} |X_k| > \frac{\gamma_n}{\varepsilon}\right) \leq r_n n \mathbb{P}\left(|X_0| > \frac{\gamma_n}{\varepsilon}\right) \rightarrow 0$$

as first  $n \rightarrow \infty$  and then  $\varepsilon \rightarrow 0$ . Therefore, the statement of the theorem will follow once we show that

$$\lim_{\delta \downarrow 0} \limsup_{\varepsilon \downarrow 0} \limsup_{n \rightarrow \infty} r_n \mathbb{E}[|f(\gamma_n^{-1} S_n) - f(N_n^0(g_\varepsilon))| \mathbf{1}\{|\gamma_n^{-1} S_n - N_n^0(g_\varepsilon)| \leq \delta\}] = 0 \tag{6.10}$$

and

$$\lim_{\varepsilon \downarrow 0} \tilde{m}_\varepsilon(f) = \tilde{m}(f). \tag{6.11}$$

Indeed, in that case we could write, for each  $\varepsilon > 0$  and  $\delta > 0$ ,

$$\begin{aligned} |\tilde{m}_n(f) - \tilde{m}(f)| &\leq r_n \mathbb{E}[|f(\gamma_n^{-1} S_n) - f(N_n^0(g_\varepsilon))| \mathbf{1}\{|\gamma_n^{-1} S_n - N_n^0(g_\varepsilon)| \leq \delta\}] \\ &\quad + r_n \mathbb{E}[f(\gamma_n^{-1} S_n) \mathbf{1}\{|\gamma_n^{-1} S_n - N_n^0(g_\varepsilon)| > \delta\}] \\ &\quad + |r_n \mathbb{E}[f(N_n^0(g_\varepsilon)) \mathbf{1}\{|\gamma_n^{-1} S_n - N_n^0(g_\varepsilon)| \leq \delta\}] - r_n \mathbb{E}[f(N_n^0(g_\varepsilon))]| \\ &\quad + |r_n \mathbb{E}[f(N_n^0(g_\varepsilon))] - m_\varepsilon(f)| + |\tilde{m}_\varepsilon(f) - \tilde{m}(f)|. \end{aligned}$$

By (6.10), the argument in (6.9), (6.8), and (6.11), each term converges to 0 as first  $n \rightarrow \infty$ , then  $\varepsilon \downarrow 0$ , and finally  $\delta \downarrow 0$ .

It remains to prove (6.10) and (6.11). We start with (6.10). Choose  $\delta$  above to be smaller than  $\eta/2$ . The reason for this is that if either  $f(\gamma_n^{-1} S_n) > 0$  or  $f(N_n^0(g_\varepsilon)) > 0$ , then, on  $\{|\gamma_n^{-1} S_n - N_n^0(g_\varepsilon)| \leq \delta\}$ , we have  $|N_n^0(g_\varepsilon)| > \eta/2$ . Since  $f$  is uniformly continuous, the expression in (6.10) is bounded above by

$$o_\delta(1)r_n \mathbb{E}[\mathbf{1}\{|N_n^0(g_\varepsilon)| > \eta/2\} \mathbf{1}\{|\gamma_n^{-1} S_n - N_n^0(g_\varepsilon)| \leq \delta\}] \leq o_\delta(1)r_n \mathbb{P}(|N_n^0(g_\varepsilon)| > \eta/2).$$

As  $n \rightarrow \infty$  and  $\varepsilon \downarrow 0$ , (6.8) and (6.11) (still to be proved) show that this remains bounded by (constant)  $o_\delta(1)$ . As  $\delta \downarrow 0$ , this converges to 0.

It remains to show (6.11). We have, as  $\varepsilon \downarrow 0$ ,

$$\begin{aligned} \tilde{m}_\varepsilon(f) &= \int_{\Omega} \int_{\mathbb{R}^d \setminus \{0\}} f\left(\sum g_\varepsilon(A_{j,j}z)\right) \mu(dz) \mathbb{P}(d\omega) \\ &\rightarrow \int_{\Omega} \int_{\mathbb{R}^d \setminus \{0\}} f\left(\sum A_{j,j}z\right) \mu(dz) \mathbb{P}(d\omega) \\ &= \tilde{m}(f), \end{aligned}$$

by dominated convergence. Indeed,  $\sum g_\varepsilon(A_{j,j}z) \rightarrow \sum A_{j,j}z$ ,  $(\mu \times \mathbb{P})$ -almost everywhere as  $\varepsilon \downarrow 0$ ,  $f$  is continuous, and

$$\left| f\left(\sum g_\varepsilon(A_{j,j}z)\right) \right| \leq |f|_\infty \mathbf{1}\left\{ \sup_{J \subset \mathbb{Z}} \left| \sum_{j \in J} A_{j,j}z \right| > \eta \right\},$$

which is  $(\mu \times \mathbb{P})$ -integrable by the scaling property of the measure  $\mu$  and assumption (6.2).

### 6.1. Checking the conditions of Theorem 6.1

To apply Theorem 6.1, we need to verify the extra assumptions imposed there. In this section we provide conditions that are easier to check for some more specific models.

**Proposition 6.1.** *Let  $(X_k)$  be the stationary process in (1.1) satisfying the conditions of Theorem 4.1. If  $0 < \alpha < 1$  then (6.5) holds and, hence, (6.1) holds as well. If  $0 < \alpha \leq 1$  then (6.4) holds and, hence, (6.2) holds as well.*

*Proof.* Assume that  $0 < \alpha < 1$ . By Markov’s inequality, Karamata’s theorem, and, finally, Theorem 2.1,

$$\begin{aligned} & \lim_{\varepsilon \downarrow 0} \limsup_n r_n \mathbb{P} \left( \sum_{k=1}^n |X_k| \mathbf{1}\{|X_k| \leq \gamma_n \varepsilon\} > \gamma_n \delta \right) \\ & \leq \lim_{\varepsilon \downarrow 0} \limsup_n r_n (\gamma_n \delta)^{-1} \mathbb{E} \left( \sum_{k=1}^n |X_k| \mathbf{1}\{|X_k| \leq \gamma_n \varepsilon\} \right) \\ & = \lim_{\varepsilon \downarrow 0} \limsup_n r_n n (\gamma_n \delta)^{-1} \mathbb{E} |X_0| \mathbf{1}\{|X_0| \leq \gamma_n \varepsilon\} \\ & = \lim_{\varepsilon \downarrow 0} \limsup_n r_n n (\gamma_n \delta)^{-1} C(\gamma_n \varepsilon) \mathbb{P}(|X_0| > \gamma_n \varepsilon) \\ & = \lim_{\varepsilon \downarrow 0} \limsup_n C \frac{\varepsilon \mathbb{P}(|X_0| > \gamma_n \varepsilon)}{\delta \mathbb{P}(|Z| > \gamma_n)} \\ & = \lim_{\varepsilon \downarrow 0} C \frac{\varepsilon^{1-\alpha}}{\delta} \\ & = 0, \end{aligned}$$

and so (6.1) holds. If  $0 < \alpha \leq 1$  then, by (2.2) and the Cauchy–Schwarz inequality,

$$\begin{aligned} \mathbb{E} \left( \sum \|A_j\| \right)^\alpha & \leq \sum \mathbb{E} \|A_j\|^\alpha \\ & \leq \sum (\mathbb{E} \|A_j\|^{\alpha-\varepsilon})^{1/2} (\mathbb{E} \|A_j\|^{\alpha+\varepsilon})^{1/2} \\ & \leq \left( \sum \mathbb{E} \|A_j\|^{\alpha-\varepsilon} \right)^{1/2} \left( \sum \mathbb{E} \|A_j\|^{\alpha+\varepsilon} \right)^{1/2} \\ & < \infty, \end{aligned}$$

and so (6.4) holds.

If the sum (1.1) defining the process  $(X_k)$  is finite, then modest additional assumptions on the sequence  $(\mathbb{A}_k)_{k \in \mathbb{Z}}$  guarantee applicability of Theorem 6.1. We present one such situation.

**Proposition 6.2.** *Let  $(X_k)$  be the stationary process in (1.1) satisfying the conditions of Theorem 4.1. Suppose, further, that the sequence  $(\mathbb{A}_k)_{k \in \mathbb{Z}}$  is i.i.d. such that, for some  $M = 0, 1, 2, \dots$ ,  $A_{k,j} = 0$  a.s. for  $|j| > M$ . Then (6.1) holds and, further, (6.5) and (6.6) (as appropriate) hold. Also, both (6.2) and (6.4) hold, and so Theorem 6.1 applies.*

When the i.i.d. assumption of the sequence  $(\mathbb{A}_k)_{k \in \mathbb{Z}}$  is dropped, we can still obtain sufficient conditions for (6.1). See Lemma 7.3, below.

*Proof of Proposition 6.2.* For finite sums, condition (6.4) is a trivial consequence of (2.2)–(2.4). We will show that (6.1) holds; the proofs for (6.5) and (6.6) are similar. It is clearly enough to consider the case  $d = 1$ . Note further that

$$\begin{aligned} & \mathbb{P} \left( \left| \sum_{k=1}^n X_k \mathbf{1}\{|X_k| \leq \gamma_n \varepsilon\} \right| > \gamma_n \delta \right) \\ & \leq \mathbb{P}(\text{for some } k = 1, \dots, n, |A_{k,j} Z_{k-j}| > \gamma_n \varepsilon \text{ for at least two different } j) \\ & \quad + \mathbb{P} \left( \left| \sum_{k=1}^n X_k \mathbf{1}\{|A_{k,j} Z_{k-j}| \leq \gamma_n M \varepsilon \text{ for all } k = 1, \dots, n \text{ and } |j| \leq M\} \right| > \gamma_n \delta \right). \end{aligned} \tag{6.12}$$

The first term on the right-hand side of (6.12) is bounded by

$$n \sum_{\substack{i,j=-M,\dots,M \\ i \neq j}} P(|A_{0,i}Z_{-i}| > \gamma_n \varepsilon, |A_{0,j}Z_{-j}| > \gamma_n \varepsilon) = o(1)n P(|Z| > \gamma_n) = o(1) \left( \frac{1}{r_n} \right),$$

as in Lemma 3.4 of Hult and Samorodnitsky (2008). The second term on the right-hand side of (6.12) does not exceed

$$\sum_{j=-M}^M P \left( \left| \sum_{k=1}^n A_{k,j} Z_{k-j} \mathbf{1}\{|A_{k,j} Z_{k-j}| \leq \gamma_n M \varepsilon \text{ for all } k = 1, \dots, n \text{ and } |j| \leq M\} \right| > \frac{\gamma_n \delta}{2M + 1} \right). \tag{6.13}$$

By the assumed independence, for every  $|j| \leq M$ ,

$$\lim_{\varepsilon \downarrow 0} \limsup_n r_n P \left( \left| \sum_{k=1}^n A_{k,j} Z_{k-j} \mathbf{1}\{|A_{k,j} Z_{k-j}| \leq \gamma_n M \varepsilon \text{ for } k = 1, \dots, n, |j| \leq M\} \right| > \frac{\gamma_n \delta}{2M + 1} \right) = 0; \tag{6.14}$$

see the argument in Lemma 2.1 of Hult *et al.* (2005). Therefore, (6.1) follows.

Finally, for certain symmetric stochastic recurrence equations as in Examples 1.2 and 2.2, we provide sufficient conditions for the applicability of Theorem 6.1.

**Proposition 6.3.** *Suppose that the i.i.d. pairs  $(Y_k, Z_k)_{k \in \mathbb{Z}}$  are symmetric (i.e.  $(-Y_k, -Z_k) \stackrel{D}{=} (Y_k, Z_k)$ ),  $Z \in \text{RV}(\mu, \alpha)$  for some  $0 < \alpha < 2$ , and  $E \|Y\|^{\alpha+\varepsilon} < 1$  for some  $\varepsilon > 0$ . Then the random recursion (1.3) has a unique stationary solution, and it satisfies Theorem 6.1.*

*Proof.* Existence and uniqueness of a stationary solution follows from Corollary 2.3 of Hult and Samorodnitsky (2008), which also shows that this solution is of the form (1.1) and satisfies the assumptions of Theorem 4.1. For  $0 < \alpha < 1$ , the statement follows from Proposition 6.1. For  $\alpha \geq 1$ , we have, by convexity (see Lemma 3.3.1 of Kwapien and Woyczyński (1992)),

$$\begin{aligned} E \left( \sum \|A_j\| \right)^\alpha &= E \left( \sum_{j=0}^\infty \left\| \prod_{i=0}^j Y_i \right\| \right)^\alpha \\ &\leq E \left( \sum_{j=0}^\infty \prod_{i=0}^j \|Y_i\| \right)^\alpha \\ &\leq \left[ \sum_{j=0}^\infty \left( E \prod_{i=0}^j \|Y_i\|^\alpha \right)^{1/\alpha} \right]^\alpha \\ &< \infty, \end{aligned}$$

since  $E \|Y\|^\alpha < 1$ . Therefore, (6.4) holds. Furthermore, the symmetry assumption in the proposition guarantees that the stationary process  $(X_k)$  is symmetric in the sense that  $(X_k) \stackrel{D}{=} (\epsilon_k X_k)$ ,

where  $(\epsilon_k)$  is a sequence of i.i.d. Rademacher random variables independent of  $(X_k)$ . We conclude, as in the proof of Proposition 6.1, that

$$\begin{aligned} & \lim_{\epsilon \downarrow 0} \limsup_n r_n \mathbb{P} \left( \left| \sum_{k=1}^n X_k \mathbf{1}\{|X_k| \leq \gamma_n \epsilon\} \right| > \gamma_n \delta \right) \\ & \leq \lim_{\epsilon \downarrow 0} \limsup_n r_n (\gamma_n \delta)^{-2} \mathbb{E} \left( \left| \sum_{k=1}^n X_k \mathbf{1}\{|X_k| \leq \gamma_n \epsilon\} \right|^2 \right) \\ & = \lim_{\epsilon \downarrow 0} \limsup_n r_n (\gamma_n \delta)^{-2} \mathbb{E} \left( \sum_{k=1}^n X_k^2 \mathbf{1}\{|X_k| \leq \gamma_n \epsilon\} \right) \\ & = \lim_{\epsilon \downarrow 0} \limsup_n r_n n (\gamma_n \delta)^{-2} \mathbb{E} X_k^2 \mathbf{1}\{|X_k| \leq \gamma_n \epsilon\} \\ & = \lim_{\epsilon \downarrow 0} \limsup_n r_n n (\gamma_n \delta)^{-2} C (\gamma_n \epsilon)^2 \mathbb{P}(|X_k| > \gamma_n \epsilon) \\ & = \lim_{\epsilon \downarrow 0} \limsup_n C \frac{\epsilon^2 \mathbb{P}(|X_k| > \gamma_n \epsilon)}{\delta^2 \mathbb{P}(|Z| > \gamma_n)} \\ & = \lim_{\epsilon \downarrow 0} C \frac{\epsilon^{2-\alpha}}{\delta^2} \\ & = 0, \end{aligned}$$

proving (6.1).

### 7. Ruin probabilities

In this section we consider the univariate ( $d = p = 1$ ) ruin problem based on the sequence  $(X_k)$  in (1.1). Throughout this section, we assume that  $\alpha > 1$  (which requires, according to our assumptions, that  $\mathbb{E} Z = 0$ ), and let  $c > 0$  be the ‘drift’. We are interested in deriving the asymptotic decay of the so-called infinite-horizon ruin probability

$$\psi(u) = \mathbb{P} \left( \sup_n (S_n - cn) > u \right)$$

as  $u \rightarrow \infty$ . Here  $S_n = X_1 + \dots + X_n$  is the partial sum process.

As in Section 6, we need to assume extra technical conditions, mostly in order to control the contributions of the small jumps to the ruin probability. We start with some notation. For integer  $q \geq -1$ , let

$$\hat{X}_k^q = \sum_{|j|>q} A_{k,j} Z_{k-j}, \quad k \in \mathbb{Z}. \tag{7.1}$$

We assume that, for each  $q \geq -1$  and each  $\delta > 0$ ,

$$\lim_{\epsilon \downarrow 0} \limsup_n r_n \mathbb{P} \left( \left| \sum_{k=1}^n (|\hat{X}_k^q| - \mathbb{E} |\hat{X}_0^q|) \mathbf{1}\{|\hat{X}_k^q| \leq n\epsilon\} \right| > n\delta \right) = 0 \tag{7.2}$$

and that, for every  $q \geq 0$  and  $\gamma > 0$ ,

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{\mathbb{P}(\sup_{k \leq n} |\sum_{i=1}^k X_i^q \mathbf{1}\{|X_i^q| \leq n\delta\}| > n\gamma)}{n \mathbb{P}(|Z| > n)} = 0. \tag{7.3}$$

It is easy to check that condition (7.2) holds, for example, under the assumptions of Proposition 6.2. Sufficient conditions for (7.3) are given in Lemma 7.3, below.

**Theorem 7.1.** *Suppose that the conditions of Theorem 4.1 hold with  $\alpha > 1$ . Suppose, in addition, that (6.4), (7.2), and (7.3) hold. Then*

$$\lim_{u \rightarrow \infty} \frac{\psi(u)}{u P(|Z| > u)} = E \left[ w \left( \sup_{j \in \mathbb{Z}} \sum_{k=-\infty}^j A_{k,k} \right)^\alpha + (1-w) \left( \sup_{j \in \mathbb{Z}} \sum_{k=-\infty}^j -A_{k,k} \right)^\alpha \right] \frac{1}{c(\alpha-1)}. \tag{7.4}$$

**Example 7.1.** *(Independent and identically distributed.)* In the i.i.d. case  $A_{k,j} = \mathbf{1}\{j = 0\}$  we get the classical result (see, e.g. Embrechts *et al.* (1997, p. 43))

$$\lim_{u \rightarrow \infty} \frac{\psi(u)}{u P(|Z| > u)} = \frac{w}{c(\alpha-1)}.$$

**Example 7.2.** *(SRE.)* Consider the univariate SRE of Examples 1.2 and 2.2. Assume that the i.i.d. pairs  $(Y_k, Z_k)_{k \in \mathbb{Z}}$  are symmetric and that (7.2) holds. Set

$$M_+ = \sup_{j \geq 0} \left( \sum_{k=0}^j Y_1 \cdots Y_k \right)^+, \quad M_- = \sup_{j \geq 0} \left( \sum_{k=0}^j Y_1 \cdots Y_k \right)^-.$$

Then

$$\lim_{u \rightarrow \infty} \frac{\psi(u)}{u P(|Z| > u)} = (w E[M_+^\alpha] + (1-w) E[M_-^\alpha]) \frac{1}{c(\alpha-1)}.$$

This result is believed to be new.

*Proof of Theorem 7.1.* For  $q \geq 0$ , we define a counterpart to (7.1) by

$$X_k^q = \sum_{|j| \leq q} A_{k,j} Z_{k-j}, \quad k \in \mathbb{Z},$$

and let

$$S_n^q = X_1^q + \cdots + X_n^q, \quad \hat{S}_n^q = \hat{X}_1^q + \cdots + \hat{X}_n^q, \quad n = 1, 2, \dots$$

Let  $R$  denote the right-hand side of (7.4). The first step is to prove the upper bound

$$\limsup_{u \rightarrow \infty} \frac{\psi(u)}{u P(|Z| > u)} \leq R. \tag{7.5}$$

For (a large) integer  $M = 1, 2, \dots$ ,  $\psi(u)$  is bounded above by

$$P \left( \sup_{k \leq [u]M} (S_k - ck) > [u] \right) + P \left( \sup_{k > [u]M} (S_k - ck) > [u] \right) =: p_M^{(11)}(u) + p_M^{(12)}(u).$$

By Lemma 7.1,

$$\lim_{M \rightarrow \infty} \limsup_{u \rightarrow \infty} \frac{p_M^{(12)}(u)}{u P(|Z| > u)} = 0,$$

so the main contribution comes from  $p_M^{(11)}(u)$ . For any  $\varepsilon > 0$  and any integer  $q \geq 0$ , we have the upper bound

$$p_M^{(11)}(u) \leq P\left(\sup_{k \leq [u]M} (S_k^q - ck) > [u](1 - \varepsilon)\right) + P\left(\sup_{k \leq [u]M} \hat{S}_k^q > [u]\varepsilon\right).$$

It follows from Remark 6.2 and assumptions (6.4) and (7.2) that

$$\lim_{q \rightarrow \infty} \lim_{u \rightarrow 0} \frac{P(\sup_{k \leq [u]M} \hat{S}_k^q > [u]\varepsilon)}{u P(|Z| > u)} = 0.$$

It remains to show that

$$\lim_{M \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \limsup_{q \rightarrow \infty} \limsup_{u \rightarrow \infty} \frac{P(\sup_{k \leq [u]M} (S_k^q - ck) > [u](1 - \varepsilon))}{u P(|Z| > u)} \leq R.$$

Setting  $n = [u]M$  and taking  $0 < \gamma < 1$ , and a small  $\delta > 0$ , we see that

$$\begin{aligned} & P\left(\sup_{k \leq [u]M} (S_k^q - ck) > [u](1 - \varepsilon)\right) \\ &= P\left(\sup_{k \leq n} n^{-1}(S_k^q - ck) > (1 - \varepsilon)M^{-1}, \sup_{k \leq n} \left| \sum_{i=1}^k X_i^q \mathbf{1}\{|X_i^q| \leq n\delta\} \right| \leq n\gamma\right) \\ &+ P\left(\sup_{k \leq n} n^{-1}(S_k^q - ck) > (1 - \varepsilon)M^{-1}, \sup_{k \leq n} \left| \sum_{i=1}^k X_i^q \mathbf{1}\{|X_i^q| \leq n\delta\} \right| > n\gamma\right). \end{aligned} \tag{7.6}$$

Note that, by the regular variation and (7.3), for every  $M$  (recall that  $n = [u]M$ ) and  $0 < \gamma < 1$ ,

$$\lim_{\delta \rightarrow 0} \limsup_{u \rightarrow \infty} \frac{P(\sup_{k \leq [u]M} \left| \sum_{i=1}^k X_i^q \mathbf{1}\{|X_i^q| \leq n\delta\} \right| > n\gamma)}{u P(|Z| > u)} = 0. \tag{7.7}$$

Hence, we are left with estimating (7.6).

Since each noise variable  $Z$  affects at most  $2q + 1$  values of the process  $(X^q)$ , it follows from the obvious fact that, for every  $\delta > 0$ ,

$$\lim_{n \rightarrow \infty} \frac{P(|Z_j| > n\delta \text{ for at least two different } j = -q, \dots, n + q)}{n P(|Z| > n)} = 0,$$

and Remark 4.1 of Hult and Samorodnitsky (2008), that

$$P(D_n) := P(|X_{j_1}^q| > n\delta \text{ for } j_1, j_2 = 1, \dots, n, |j_1 - j_2| > 2q) = o(n P(|Z| > n)). \tag{7.8}$$

We conclude by (7.8) and (7.9) that, for the upper bound, we need to prove that

$$\lim_{M \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \limsup_{q \rightarrow \infty} \lim_{\gamma \rightarrow 0} \limsup_{\delta \rightarrow 0} \limsup_{u \rightarrow \infty} \frac{p_n(\delta)}{u P(|Z| > u)} \leq R. \tag{7.9}$$

Here  $p_n(\delta)$  is a modification of the probability in (7.6), defined as follows.

For  $n \geq 1$  and  $\delta > 0$ , let

$$K_\delta(n) = \inf\{i = 1, \dots, n : |X_i| > n\delta\},$$

defined to be equal to  $n + 1$  if the infimum is taken over the empty set. Then we set

$$p_n(\delta) = \mathbb{P}\left(\sup_{0 \leq k \leq 2q} n^{-1} \sum_{i=K_\delta(n)}^{K_\delta(n)+k} X_i - cK_\delta(n) > (1 - \varepsilon)M^{-1} - 2\gamma\right).$$

This puts us in a situation where we can use the large deviations for point processes in Theorem 4.1 and the mapping theorem in Lemma A.2.

Let  $q' = 6q + 1$ . This will correspond to the dimension of the point processes we will work with. Specifically,  $q'$  is the number of values of the process we are keeping track of in (4.2), and we will use the statement of Theorem 4.1 in the space  $\mathbb{M}_0(\mathbb{N}_p^{q'})$ . We now define a functional  $h^* : \mathbb{N}_p^{q'} \rightarrow \mathbb{R}$  as follows. Let

$$\xi = \sum_{k \in F} \delta_{(t_k, x_k^{(1)}, \dots, x_k^{(q')})} \in \mathbb{N}_p^{q'}.$$

Consider all points  $(t_k, x_k^{(1)}, \dots, x_k^{(q')})$  of  $\xi$  satisfying the following two conditions:

1. for some  $m = 2q + 1, \dots, q' - 2q, |x_k^{(m)}| > \delta$ ;
2.  $|x_k^{(j)}| \leq \delta$  for all  $j = 1, \dots, 2q$  and all  $j = q' - 2q + 1, \dots, q'$ .

Note that, by the definition of the space  $\mathbb{N}_p^{q'}$ , the set  $H_\delta(\xi)$  of such points is finite. If  $H_\delta(\xi) = \emptyset$ , we set  $h^*(\xi) = 0$ .

With the obvious convention for the expression  $k \in H_\delta(\xi)$ , we set, for each such  $k$ ,

$$m_k = \min\{m = 2q + 1, \dots, q' - 2q : |x_k^{(m)}| > \delta\},$$

and define

$$h(t_k, x_k^{(1)}, \dots, x_k^{(q')}) = \max_{j=0,1,\dots,2q} \left( \sum_{i=m_k}^{m_k+j} x_k^{(i)} \right) - ct_k.$$

Finally, we define

$$h^*(\xi) = \max_{k \in H_\delta(\xi)} h(t_k, x_k^{(1)}, \dots, x_k^{(q')}). \tag{7.10}$$

It follows from Lemma 7.2 that the measure  $m^{q'}$  in Theorem 4.1 assigns zero value to the set of discontinuities of  $h^*$ .

By the mapping theorem (Lemma A.2) we conclude that, for any  $\tau > 0$ ,

$$r_n \mathbb{P}(h^*(N_n^{q'}) > \tau) = m_n^{q'}(h^*(N_n^{q'}) > \tau) \rightarrow m^{q'}\{\xi : h^*(\xi) > \tau\}, \tag{7.11}$$

using the fact that the right-hand side of (7.11) is continuous in  $\tau > 0$ .

Now taking into account the definition of  $r_n$ , the estimate (7.8), and the fact that  $n = [u]M$ , we obtain, from (7.11),

$$\limsup_{u \rightarrow \infty} \frac{p_n(\delta)}{u \mathbb{P}(|Z| > u)} \leq M^{-(\alpha-1)} m^{q'}\{\xi : h^*(\xi) > (1 - \varepsilon)M^{-1} - 2\gamma\}.$$

It follows from the form of the limiting measure  $m^{q'}$  in the one-dimensional case (see (2.6)) that, for any  $\tau > 0$ ,

$$\begin{aligned} &\lim_{\delta \rightarrow 0} m^{q'} \{ \xi : h^*(\xi) > \tau \} \\ &= \frac{1}{c(\alpha - 1)} (\tau^{-(\alpha-1)} - (c + \tau)^{-(\alpha-1)}) \\ &\quad \times \mathbb{E} \left[ w \left( \max_{j=0,1,\dots,2q} \sum_{k=-q}^{-q+j} A_{k,k} \right)_+^\alpha + (1-w) \left( \max_{j=0,1,\dots,2q} \sum_{k=-q}^{-q+j} -A_{k,k} \right)_+^\alpha \right], \end{aligned}$$

from which we see that

$$\begin{aligned} &\limsup_{q \rightarrow \infty} \lim_{\gamma \rightarrow 0} \limsup_{\delta \rightarrow 0} \limsup_{u \rightarrow \infty} \frac{p_n(\delta)}{u \mathbb{P}(|Z| > u)} \\ &\leq M^{-(\alpha-1)} [((1 - \varepsilon)M)^{\alpha-1} - (c + ((1 - \varepsilon)M^{-1}))^{-(\alpha-1)}] R, \end{aligned}$$

from which (7.9) follows. This proves the upper bound (7.5).

The lower bound requires a similar estimate. Take  $\varepsilon > 0$ , and let  $u$  be sufficiently large that  $([u] + 1)/[u] < 1 + \varepsilon$ . For (a large) integer  $M = 1, 2, \dots$ , we have

$$\begin{aligned} \psi(u) &\geq \mathbb{P} \left( \sup_k (S_k - ck) > [u] + 1 \right) \\ &\geq \mathbb{P} \left( \sup_{k \leq [u]M} (S_k - ck) > [u](1 + \varepsilon) \right) \\ &\geq \mathbb{P} \left( \sup_{k \leq [u]M} (S_k^q - ck) > [u](1 + 2\varepsilon) \right) - \mathbb{P} \left( \sup_{k \leq [u]M} \hat{S}_k^q > [u]\varepsilon \right). \end{aligned}$$

Hence, by Remark 6.2 and assumptions (6.4) and (7.2), it is sufficient to prove that

$$\lim_{M \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \liminf_{q \rightarrow \infty} \liminf_{u \rightarrow \infty} \frac{\mathbb{P}(\sup_{k \leq [u]M} (S_k^q - ck) > [u](1 + 2\varepsilon))}{u \mathbb{P}(|Z| > u)} \geq R.$$

Using (7.7) again, it is sufficient to consider

$$\frac{\mathbb{P}(\sup_{k \leq [u]M} n^{-1} (\sum_{i=1}^k X_i \mathbf{1}\{n\delta < |X_i| < n/\delta\} - ck) > (1 + 2\varepsilon)/M + \gamma)}{u \mathbb{P}(|Z| > u)},$$

and the argument from here is the same as in the case of the upper bound.

Below are the lemmas used in the proof of Theorem 7.1.

**Lemma 7.1.** *Under the assumptions of Theorem 7.1,*

$$\lim_{M \rightarrow \infty} \limsup_{u \rightarrow \infty} \frac{\mathbb{P}(\sup_{k > uM} (S_k - ck) > u)}{u \mathbb{P}(|Z| > u)} = 0.$$

*Proof.* We use Theorem 6.1 and Remarks 6.1 and 6.2 with  $\gamma_n \equiv n$ . Choose  $\beta > 1$  and  $A > 1$  such that

$$\mathbb{E} |X_1| < \frac{c(1 - 1/A)}{\beta - 1}, \tag{7.12}$$

and write

$$P\left(\sup_{k > uM} (S_k - ck) > u\right) \leq \sum_{j=1}^{\infty} P(S_k > ck \text{ for some } Mu\beta^{j-1} \leq k \leq Mu\beta^j).$$

By stationarity of  $(X_k)$ , for every  $j = 1, 2, \dots$ ,

$$\begin{aligned} &P(S_k > ck \text{ for some } Mu\beta^{j-1} \leq k \leq Mu\beta^j) \\ &\leq P\left(S_{\lceil Mu\beta^{j-1} \rceil} > \frac{cMu\beta^{j-1}}{A}\right) \\ &\quad + P\left(S_k > ck + cMu\beta^{j-1}\left(1 - \frac{1}{A}\right) \text{ for some } 0 \leq k \leq Mu(\beta^j - \beta^{j-1})\right). \end{aligned}$$

Using Theorem 6.1, we see that, for some positive constant  $C$  (that, as usual, may change in the sequel), we have, for large enough  $u$ ,

$$P\left(S_{\lceil Mu\beta^{j-1} \rceil} > \frac{cMu\beta^{j-1}}{A}\right) \leq CMu\beta^{j-1} P(|Z| > Mu\beta^{j-1}),$$

and, by Potter’s bound, for large enough  $M$ ,

$$\frac{P(|Z| > Mu\beta^{j-1})}{P(|Z| > u)} \leq C(M\beta^{j-1})^{-\alpha}.$$

It follows that

$$\limsup_{u \rightarrow \infty} \frac{\sum_{j=1}^{\infty} P(S_{\lceil Mu\beta^{j-1} \rceil} > cMu\beta^{j-1}/A)}{u P(|Z| > u)} \leq CM^{-(\alpha-1)}.$$

Using the fact that  $\alpha > 1$ , we let  $M \rightarrow \infty$  and see that the above expression converges to 0.

Furthermore, for every  $j = 1, 2, \dots$ ,

$$\begin{aligned} &P\left(S_k > ck + cMu\beta^{j-1}\left(1 - \frac{1}{A}\right) \text{ for some } 0 \leq k \leq Mu(\beta^j - \beta^{j-1})\right) \\ &\leq P\left(\sum_{k=0}^{Mu(\beta^j - \beta^{j-1})} |X_k| > cMu\beta^{j-1}\left(1 - \frac{1}{A}\right)\right) \\ &\leq P\left(\frac{1}{Mu(\beta^j - \beta^{j-1})} \sum_{k=0}^{Mu(\beta^j - \beta^{j-1})} (|X_k| - E|X_1|) > \frac{c(1 - 1/A)}{\beta - 1} - E|X_1|\right). \end{aligned}$$

By the choice of  $\beta$  and  $A$  as in (7.12) and assumption (7.2), we can use the large deviation result (6.7) to conclude that, just as above, for all large enough  $M$ ,

$$\begin{aligned} &\limsup_{u \rightarrow \infty} \frac{P(S_k > ck + cMu\beta^{j-1}(1 - 1/A) \text{ for some } 0 \leq k \leq Mu(\beta^j - \beta^{j-1}))}{u P(|Z| > u)} \\ &\leq C(M\beta^{j-1})^{-(\alpha-1)}, \end{aligned}$$

and, as before, these bounds can be summed up over  $j$  and, then, we let  $M \rightarrow \infty$  and use the fact that  $\alpha > 1$ . This proves the statement of the lemma.

**Lemma 7.2.** *Under the assumptions of Theorem 7.1, the measure  $m^{q'}$  in Theorem 4.1 does not charge the set of discontinuities of  $h^*$  in (7.10).*

*Proof.* Let  $\Xi$  be the subset of  $\mathbb{N}_p^{q'}$  consisting of point measures  $\xi$  such that

$$\xi(\{[0, 1] \times \{(x_0, \dots, x_{q'}) : |x_i| = \delta, \text{ some } i \in \{0, \dots, q'\}\}) = 0.$$

According to Remark 4.1, the measure  $m^{q'}$  is concentrated on the set  $\Xi$ , and so it is enough to prove that the functional  $h^*$  is continuous at each  $\xi \in \Xi$ . Let  $(\xi_n)$  be a sequence in  $\mathbb{N}_p^{q'}$  such that  $\xi_n \xrightarrow{v} \xi$ . If  $H_\delta(\xi) = \emptyset$  then  $H_\delta(\xi_n) = \emptyset$  for all large enough  $n$ , and so  $h^*(\xi_n) = 0 \rightarrow 0 = h^*(\xi)$ .

Suppose now that  $H_\delta(\xi) \neq \emptyset$ . By the definition of the set  $\Xi$  we see that, for all large enough  $n$  (say  $n \geq n_0$ ), the cardinality of  $H_\delta(\xi_n)$  is equal to the (finite) cardinality of  $H_\delta(\xi)$ . Moreover, the vague convergence  $\xi_n \xrightarrow{v} \xi$  implies that, for every  $n \geq n_0$ , there is an enumeration  $\{((t_k)^n, (x_k^{(1)})^n, \dots, (x_k^{(q')})^n)\}$  of  $H_\delta(\xi_n)$  such that, for every  $k \in H_\delta(\xi)$ ,

$$((t_k)^n, (x_k^{(1)})^n, \dots, (x_k^{(q')})^n) \rightarrow (t_k, x_k^{(1)}, \dots, x_k^{(q')})$$

componentwise as  $n \rightarrow \infty$  (see Resnick (1987, Proposition 3.13)). Therefore, for each such  $k$ ,

$$h((t_k)^n, (x_k^{(1)})^n, \dots, (x_k^{(q')})^n) \rightarrow h(t_k, x_k^{(1)}, \dots, x_k^{(q')}),$$

and, since the set  $H_\delta(\xi)$  is finite, we conclude that  $h^*(\xi_n) \rightarrow h^*(\xi)$ , as required.

Finally, as promised, we provide sufficient conditions for (7.3).

**Lemma 7.3.** *Assume that the hypothesis of Theorem 4.1 holds. If  $1 < \alpha \leq 2$  then, for every  $q \geq 0$  and  $\gamma > 0$ ,*

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{\mathbb{P}(\sup_{k \leq n} |\sum_{i=1}^k X_i^q \mathbf{1}\{|X_i^q| \leq n\delta}\}| > n\gamma)}{n \mathbb{P}(|Z| > n)} = 0.$$

*If  $\alpha > 2$ , the conclusion holds if, additionally, for some  $\beta > \alpha - 1$  and all  $-q \leq j \leq q$ ,*

$$\mathbb{E} A_{0,j}^{2\beta} < \infty. \tag{7.13}$$

*Proof.* Write

$$\begin{aligned} & \mathbb{P}\left(\sup_{k \leq n} \left| \sum_{i=1}^k X_i^q \mathbf{1}\{|X_i^q| \leq n\delta}\right| > n\gamma\right) \\ & \leq \sum_{|j| \leq q} \mathbb{P}\left(\sup_{k \leq n} \left| \sum_{i=1}^k A_{i,j} Z_{i-j} \mathbf{1}\{|X_i^q| \leq n\delta}\right| > \frac{n\gamma}{2q+1}\right). \end{aligned}$$

We replace below, for simplicity,  $\gamma/(2q + 1)$  with  $\gamma$ . Since the above sum has a finite number of terms, it is enough to prove the appropriate convergence to 0 for each one of the terms separately. For simplicity, we consider  $j = 0$ . Define

$$B_n = \left\{ \sup_{k \leq n} \left| \sum_{i=1}^k A_{i,0} Z_i \mathbf{1}\{|X_i^q| \leq n\delta}\right| > n\gamma \right\},$$

so that we can write, for a small  $\rho > 0$ ,

$$\begin{aligned} & \mathbb{P}\left(\sup_{k \leq n} \left| \sum_{i=1}^k A_{i,0} Z_i \mathbf{1}\{|X_i^q| \leq n\delta\} \right| > n\gamma\right) \\ &= \mathbb{P}(B_n \cap \{|Z_m| \leq n\rho \text{ for all } m = 1 - q, \dots, n + q\}) \\ & \quad + \mathbb{P}(B_n \cap \{|Z_m| > n\rho \text{ for exactly one } m = 1 - q, \dots, n + q\}) \\ & \quad + \mathbb{P}(B_n \cap \{|Z_m| > n\rho \text{ for two or more } m = 1 - q, \dots, n + q\}) \\ &:= p_1(n) + p_2(n) + p_3(n). \end{aligned}$$

Clearly, for every  $\rho > 0$ ,

$$\lim_{n \rightarrow \infty} \frac{p_3(n)}{n \mathbb{P}(|Z| > n)} = 0.$$

Next, select  $0 < \theta < \delta/(2q + 1)$ , and introduce the event

$$C_n = \{|A_{i_1,j} Z_{i_1-j}| > n\theta \text{ for some } i_1 = 1, \dots, n, |j| \leq q\}.$$

Then

$$\begin{aligned} p_1(n) &\leq \mathbb{P}(C_n \cap \{|Z_m| \leq n\rho \text{ for all } m = 1 - q, \dots, n + q\}) \\ & \quad + \mathbb{P}(B_n \cap C_n^c \cap \{|Z_m| \leq n\rho \text{ for all } m = 1 - q, \dots, n + q\}) \\ &:= p_{11}(n) + p_{12}(n). \end{aligned}$$

By stationarity,

$$\begin{aligned} p_{11}(n) &\leq (n + 2q) \mathbb{P}\left(|Z_0| \leq n\rho, \max_{|j| \leq q} |A_{j,j} Z_0| > n\theta\right) \\ &\leq (n + 2q) \mathbb{P}\left(\max_{|j| \leq q} |A_{j,j}| \mathbf{1}\left\{\max_{|j| \leq q} |A_{j,j}| > \frac{\theta}{\rho}\right\} |Z_0| > n\theta\right). \end{aligned}$$

Therefore,

$$\lim_{n \rightarrow \infty} \frac{p_{11}(n)}{n \mathbb{P}(|Z| > n)} = \theta^{-\alpha} \mathbb{E}\left[\max_{|j| \leq q} |A_{j,j}|^\alpha \mathbf{1}\left\{\max_{|j| \leq q} |A_{j,j}| > \frac{\theta}{\rho}\right\}\right],$$

and this expression can be made arbitrarily small by selecting  $\rho$  small in comparison with  $\theta$ . Furthermore, the choice of  $\theta$  guarantees that, on the event  $C_n^c$ , we automatically have

$$|X_i^q| \leq n\delta \quad \text{for each } i = 1, \dots, n.$$

Therefore,

$$p_{12}(n) \leq \mathbb{P}\left(\sup_{k \leq n} \left| \sum_{i=1}^k A_{i,0} Z_i \mathbf{1}\{|Z_i| \leq n\rho\} \right| > n\gamma\right).$$

Set

$$m_{n\rho} = \mathbb{E} Z \mathbf{1}\{|Z| \leq n\rho\} \quad \text{and} \quad \tilde{S}_k = \sum_{i=1}^k A_{i,0} (Z_i \mathbf{1}\{|Z_i| \leq n\rho\} - m_{n\rho}),$$

and take  $p > \alpha$  such that  $E |A_{0,0}|^p < \infty$ . Then

$$\begin{aligned} & \mathbb{P}\left(\max_{k \leq n} \left| \sum_{i=1}^k A_{i,0} Z_i \mathbf{1}\{|Z_i| \leq n\rho\} \right| > n\gamma\right) \\ & \leq \mathbb{P}\left(\max_{k \leq n} |\tilde{S}_k| > \frac{n\gamma}{2}\right) + \mathbb{P}\left(\max_{k \leq n} \left| \sum_{i=1}^k A_{i,0} m_{n\rho} \right| > \frac{n\gamma}{2}\right). \end{aligned} \tag{7.14}$$

By Markov’s inequality, the second term in (7.14) is bounded above by

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^n |A_{i,0}| > \frac{n\gamma}{2|m_{n\rho}|}\right) & \leq \left(\frac{n\gamma}{2|m_{n\rho}|}\right)^{-p} E\left(\sum_{i=1}^n |A_{i,0}|^p\right) \\ & \leq \left(\frac{n\gamma}{2|m_{n\rho}|}\right)^{-p} n^{p-1} E \sum_{i=1}^n |A_{i,0}|^p \\ & = \left(\frac{\gamma}{2|m_{n\rho}|}\right)^{-p} E |A_{0,0}|^p. \end{aligned}$$

Since  $p > \alpha > 1$ ,  $E |A_{0,0}|^p < \infty$ , and  $|m_{n\rho}| \sim Cn\rho \mathbb{P}(|Z| > n\rho)$ , it follows that

$$\limsup_{n \rightarrow \infty} \frac{\mathbb{P}(\sup_{k \leq n} |\sum_{i=1}^k A_{i,0} m_{n\rho}| > n\gamma/2)}{n \mathbb{P}(|Z| > n)} \leq \limsup_{n \rightarrow \infty} C \frac{[n\rho \mathbb{P}(|Z| > n\rho)]^p}{n \mathbb{P}(|Z| > n)} = 0.$$

To handle the first term in (7.14), we divide it into two cases. For  $1 < \alpha < 2$ , we can take  $\alpha < p < 2$  and use the fact that  $\tilde{S}_k$  is a martingale with respect to  $\mathcal{F}_k = \sigma(\{A_{i,0}\}_{i=1}^n, Z_1, \dots, Z_k)$ . Then the Burkholder–Davis–Gundy inequality implies that

$$\begin{aligned} \mathbb{P}\left(\max_{k \leq n} |\tilde{S}_k| > \frac{n\gamma}{2}\right) & \leq \frac{C}{(n\gamma)^p} E([\tilde{S}]_n^{p/2}) \\ & \leq \frac{C}{(n\gamma)^p} E\left(\sum_{i=1}^n |A_{i,0}|^p |Z_i| \mathbf{1}\{|Z_i| \leq n\rho\} - m_{n\rho}|^p\right) \\ & \leq \frac{C}{(n\gamma)^p} n E |A_{0,0}|^p (E |Z|^p \mathbf{1}\{|Z| \leq n\rho\} + |m_{n\rho}|^p) \\ & \sim \frac{C}{(n\gamma)^p} n (E |A_{0,0}|^p (n\rho)^p \mathbb{P}(|Z| > n\rho) + |m_{n\rho}|^p), \end{aligned} \tag{7.15}$$

where, in the last step, we used Karamata’s theorem. In particular,

$$\lim_{\rho \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{(C/(n\gamma)^p)n(E |A_{0,0}|^p (n\rho)^p \mathbb{P}(|Z| > n\rho) + |m_{n\rho}|^p)}{n \mathbb{P}(|Z| > n)} = 0.$$

For  $\alpha \geq 2$ , a variant of the Fuk–Nagaev inequality (see Petrov (1995, Section 2.6.6, p. 79))

implies that

$$\begin{aligned} \mathbb{P}\left(\max_{k \leq n} |\tilde{S}_k| > \frac{n\gamma}{2}\right) &= \mathbb{E}\left[\mathbb{P}\left(\max_{k \leq n} |\tilde{S}_k| > \frac{n\gamma}{2} \mid \{A_{i,0}\}\right)\right] \\ &\leq \mathbb{E}\left[C_1(n\gamma)^{-p} \sum_{i=1}^n |A_{i,0}|^p \mathbb{E}|Z_i \mathbf{1}\{|Z_i| \leq n\rho\} - m_{n\rho}|^p\right] \\ &\quad + \mathbb{E} \exp\left\{-C_2 n^2 \left(\sum_{i=1}^n A_{i,0}^2 \text{var}(Z \mathbf{1}\{|Z| \leq n\rho\})\right)^{-1}\right\}. \end{aligned}$$

The first of these terms can be bounded just as (7.15). To handle the second term, we write  $W_n := \sum_{i=1}^n A_{i,0}^2$  and note that, since  $\alpha \geq 2$ ,  $\text{var}(Z \mathbf{1}\{|Z| \leq n\rho\})$  is a slowly varying function (this quantity is even bounded when  $\text{var}(Z) < \infty$ ). Therefore, it is bounded by  $n^\varepsilon$  for all sufficiently large  $n$ , where we choose  $\varepsilon$  to satisfy  $\beta > (\alpha - 1)/(1 - \varepsilon)$ . Then it follows that, for each  $\lambda > 0$ ,

$$\begin{aligned} &\mathbb{E}[\exp\{-C_2 n^2 (W_n \text{var}(Z \mathbf{1}\{|Z| \leq n\rho\}))^{-1}\}] \\ &\leq \mathbb{E}\left[\exp\left\{-C_3 \frac{n^{2-\varepsilon}}{W_n}\right\}\right] \\ &= \mathbb{E}\left[\exp\left\{-C_3 \frac{n^{2-\varepsilon}}{W_n}\right\} \mathbf{1}\left\{W_n \leq \frac{\lambda n^{2-\varepsilon}}{\log n}\right\}\right] + \mathbb{E}\left[\exp\left\{-C_3 \frac{n^{2-\varepsilon}}{W_n}\right\} \mathbf{1}\left\{W_n > \frac{\lambda n^{2-\varepsilon}}{\log n}\right\}\right] \\ &\leq n^{-C_3 \lambda} + \mathbb{P}\left(W_n > \frac{\lambda n^{2-\varepsilon}}{\log n}\right). \end{aligned}$$

In particular, we may choose  $\lambda > (\alpha - 1)/C_3$ , which will imply that

$$\frac{n^{-C_3 \lambda}}{n \mathbb{P}(|Z| > n)} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

We also have, for large  $n$ , by the choice of  $\varepsilon$ ,

$$\begin{aligned} \frac{\mathbb{P}(W_n > \lambda n^{2-\varepsilon} / \log n)}{n \mathbb{P}(|Z| > n)} &\leq \frac{n^{-\beta(2-\varepsilon)}}{n \mathbb{P}(|Z| > n)} \mathbb{E}\left(\sum_{i=1}^n A_{i,0}^2\right)^\beta \\ &\leq \frac{n^{-\beta(1-\varepsilon)}}{n \mathbb{P}(|Z| > n)} \mathbb{E} A_{0,0}^{2\beta} \\ &\rightarrow 0, \end{aligned}$$

by assumption (7.13).

Finally, the term  $p_2(n)$  can be treated in the same way as the term  $p_1(n)$ , if we note that the single large value of  $Z_m$  can contribute to at most  $2q + 1$  different  $X_i$ . If we choose  $\delta$  small enough so that  $(2q + 1)\delta < \gamma$ , then these terms can be excluded from the sum  $\sum_{i=1}^k X_i^q \mathbf{1}\{|X_i^q| \leq n\delta\}$  in the first place. Hence, the statement of the lemma follows.

### Appendix A. Framework

Let  $\mathbb{E}$  be a locally compact complete separable metric space and consider the space  $\mathbb{N}_p$  of Radon point measures on  $\mathbb{E}$ . In the main part of the paper  $\mathbb{E}$  is the space  $[0, 1] \times (\mathbb{R}^{d(q+1)} \setminus \{0\})$

for some  $q \geq 0$  and  $d \geq 1$ , but here it can be quite arbitrary. Let  $(h_i)_{i \geq 1}$  be a countable dense collection of functions in  $C_K^+(\mathbb{E})$ , the space of nonnegative continuous functions on  $\mathbb{E}$  with compact support, such that  $\xi_n(h_i) \rightarrow \xi(h_i)$  as  $n \rightarrow \infty$  for each  $i \geq 1$  implies that  $\xi_n \xrightarrow{v} \xi$  in  $\mathbb{N}_p$ . Here ‘ $\xrightarrow{v}$ ’ denotes vague convergence. The existence of such a sequence  $(h_i)_{i \geq 1}$  is established in Kallenberg (1983, Section 15.7.7) (see also Resnick (1987, Proposition 3.17)). Note also that the functions  $h_i$  may be chosen to be Lipschitz with respect to the metric on  $\mathbb{E}$ . This follows from the fact that the approximating functions in the version of the Urysohn lemma used for the purpose of this construction are already Lipschitz (see Resnick (1987, Lemma 3.11)). In particular, a measure  $\xi$  in  $\mathbb{N}_p$  is uniquely determined by the sequence  $(\xi(h_i))_{i \geq 1}$ . We may and will assume that the collection  $(h_i)_{i \geq 1}$  is closed under multiplication by positive rational numbers.

We can identify  $\mathbb{N}_p$  with a closed subspace of  $[0, \infty)^\infty$  via the mapping  $h: \mathbb{N}_p \rightarrow [0, \infty)^\infty$  given by  $h(\xi) = (\xi(h_i))_{i \geq 1}$ . To see that  $h(\mathbb{N}_p)$  is closed in  $[0, \infty)^\infty$ , let  $(x_i^n)_{i \geq 1}$  be a convergent sequence in  $h(\mathbb{N}_p)$ . That is,  $x_i^n \rightarrow x_i$  for each  $i$ . Then there exist  $\xi_n \in \mathbb{N}_p$  such that  $\xi_n(h_i) = x_i^n$  for each  $i \geq 1$ . The collection  $(\xi_n)_{n \geq 1}$  is relatively compact in  $\mathbb{N}_p$  because

$$\sup_n \xi_n(h_i) = \sup_n x_i^n < \infty \quad \text{for each } i.$$

Hence, there is a convergent subsequence  $\xi_{n_k} \rightarrow \xi$ . This  $\xi$  necessarily satisfies  $\xi(h_i) = x_i$  and we conclude that  $(x_i)_{i \geq 1} \in h(\mathbb{N}_p)$ . Thus,  $h(\mathbb{N}_p)$  is closed.

The vague convergence on  $\mathbb{N}_p$  can be metrized via a metric  $d$  induced from  $[0, \infty)^\infty$ , defined by

$$d(x, y) = \sum_{i=1}^{\infty} 2^{-i} \frac{|x_i - y_i|}{1 + |x_i - y_i|} \tag{A.1}$$

for elements  $x = (x_i)_{i \geq 1}$  and  $y = (y_i)_{i \geq 1}$  in  $[0, \infty)^\infty$ . This makes  $\mathbb{N}_p$  into a complete separable metric space (since it is a closed subspace of the complete separable metric space  $[0, \infty)^\infty$ ). The open ball of radius  $r > 0$  in  $\mathbb{N}_p$  centered at  $\xi$  is denoted by  $B_{\xi,r}$ . Recall that we denote by  $\emptyset$  the null measure in  $\mathbb{N}_p$ .

We will consider convergence of Radon measures  $m$  on the space  $\mathbb{N}_p$ . The framework considered here is that of Hult and Lindskog (2006) where the underlying space, denoted by  $\mathbf{S}$  in Hult and Lindskog (2006), is taken to be  $\mathbb{N}_p$ . The space of Radon measures on  $\mathbb{N}_p$  whose restriction to  $\mathbb{N}_p \setminus B_{\emptyset,r}$  is finite for each  $r > 0$  is denoted by  $\mathbb{M}_0 = \mathbb{M}_0(\mathbb{N}_p)$ . The convergence  $m_n \rightarrow m$  in  $\mathbb{M}_0$  is defined as the convergence  $m_n(f) \rightarrow m(f)$  for all  $f \in C_0(\mathbb{N}_p)$ , the space of bounded continuous functions on  $\mathbb{N}_p$  that vanish in a neighborhood of ‘the origin’  $\emptyset$ .

The typical situation in this paper is that we have a sequence of random point measures  $(N_n)$  on  $\mathbb{E}$ , and we are interested in the convergence

$$m_n(\cdot) := r_n P(N_n \in \cdot) \rightarrow m(\cdot) \quad \text{in } \mathbb{M}_0.$$

**A.1. Convergence in  $\mathbb{M}_0(\mathbb{N}_p)$**

We start with relative compactness criteria. For measures on a general metric space, such criteria are given in Theorem 2.7 of Hult and Lindskog (2006).

**Theorem A.1.** *Let  $M \subset \mathbb{M}_0(\mathbb{N}_p)$ . Then  $M$  is relatively compact if*

(i) *for each  $\varepsilon > 0$ ,*

$$\sup_{m \in M} m \left( \xi : \sum_{i=1}^{\infty} \frac{2^{-i} \xi(h_i)}{1 + \xi(h_i)} > \varepsilon \right) < \infty,$$

and

(ii) for each  $h \in C_K^+(\mathbb{E})$  and  $\delta > 0$ , there exists  $R$  such that

$$\sup_{m \in M} m(\xi : \xi(h) > R) \leq \delta.$$

*Proof.* We need to check (2.2) and (2.3) in Theorem 2.7 of Hult and Lindskog (2006). Since the metric on  $\mathbb{N}_p$  is given by (A.1), (i) immediately implies (2.2) in that reference.

Next note that any set of the form  $\prod_{i=1}^\infty [0, R_i]$  is a compact subset of  $[0, \infty)^\infty$ . Hence,

$$C = \{\xi : \xi(h_i) \leq R_i \text{ for each } i\} \setminus B_{\emptyset, \varepsilon}$$

is a compact subset of  $\mathbb{N}_p \setminus B_{\emptyset, \varepsilon}$  and

$$\begin{aligned} \sup_{m \in M} m(\mathbb{N}_p \setminus (B_{\emptyset, \varepsilon} \cup C)) &\leq \sup_{m \in M} m(\xi : \xi(h_i) > R_i \text{ some } i \geq 1) \\ &\leq \sup_{m \in M} \sum_{i=1}^\infty m(\xi : \xi(h_i) > R_i). \end{aligned}$$

By (ii) we can take  $R_i$  such that

$$\sup_{m \in M} m(\xi : \xi(h_i) > R_i) < 2^{-i} \delta,$$

which implies (2.3) of Hult and Lindskog (2006).

To show actual convergence, we need, in addition to relative compactness, to identify subsequential limits. For this purpose, we define, for  $g_1, g_2 \in C_K^+(\mathbb{E})$  and  $\varepsilon_1, \varepsilon_2 > 0$ , a function  $F_{g_1, g_2, \varepsilon_1, \varepsilon_2} : \mathbb{N}_p \rightarrow [0, \infty)$  by

$$F_{g_1, g_2, \varepsilon_1, \varepsilon_2}(\xi) = (1 - \exp\{-(\xi(g_1) - \varepsilon_1)_+\})(1 - \exp\{-(\xi(g_2) - \varepsilon_2)_+\}). \tag{A.2}$$

Note that each  $F_{g_1, g_2, \varepsilon_1, \varepsilon_2}$  is a bounded continuous function that vanishes on a neighborhood of the null measure  $\emptyset$ .

**Lemma A.1.** *Let  $m_1$  and  $m_2$  be measures in  $\mathbb{M}_0(\mathbb{N}_p)$ . If the equality*

$$m_1(F_{g_1, g_2, \varepsilon_1, \varepsilon_2}) = m_2(F_{g_1, g_2, \varepsilon_1, \varepsilon_2})$$

*holds for all Lipschitz functions  $g_1, g_2 \in C_K^+(\mathbb{E})$  and  $\varepsilon_1, \varepsilon_2 > 0$ , then  $m_1 = m_2$ .*

*Proof.* We use the assumption with  $g_i = h_{j_i}$ ,  $i = 1, 2$ . Replacing  $h_{j_i}$  by  $bh_{j_i}$  and  $\varepsilon_1$  by  $b\varepsilon_1$  with positive rational  $b$ , and letting  $b \rightarrow \infty$  and  $\varepsilon_2 \rightarrow 0$ , we obtain

$$\int_{\mathbb{N}_p} \mathbf{1}\{\xi(h_{j_1}) \geq \varepsilon_1\} e^{-\xi(h_{j_2})} m_1(d\xi) = \int_{\mathbb{N}_p} \mathbf{1}\{\xi(h_{j_1}) \geq \varepsilon_1\} e^{-\xi(h_{j_2})} m_2(d\xi). \tag{A.3}$$

Replacing, in (A.3),  $h_{j_2}$  by  $bh_{j_2}$  as above, and letting  $b \rightarrow 0$ , we also obtain

$$m_1(\mathbf{1}\{\xi(h_{j_1}) \geq \varepsilon_1\}) = m_2(\mathbf{1}\{\xi(h_{j_1}) \geq \varepsilon_1\}). \tag{A.4}$$

Since the family  $(h_i)_{i \geq 1}$  is dense in  $C_K^+(\mathbb{E})$ , we conclude that (A.3) holds with  $h_{j_2}$  replaced by any function in  $C_K^+(\mathbb{E})$ . To see that (A.3) and (A.4) imply that  $m_1 = m_2$ , we define, for any  $j_1 \geq 1$  and  $\varepsilon_1 > 0$ , probability measures on  $\mathbb{N}_p$  by

$$\begin{aligned} \tilde{m}_1(\cdot) &= \frac{m_1(\cdot \cap \{\xi : \xi(h_{j_1}) \geq \varepsilon_1\})}{m_1(\xi : \xi(h_{j_1}) \geq \varepsilon_1)}, \\ \tilde{m}_2(\cdot) &= \frac{m_2(\cdot \cap \{\xi : \xi(h_{j_1}) \geq \varepsilon_1\})}{m_2(\xi : \xi(h_{j_1}) \geq \varepsilon_1)}. \end{aligned}$$

The uniqueness property of the Laplace functionals (see Resnick (1987, Section 3.2)) (A.3) implies that  $\tilde{m}_1$  and  $\tilde{m}_2$  coincide. Hence,  $m_1$  and  $m_2$  coincide on the set  $\{\xi(h_j) \geq \varepsilon\}$  for any  $j$  and  $\varepsilon$ . Letting  $\varepsilon \rightarrow 0$  we obtain the claim.

Finally, we are ready to state necessary and sufficient conditions for convergence in  $\mathbb{M}_0(\mathbb{N}_p)$ .

**Theorem A.2.** *Let  $m, m_1, m_2, \dots$  be measures in  $\mathbb{M}_0(\mathbb{N}_p)$ . The condition*

$$\lim_{n \rightarrow \infty} m_n(F_{g_1, g_2, \varepsilon_1, \varepsilon_2}) = m(F_{g_1, g_2, \varepsilon_1, \varepsilon_2})$$

for all  $g_1, g_2 \in C_K^+(\mathbb{E})$  and  $\varepsilon_1, \varepsilon_2 > 0$  is necessary and sufficient for the convergence  $m_n \rightarrow m$  in  $\mathbb{M}_0(\mathbb{N}_p)$ . Furthermore, it is sufficient to check the condition only for the Lipschitz functions in  $C_K^+(\mathbb{E})$ .

*Proof.* The necessity of the condition is obvious. For the sufficiency, we start by checking that the sequence  $(m_n)_{n \geq 1}$  is relatively compact in  $\mathbb{M}_0(\mathbb{N}_p)$ , for which we check (i) and (ii) in Theorem A.1.

Start by choosing a Lipschitz collection  $(h_i)_{i \geq 1}$  as above. Take  $\varepsilon > 0$ . With

$$J_\varepsilon = \lceil -\log_2 \varepsilon \rceil + 1$$

we have

$$\begin{aligned} m_n \left( \left\{ \xi : \sum_{i=1}^{\infty} 2^{-i} \frac{\xi(h_i)}{1 + \xi(h_i)} \geq \varepsilon \right\} \right) &\leq m_n \left( \left\{ \xi : \sum_{i=1}^{J_\varepsilon} 2^{-i} \frac{\xi(h_i)}{1 + \xi(h_i)} > \frac{\varepsilon}{3} \right\} \right) \\ &\leq \sum_{i=1}^{J_\varepsilon} m_n \left( \left\{ \xi : \frac{\xi(h_i)}{1 + \xi(h_i)} > \frac{\varepsilon}{3J_\varepsilon} \right\} \right) \\ &\leq \sum_{i=1}^{J_\varepsilon} m_n \left( \left\{ \xi : \xi(h_i) > \frac{1}{(3J_\varepsilon/\varepsilon - 1)} \right\} \right). \end{aligned}$$

Note that, for any  $h \in C_K^+(\mathbb{E})$  and  $R > 0$ ,

$$\begin{aligned} m(F_{h, h, R/2, R/2}) &= \int (1 - e^{-(\xi(h) - R/2)_+})^2 m(d\xi) \tag{A.5} \\ &\geq \int (1 - e^{-(\xi(h) - R/2)_+})^2 \mathbf{1}\{\xi(h) > R\} m(d\xi) \\ &\geq (1 - e^{-R/2})^2 m(\xi : \xi(h) > R). \end{aligned}$$

For  $\varepsilon > 0$ , we choose  $R = R(\varepsilon) = 2/(3J_\varepsilon/\varepsilon - 1)$ . By the assumption of the proposition, there is an  $n_1$  such that, for all  $n \geq n_1$ , the bound

$$m_n(F_{h_i, h_i, R/2, R/2}) \leq m(F_{h_i, h_i, R/2, R/2}) + 1$$

holds for each  $i = 1, \dots, J_\varepsilon$ . It follows from (A.5) that, for all such  $n$ ,

$$m_n\left(\left\{\xi \in \mathbb{N}_p : \sum_{i=1}^{\infty} 2^{-i} \frac{\xi(h_i)}{1 + \xi(h_i)} \geq \varepsilon\right\}\right) \leq (1 - e^{-R(\varepsilon)})^{-2} \sum_{i=1}^{J_\varepsilon} [m(F_{h_i, h_i, R(\varepsilon)/2, R(\varepsilon)/2}) + 1],$$

which is finite, establishing (i) in Theorem A.1.

The next step is to check (ii) in Theorem A.1. For  $h \in C_K^+(\mathbb{E})$  and  $R > 0$ , we have, by (A.5),

$$\begin{aligned} \limsup_{n \rightarrow \infty} m_n(\{\xi : \xi(h) > R\}) &\leq \limsup_{n \rightarrow \infty} (1 - e^{-R/2})^{-2} m_n(F_{h, h, R/2, R/2}) \\ &= (1 - e^{-R/2})^{-2} m(F_{h, h, R/2, R/2}). \end{aligned}$$

The latter expression converges to 0 as  $R \rightarrow \infty$ , which implies (ii) in Theorem A.1.

We conclude that  $(m_n)$  is relatively compact in  $\mathbb{M}_0(\mathbb{N}_p)$ .

Since the assumptions of Lemma A.1 are satisfied for any subsequential vague limit point of the sequence  $(m_n)$  and the measure  $m$ , we conclude that all subsequential vague limit points of the sequence  $(m_n)$  coincide with  $m$  and, hence,  $m_n \rightarrow m$  in  $\mathbb{M}_0(\mathbb{N}_p)$ .

### A.2. A mapping theorem

The general version of the mapping theorem is given in Theorem 2.5 of Hult and Lindskog (2006). Here we will state a useful special case.

**Lemma A.2.** *Suppose that  $m_n \rightarrow m$  in  $\mathbb{M}_0(\mathbb{N}_p)$  and that  $f : \mathbb{E} \rightarrow \mathbb{R}^d$  is a measurable function with a bounded support, such that*

$$m(\xi : \xi(\mathcal{D}_f) > 0) = 0,$$

where  $\mathcal{D}_f$  is the set of discontinuities of the function  $f$ . Define  $T : \mathbb{N}_p \rightarrow \mathbb{R}^d$  by  $T(\xi) = \xi(f)$ . Then

$$m_n \circ T^{-1}(\cdot) \rightarrow m \circ T^{-1}(\cdot) \quad \text{in } \mathbb{M}_0(\mathbb{R}^d).$$

*Proof.* This follows from Theorem 2.5 of Hult and Lindskog (2006) since  $T$  is discontinuous on a set of measure  $m$  equal to 0,  $T(\emptyset) = 0$ , and  $T$  is continuous at  $\xi_0$ .

### Acknowledgements

It is a pleasure to thank the anonymous referees for the careful reading of the manuscript and making numerous suggestions for improving the presentation of the paper.

### References

ADLER, R. J., FELDMAN, R. E. AND TAQQU, M. S. (eds) (1998). *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. Birkhäuser, Boston, MA.  
 ALDOUS, D. (1989). *Probability Approximations via the Poisson Clumping Heuristic*. Springer, Berlin.  
 ASMUSSEN, S. (2000). *Ruin Probabilities*. World Scientific, River Edge, NJ.

- BASRAK, B. (2000). The sample autocorrelation function of non-linear time series. Doctoral Thesis, University of Groningen.
- BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. John Wiley, New York.
- DALEY, D. J. AND VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes* Vol. I, 2nd edn. Springer, New York.
- DAVIS, R. A. AND MIKOSCH, T. (2008). Extremes of stochastic volatility models. In *Handbook of Financial Time Series*, eds T. G. Andersen et al. Springer, Berlin, pp. 355–364.
- DEMBO, A. AND ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd edn. Springer, New York.
- EMBRECHTS, P., KLÜPPELBERG, C. AND MIKOSCH, T. (1997). *Modelling Extremal Events*. Springer, Berlin.
- HULT, H. AND LINDSKOG, F. (2006). Regular variation for measures on metric spaces. *Publ. Inst. Math.* **80**, 121–140.
- HULT, H. AND SAMORODNITSKY, G. (2008). Tail probabilities for infinite series of regularly varying random vectors. *Bernoulli* **14**, 838–864.
- HULT, H., LINDSKOG, F., MIKOSCH, T. AND SAMORODNITSKY, G. (2005). Functional large deviations for multivariate regularly varying random walks. *Ann. Appl. Prob.* **15**, 2651–2680.
- KALLENBERG, O. (1983). *Random Measures*, 3rd edn. Akademie, Berlin.
- KESTEN, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207–248.
- KWAPIEŃ, S. AND WOYCZYŃSKI, W. (1992). *Random Series and Stochastic Integrals: Single and Multiple*. Birkhäuser, Boston, MA.
- 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.
- PETROV, V. V. (1995). *Limit Theorems of Probability Theory*. Oxford University Press.
- RACHEV, S. (ed.) (2003). *Handbook of Heavy Tailed Distributions in Finance*. Elsevier, Amsterdam.
- RESNICK, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.
- RESNICK, S. I. (2006). *Probabilistic and Statistical Modeling of Heavy Tail Phenomena*. Springer, Berlin.
- TSONIS, A. AND ELSNER, J. B. (2007). *Nonlinear Dynamics in Geosciences*. Springer, New York.