The Annals of Probability
1986, Vol. 14, No. 2, 612–652

EXTREME VALUE THEORY FOR MOVING AVERAGE PROCESSES[1]

BY HOLGER ROOTZÉN

University of Copenhagen and University of North Carolina at Chapel Hill

This paper studies extreme values in infinite moving average processes $X_t = \Sigma_\lambda c_{\lambda-t} Z_\lambda$ defined from an i.i.d. noise sequence $\{Z_\lambda\}$. In particular this includes the ARMA-processes often used in time series analysis. A fairly complete extremal theory is developed for the cases when the d.f. of the Z_λ's has a smooth tail which decreases approximately as $\exp\{-z^p\}$ as $z \to \infty$, for $0 < p < \infty$, or as a power of z. The influence of the averaging on extreme values depends on p and the c_λ's in a rather intricate way. For $p = 2$, which includes normal sequences, the correlation function $r_t = \Sigma_\lambda c_{\lambda-t} c_\lambda / \Sigma_\lambda c_\lambda^2$ determines extremal behavior while, perhaps more surprisingly, for $p \neq 2$ correlations have little bearing on extremes. Further, the sample paths of $\{X_t\}$ near extreme values asymptotically assume a specific nonrandom form, which again depends on p and $\{c_\lambda\}$ in an interesting way. One use of this latter result is as an informal quantitative check of a fitted moving average (or ARMA) model, by comparing the sample path behavior predicted by the model with the observed sample paths.

1. Introduction. Let $\{X_t = \Sigma c_{\lambda-t} Z_\lambda\}$ be an infinite moving average process, with $\{c_\lambda\}$ given constants and with the noise sequence $\{Z_\lambda\}$ consisting of independent identically distributed (i.i.d.) random variables. Such processes have been extensively studied for both practical and theoretical reasons, and, in particular, include the ARMA (autoregressive-moving average) processes often used in time series analysis (as can be seen by inverting the autoregressive part of the process). In fact also more general, infinite, autoregressions fit into this framework, as discussed in Section 9. In the present paper we study extremal properties connected with such processes, for the case when the marginal distribution of the noise variables, $\{Z_\lambda\}$, has a tail which decreases approximately as a polynomial times $\exp\{-z^p\}$ as $z \to \infty$, for the parameter p ranging over the interval $(0, \infty)$. In the last section, we also comment briefly on earlier results for polynomially decreasing tails.

In addition to extreme values of $\{X_t\}$ itself, we study their relation to extremes of the Z_λ's and of a third related sequence $\hat{X}_1, \hat{X}_2, \ldots,$ the *associated independent sequence*. By definition this is the i.i.d. sequence which has the same marginal distribution function (d.f.) as the X_t's. Extremes of the associated independent sequence are of course completely determined by the tail of the d.f. of $\hat{X}_0$, or equivalently of $\Sigma c_\lambda Z_\lambda$. Hence, to determine the extremal behavior of

Received August 1983; revised April 1985.
[1]Research supported in part by the Air Force Office of Scientific Research Contract No. F49620-82-C-0009.
AMS 1980 *subject classifications*. Primary 60F05; secondary 60G17.
Key words and phrases. Extreme values, moving averages, ARMA processes, sample path properties, distributions of weighted sums.

$\{\hat{X}_t\}$, we have to find accurate approximations for the tails of the d.f. of weighted sums, which may be of interest also outside the present context.

Specifically, writing $M_n = \max\{X_1, \dots, X_n\}$, $\tilde{M}_n = \max\{Z_1, \dots, Z_n\}$, and $\hat{M}_n = \max\{\hat{X}_1, \dots, \hat{X}_n\}$ for any $p > 0$ we find norming constants $a_n, \tilde{a}_n, \hat{a}_n > 0$ and $b_n, \tilde{b}_n, \hat{b}_n$, such that the d.f. of each of $a_n(M_n - b_n)$, $\tilde{a}_n(\tilde{M}_n - \tilde{b}_n)$, and $\hat{a}_n(\hat{M}_n - \hat{b}_n)$ converges to the type I extreme value d.f. $\exp\{-e^{-x}\}$. Here the norming constants depend on p and on the c_λ's in a rather intricate way. In all cases, the b's which give the center of the distribution of the maxima are of the order $(\log n)^{1/p}$, which tends to infinity with n. The a's are of the order $(\log n)^{1-1/p}$, which tends to infinity for $p > 1$, thus showing that the scale of extremes decreases in this case, while it tends to zero for $0 < p < 1$, corresponding to an increasing scale of extremes, and remains constant for $p = 1$. Further, for $p > 1$, $a_n = \hat{a}_n$, $b_n = \hat{b}_n$, and a_n and $\tilde{a}_n$ are of the same order, but b_n may be significantly different from $\tilde{b}_n$ (i.e., often $a_n|b_n - \tilde{b}_n| \to \infty$), and a_n, b_n depend on the weights $\{c_\lambda\}$ through the quantity $\Sigma|c_\lambda|^q$, where q is the conjugate exponent to p, defined by $1/q + 1/p = 1$. For $0 < p < 1$, typically a_n, b_n resemble $\tilde{a}_n$, $\tilde{b}_n$, while $\hat{a}_n$, $\hat{b}_n$ may be slightly different, and in this case it is the maximum of the c_λ's which enters into the normings. The case $p = 1$ provides intermediate behavior.

The convergence results for maxima are obtained as corollaries to much more general point process convergence results for normalized heights and locations of extreme values. This point process convergence also has many other corollaries, e.g., concerning the joint asymptotic distribution of several extreme order statistics, and convergence of so called record time processes and extremal processes. However, these corollaries will not be explicitly stated, and instead the reader is referred to [5], Chapter 5, for a detailed discussion. Moreover, the results are further generalized to take into account also the behavior of sample paths near extremes, showing that asymptotically they assume a specific deterministic form, which depends on p and $\{c_\lambda\}$ in an interesting way. E.g., in the simplest case, when all the c_λ's are nonnegative, for $p > 1$ the suitably normalized sample paths around extremes approach the function

$$y_\tau = \sum_\lambda c_{\lambda-\tau} c_\lambda^{q/p} \Big/ \sum_\lambda c_\lambda^q, \qquad \tau = 0, \pm 1, \dots, \tag{1.1}$$

and for $0 < p < 1$ approach a specific translate of the function

$$y_\tau = c_{-\tau}/\max\{c_\lambda;\, \lambda = 0, \pm 1, \dots\}, \qquad \tau = 0, \pm 1, \dots, \tag{1.2}$$

while the borderline case $p = 1$ mainly resembles $0 < p < 1$. The case of negative c_λ's involves some further complexity. In passing we note that for $p = 2$, which includes the normal distribution, y_τ is in fact the correlation function of $\{X_t\}$. This of course agrees with the well-known extreme value theory for normal sequences. However, perhaps more surprising, for $p \neq 2$ the correlation function does not seem to have any bearing on extremal behavior, and the important role is instead played by the function $\{y_\tau\}$ given by (1.1) or (1.2).

Some "geometrical" heuristics, which originally suggested the results, are illustrated in Figure 1. In the figure it is assumed that $c_0 > 0$, $c_1 > 0$, that the

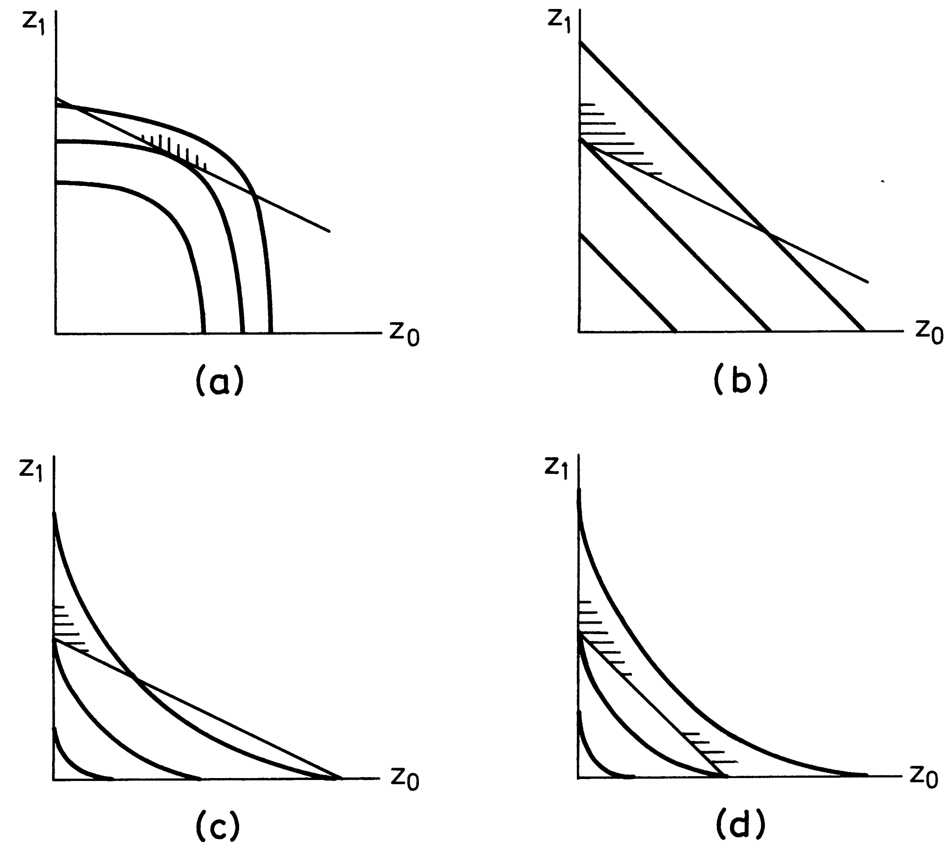

FIG. 1. *Level curves* $\exp\{-\|z\|_p^p\} = \eta$ *for* $\eta = 0.1, 0.01,$ *and* 0.001. *Shaded area contains most of the probability outside the line* $c_0 z_0 + c_1 z_1 = u$. (a) $p = 3$, $c_0 = 1$, $c_1 = 2$; (b) $p = 1$, $c_0 = 1$, $c_1 = 2$; (c) $p = \frac{2}{3}$, $c_0 = 1$, $c_1 = 2$; *and* (d) $p = \frac{2}{3}$, $c_0 = c_1 = 1$. *Different scales in different figures.*

remaining c_λ's are zero, and that the d.f. of the Z_λ's has a density of the form $\exp\{-z^p\}$ for all sufficiently large values of z. From Figure 1a, it can be seen that for $p > 1$ and for large u, most of the probability mass outside the line $c_0 z_0 + c_1 z_1 = u$ is concentrated in a small region around the point where the line is tangent to a level curve of the bivariate density of (Z_0, Z_1). Thus, in general notation, if $X_0 = \Sigma c_\lambda Z_\lambda$ exceeds u, then with high probability $(\dots, Z_0, Z_1, \dots)$ is close to $(\dots, u c_0^{q/p}/\Sigma c_\lambda^q, u c_1^{q/p}/\Sigma c_\lambda^q, \dots)$ and one would expect that for τ close to zero, $X_\tau/u = \Sigma c_{\lambda-\tau} Z_\lambda/u$ would be close to $\Sigma c_{\lambda-\tau} c_\lambda^{q/p} = y_\tau$. In particular, for $p > 1$, large values of X_0 are hence caused by rare combinations of many moderately large noise variables. For $0 < p < 1$ and $c_0 < c_1$, the probability mass outside $c_0 z_0 + c_1 z_1 = u$ is concentrated around the point $z_0 = 0$, $z_1 = u/c_1$, cf. Figure 1c so that by similar reasoning, if X_0 exceeds u for τ close to zero, one would expect X_τ/u to be close to $c_{-\tau}/c_0 = y_\tau$. If $c_0 = c_1$, half of the probability mass outside $c_0 z_0 + c_1 z_1 = u$ is concentrated near $z_0 = u/c_0$, $z_1 = 0$, and the other half near $z_0 = 0$, $z_1 = u/c_1$, as shown in Figure 1d which leads X_τ/u to be close to $c_{-\tau}/c_0 = y_\tau$ with probability $\frac{1}{2}$ and close to $c_{1-\tau}/c_0 = y_{\tau-1}$ with probability $\frac{1}{2}$, if X_0 exceeds u and τ is small. Thus, in both cases, extremes of X_0 are caused by just one Z_λ being large, but if $c_0 = c_1$, it may be either one of Z_0 and Z_1. Again the case $p = 1$ is similar to $0 < p < 1$, but with the added complexity that if, say, $c_0 = c_1 > 0$, $c_\lambda = 0$; $\lambda \neq 0, 1$, then large values of X_0 may be caused by more than one of the Z_λ's being simultaneously large, as can be guessed from Figure 1b.

A main part of the proofs for each of the three cases $p > 1$, $p = 1$, and $0 < p < 1$ is to obtain accurate approximations for the tail of the d.f. of $\Sigma c_\lambda Z_\lambda$. For the hardest case, $p > 1$, this is done in a companion paper, ref. [8]. For $p = 1$ the tail behavior is simpler, and the proof given below is made easier by the possibility to use moment generating functions rather straightforwardly. Finally, for $0 < p < 1$, convolution integrals are easy to estimate and give the desired approximation for the tail of the d.f.

Furthermore, for $p \geq 1$, extremal theory for the moving average process $\{X_t\}$ itself is obtained via Leadbetter's "distributional mixing conditions" as given in [5], while for the case $0 < p < 1$ we use a direct approach related to methods in [7]. Finally, the sample path results are obtained via direct calculations, which are closely related to the heuristics presented above.

There is a large literature on general extreme value theory for independent and dependent sequences, and in particular, normal sequences have been studied in extensive detail (for a recent survey, see [5]), but there is not much written on the present subject. Moving averages of stable variables (which have polynomially decreasing tails) are extensively discussed in Rootzén (1978) (see also Section 9). Finster (1982) found the asymptotic distribution of maxima of autoregressive processes when the noise variables have exponential tails (corresponding to the case $p = 1$, $\alpha = 0$, $k_+ = 1$ in Section 7) and for noise variables with polynomially decreasing tails. (There is some overlap, apparently not noticed by Finster, between the latter result and those of [7]). Finster's conditions are in terms of an autoregressive representation of the process, although many of the computations are made after inverting to a moving average representation. This seems to make them somewhat less directly connected with the core of the problem. Chernick (1981) has exhibited further qualitatively different behavior of extreme values of autoregressive processes, which by inversion can be translated to moving average processes, for a case when the noise variables have nonsmooth tails. Finally the extensive literature on normal sequences (see e.g., [5]) of course also concerns moving averages, since any normal sequence which has an absolutely continuous spectral distribution also has a moving average representation.

The present paper is an attempt at a rather complete qualitative and quantitative study of extreme values of moving averages of variables with smooth tails. As alluded to above, the practical motivation for the study is the importance of moving averages (or "filtered white noise") models, and that extreme values are inherently important in many of their applications. Further, as a byproduct, the results on sample path behavior near extremes may be used as an informal, quantitative check of a fitted moving average (or ARMA) model, by comparing the sample path behavior predicted by the model with the observed sample paths. A theoretical motivation is to provide a testing ground for the general extreme value theory for dependent sequences and impetus for further development of that theory and to provide a mathematically interesting example of some of the quite complex ways in which dependence affects extremal behavior.

Each of Sections 5–9 starts with a more detailed overview of that section. Sections 5, 6, and 7 and 9 on $p > 1$, on $p = 1$, and on $0 < p < 1$ can be read independently of one another.

2. Definitions and conditions. For the study of extreme values of the moving average process

$$X_t = \sum_{\lambda} c_{\lambda - t} Z_{\lambda}, \qquad t = 0, \pm 1, \dots, \tag{2.1}$$

we need conditions on the *noise variables* $\{Z_\lambda\}$, conditions on the *weights* $\{c_\lambda\}$, and conditions involving $\{Z_\lambda\}$ and $\{c_\lambda\}$ simultaneously. In addition the conditions will depend on the parameter p introduced in (2.2) below, being more stringent for $p > 1$ than for $p = 1$ or $0 < p < 1$.

The Z_λ's will always be i.i.d. random variables, and for convenience of notation we will let Z be another random variable with the same distribution as the Z_λ's. Throughout, it will be assumed that

$$P(Z > z) \sim K z^{\alpha} e^{-z^p} \quad \text{as } z \to \infty, \tag{2.2}$$

where p, K are positive parameters and α is a real parameter, and that the first moment exists, $E|Z| < \infty$, and for $p \geq 1$ in addition that $EZ^2 < \infty$. [Here $A(z) \sim B(z)$ has the standard meaning that $A(z)/B(z) \to 1$.] For $p > 1$, (2.2) has to be substantially strengthened. We will then suppose that the distribution of Z has a continuously differentiable density f which satisfies

$$f(z) \sim K' z^{\alpha'} e^{-z^p} \quad \text{as } z \to \infty, \tag{2.3}$$

for $\alpha' = \alpha + p - 1$, $K' = Kp$, and that

$$e^{cz} f'(z) \text{ is bounded for } z \in (-\infty, 0], \tag{2.4}$$

for some constant $c \geq 0$. Moreover, defining $D(z) = f(z) e^{z^p}$ for $z \geq 0$, and $D(z) = f(z)$ otherwise so that

$$f(z) = \begin{cases} D(z) e^{-z^p} & \text{for } z \geq 0, \\ D(z) & \text{for } z < 0, \end{cases} \tag{2.5}$$

with

$$D(z) \sim K' z^{\alpha'} \quad \text{as } z \to \infty, \tag{2.6}$$

we assume that

$$\limsup_{z \to \infty} \left| \frac{z D'(z)}{D(z)} \right| < \infty. \tag{2.7}$$

Here of course f' and D' are the derivatives of f and D. The reason for the particular choice of α', K' is that with this choice (2.3) implies (2.2), so that the parameters have the same meaning for $p > 1$ and for $0 < p \le 1$. It may be further noted that (2.7) e.g. is satisfied if $D(z)$ for large z is a rational function of z.

The conditions on the weights are that at least one c_λ is strictly positive, and that

(2.8) $$|c_\lambda| = O(|\lambda|^{-\theta}) \quad \text{as } \lambda \to \pm\infty, \text{for some } \theta > 1,$$

which again has to be strengthened for $p > 1$, to

(2.9) $$|c_\lambda| = O(|\lambda|^{-\theta}) \quad \text{as } \lambda \to \pm\infty, \text{for some } \theta > \max(1, 2/q),$$

where as in the introduction q is the conjugate exponent of p, defined by $1/p + 1/q = 1$. In particular, the condition (2.8) implies that $\Sigma|c_\lambda| < \infty$, which together with $E|Z| < \infty$ ensures a.s. convergence of the sums in (2.1), which define X_t. In the sequel, some further notation pertaining to the c_λ's will be needed. Let $c_\lambda^+ = \max(0, c_\lambda)$, $c_\lambda^- = \max(0, -c_\lambda)$, $c_+ = \max\{c_\lambda^+; \lambda = 0, \pm 1, \dots\}$, and $c_- = \max\{c_\lambda^-; \lambda = 0, \pm 1, \dots\}$ and let $\Lambda_+ = \{\lambda_1, \dots, \lambda_{k_+}\}$ be the set of λ's for which $c_\lambda = c_+$, and let $\Lambda_- = \{\lambda_1^-, \dots, \lambda_{k_-}^-\}$ be defined similarly from $\{c_\lambda^-\}$ with $\Lambda_- = \varnothing$ if $c_- = 0$. Further, with standard notation, we will write $\|c\|_q = \{\Sigma_\lambda |c_\lambda|^q\}^{1/q}$ and $\|c^+\|_q = \{\Sigma_\lambda |c_\lambda^+|^q\}^{1/q}$ for $q > 1$.

The reason that conditions involving weights and noise variables simultaneously are needed is the following. If some of the c_λ's are negative then extremes of $\{X_t\}$ may be influenced also by the left tail of the distribution of Z, and this influence is determined by how a combination of $\{c_\lambda^-\}$ and the left tail of Z compares with the corresponding combination of $\{c_\lambda^+\}$ and the right tail of Z. There are three cases of interest, which we will refer to as the case of *positive* c_λ's, the case of a *dominating right tail*, and the case of *balanced tails*. (Of course, the results for the potential fourth case, a *dominating left tail* are immediate consequences of the results for a dominating right tail.) The precise meaning of the three cases will be somewhat different for $0 < p \le 1$ and for $p > 1$, and will be formalized in three conditions, to be called A.1–A.3 for $0 < p \le 1$ and B.1–B.3 for $p > 1$, respectively. The conditions for $0 < p \le 1$ are

A.1 (2.2) and (2.8) hold, and all c_λ's are nonnegative,

A.2 (2.2) and (2.8) hold, and $P(Z < z) = O(e^{-|z|^p/\gamma})$ as $z \to -\infty$, where γ satisfies $c_-\gamma^{1/p} < c_+$ and

A.3 (2.2) and (2.8) hold, and $P(Z < z) \sim K_-|z|^\alpha e^{-|z|^p/\gamma}$, for some constant $K_- > 0$, where $c_-\gamma^{1/p} = c_+$, and α is the same as in (2.2).

The conditions for $p > 1$ are

B.1 $p > 1$, (2.3), (2.4), (2.7), and (2.9) hold, and all c_λ's are nonnegative,

B.2 $p > 1$, (2.3), (2.7), and (2.9) hold, and in addition $f(-z)$ satisfies (2.3) and (2.7), with p in (2.3) replaced by some $p' > p$, and possibly with different D, α', K', and

B.3 $p > 1$, (2.3), (2.7), and (2.9) hold, and in addition $f(-z)$ satisfies (2.3) and (2.7) with the same p as in (2.3), but possibly with different D, α', K'.

The main results of this paper, in addition to approximations for the tails of the distribution of the weighted sums $\Sigma c_\lambda Z_\lambda$, concern convergence of point processes of heights and locations of extreme values of $\{X_t\}$, and of more general "marked" point processes which retain information also about the behavior of sample paths near extremes. The reader is referred to [5] for definitions and information on point process convergence in extreme value theory, and to [4] and [6] for the general theory of point processes. Reference [4] only treats locally compact spaces, and there "bounded" has the technical meaning of being relatively compact, while [6] covers general Polish spaces. However, throughout this paper in the cases where both approaches apply, they coincide, as readily seen. Specifically, we will let N_n denote the point process in $[0,\infty) \times \mathbb{R}$ which consists of the points $(j/n, a_n(X_j - b_n))$, $j = 1, 2, \ldots$, and will for each $p > 0$ find a point process N and choose the constants $a_n > 0$, b_n so that N_n converges in distribution to N (denoted $N_n \to_d N$). As discussed in [5], Chapter 5, this implies many asymptotic results, e.g., on the joint distribution of the k largest extreme order statistics, on the so called record time process and the extremal process. However, we will only explicitly note the corollary that, for $M_n = \max\{X_1, \ldots, X_n\}$,

$$P(a_n(M_n - b_n) \le x) \to e^{-e^{-x}} \quad \text{as } n \to \infty.$$

Next, let

$$Y'_{n,j}(\tau) = X_{\tau+j}/b_n, \qquad j = 1, 2, \ldots,$$

and

$$Y''_{n,j}(\tau) = X_{\tau+j}/X_j, \qquad j = 1, 2, \ldots,$$

(defined e.g. to be zero for $X_j = 0$) be the normalized sample path around X_j, write $S = [0,\infty) \times \mathbb{R}$, and let $\mathbb{R}^\infty = \cdots \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \cdots = \{x; x = (\ldots, x_{-1}, x_0, x_1, \ldots)\}$ be the space of doubly infinite sequences of real numbers. The processes $Y'_{n,j}$ and $Y''_{n,j}$ are then the "marks" and are random variables in the "mark space" $\mathbb{R}^\infty$, and the marked point processes N'_n and N''_n are just the ordinary point processes in $S \times \mathbb{R}^\infty$ which consist of the points $((j/n, a_n(X_j - b_n)), Y'_{n,j})$, $j = 1, 2, \ldots,$ and of the points $((j/n, a_n(X_j - b_n)), Y''_{n,j})$, $j = 1, 2, \ldots$, respectively. As in [6] we will assume the mark space $\mathbb{R}^\infty$ is given some *bounded* metric which generates the product topology and will consider $S \times \mathbb{R}$ as a Polish space, with the product of this metric and the ordinary metric in S as metric. In particular, this means that a product set $A_1 \times A_2$ in $S \times \mathbb{R}^\infty$ is bounded if A_1 is bounded. Let $y = \{y_\tau\}_{\tau=-\infty}^{\infty}$ be a given point in $\mathbb{R}^\infty$. Often the limit, say N', of N'_n or N''_n is obtained by *adjoining the mark* y *to each point of* N, i.e., if N has the points (t_j, x_j), $j = 1, 2, \ldots$, then N' is defined to be the point process consisting of the points $((t_j, x_j), y)$, $j = 1, 2, \ldots$.

Further, as in the introduction we will write $\tilde{M}_n = \max\{Z_1, \ldots, Z_n\}$ and $\hat{M}_n = \max\{\hat{X}_1, \ldots, \hat{X}_n\}$, where $\hat{X}_1, \hat{X}_2, \ldots$ is the independent sequence associated with $\{X_t\}$. Similarly, for norming constants $\tilde{a}_n$, $\hat{a}_n > 0$ and $\tilde{b}_n$, $\hat{b}_n$ to be specified below, we let the point processes $\tilde{N}_n$, $\tilde{N}'_n$, $\tilde{N}''_n$ and $\hat{N}_n$, $\hat{N}'_n$, $\hat{N}''_n$ be defined from $\{Z_\lambda\}, \{\tilde{a}_n, \tilde{b}_n\}$ and from $\{\hat{X}_t\}, \{\hat{a}_n, \hat{b}_n\}$ in the same way as N_n, N'_n, N''_n are defined from $\{X_t\}, \{a_n, b_n\}$.

Finally, some general points of notation. If limits of summation or integration are deleted, then the summation or integration is always from $-\infty$ to $+\infty$ and summation from a to b, where a and b are not necessarily integers, means summation over all integers in the closed interval $[a, b]$. $N(0, \sigma^2)$ denotes the normal distribution with mean zero and variance σ^2. Often C and γ will be generic constants whose value may change from one appearance to the next. The indicator function is denoted by I, i.e., $I\{\cdot\}$ is one if the event within curly brackets occurs, and zero otherwise. Convergence in probability is denoted $\to_P$.

3. Preliminaries: extreme values of moving averages and point process convergence. For $p \geq 1$, convergence of the point process N_n of heights and locations of extreme values will be proved through verifying Leadbetter's conditions $D_r(\mathbf{u}_n)$ and $D'(u_n)$, as given in [5], pages 107 and 58. In the first part of this section, we modify the conditions to forms which are particularly convenient in the present context. Then we obtain two lemmas which will be useful for $0 < p < 1$ and for the marked point process results, respectively.

The condition $D_r(\mathbf{u}_n)$ will be established via the following lemma, which, for later reference also, is stated separately here, under general conditions. It is given in a rather crude form, which however suffices for our present purposes.

LEMMA 3.1. *Suppose that the moving average process $\{X_t\}$ given by* (2.1) *is defined by a.s. convergent sums and for some constants $a_n > 0$, b_n and nondegenerate distribution function G, it holds that*

$$P(a_n(\hat{M}_n - b_n) \leq x) \to G(x) \quad \text{as } n \to \infty \tag{3.1}$$

for each x with $G(x) > 0$, where $\hat{M}_n$ is the maximum in the associated independent sequence.

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

$$\begin{aligned} &nP\left(a_n\left|\sum_{n\nu}^{\infty} c_\lambda Z_\lambda\right| > \varepsilon\right) \to 0,\\ &nP\left(a_n\left|\sum_{-\infty}^{-n\nu} c_\lambda Z_\lambda\right| > \varepsilon\right) \to 0 \end{aligned} \tag{3.2}$$

as $n \to \infty$, then $\{X_t\}$ satisfies $D_r(\mathbf{u}_n)$ for arbitrary r and $\mathbf{u}_n = (u_n^{(1)}, \dots, u_n^{(r)})$ with $u_n^{(i)} = x_i/a_n + b_n$, for arbitrary $x_1, \dots, x_r$.

(ii) *If $a_n = O((\log n)^\beta)$ for some β, $|c_\lambda| = O(|\lambda|^{-\theta})$ for some $\theta > 1$ and $EZ^2 < \infty$, then* (3.2), *and hence also $D_r(\mathbf{u}_n)$, holds for all $\{\mathbf{u}_n\}$ of this form.*

PROOF. (i) We will only verify $D_r(\mathbf{u}_n)$ for $r = 1$, which is the same as to verify the condition $D(u_n)$ of [5], page 53. The extension to $r > 1$ is completely straightforward, involving notational problems only, and is omitted. Thus, let $u_n = x/a_n + b_n$, and assume $G(x) > 0$, since $D(u_n)$ trivially holds if $G(x) = 0$. Let $1 \leq i_1 < \cdots < i_r < j_1 < \cdots < j_s \leq n$ be integers with $j_1 - i_r \geq 2n\nu$ for fixed $\nu > 0$. For brevity of notation, write $\mathbf{X}_\mathbf{i} = (X_{i_1}, \dots, X_{i_r})$, $\mathbf{X}_\mathbf{j} = (X_{j_1}, \dots, X_{j_s})$

and similarly $\mathbf{X}'_{\mathbf{i}} = (X'_{i_1}, \ldots, X'_{i_r})$, $\mathbf{X}''_{\mathbf{j}} = (X''_{j_1}, \ldots, X''_{j_s})$, for

$$X'_t = \sum_{-\infty}^{n\nu-1} c_\lambda Z_{\lambda+t}, \qquad X''_t = \sum_{-n\nu+1}^{\infty} c_\lambda Z_{\lambda+t}.$$

Further, let $M'_n = \max\{|X_1 - X'_1|, \ldots, |X_n - X'_n|\}$ and $M''_n = \max\{|X_1 - X''_1|, \ldots, |X_n - X''_n|\}$, and in the sequel let an inequality between a real number and a vector mean that the inequality holds between the number and each component of the vector. Clearly, since $j_1 - i_r \geq 2n\nu$, $\mathbf{X}'_{\mathbf{i}}$ and $\mathbf{X}''_{\mathbf{j}}$ are independent, and hence for $\varepsilon > 0$,

$$\begin{aligned} P(\mathbf{X}_{\mathbf{i}} \leq u_n, \mathbf{X}_{\mathbf{j}} \leq u_n) &\leq P(\mathbf{X}'_{\mathbf{i}} \leq u_n + \varepsilon) P(\mathbf{X}''_{j} \leq u_n + \varepsilon) \\ &\quad + P(M'_n > \varepsilon) + P(M''_n > \varepsilon) \\ &\leq P(\mathbf{X}_{\mathbf{i}} \leq u_n + 2\varepsilon) P(\mathbf{X}_{\mathbf{j}} \leq u_n + 2\varepsilon) \\ &\quad + 2P(M'_n > \varepsilon) + 2P(M''_n > \varepsilon) \\ &\leq P(\mathbf{X}_{\mathbf{i}} \leq u_n) P(\mathbf{X}_{\mathbf{j}} \leq u_n) + \sum_{t=1}^{n} P(u_n < X_t \leq u_n + 2\varepsilon) \\ &\quad + 2P(M'_n > \varepsilon) + 2P(M''_n > \varepsilon). \end{aligned} \tag{3.3}$$

A corresponding lower bound is readily obtained, and after using stationarity and Boole's inequality to estimate the last two terms, this shows that

$$\begin{aligned} \Delta_n &\doteq \left| P(\mathbf{X}_{\mathbf{i}} \leq u_n, \mathbf{X}_{\mathbf{j}} \leq u_n) - P(\mathbf{X}_{\mathbf{i}} \leq u_n) P(\mathbf{X}_{\mathbf{j}} \leq u_n) \right| \\ &\leq nP(u_n - 2\varepsilon < X_0 \leq u_n + 2\varepsilon) + 2nP(|X_0 - X'_0| > \varepsilon) \\ &\quad + 2nP(|X_0 - X''_0| > \varepsilon). \end{aligned}$$

Here, the bounds do not depend on the specific choices of $\mathbf{i}$ and $\mathbf{j}$ (subject to $1 \leq i_1, j_s \leq n$, $j_1 - i_r \geq 2n\nu$), and hence, replacing ε by ε/a_n and writing $u_n + 2\varepsilon/a_n = (x + 2\varepsilon)/a_n + b_n$, etc., we have that

$$\begin{aligned} \sup_{\mathbf{i},\mathbf{j}} \Delta_n &\leq nP((x - 2\varepsilon)/a_n + b_n < X_0 \leq (x + 2\varepsilon)/a_n + b_n) \\ &\quad + 2nP\left(a_n \left| \sum_{n\nu}^{\infty} c_\lambda Z_\lambda \right| > \varepsilon \right) + 2nP\left(a_n \left| \sum_{-\infty}^{-n\nu} c_\lambda Z_\lambda \right| > \varepsilon \right). \end{aligned}$$

The last two terms tend to zero by assumption (3.2), and since furthermore, according to (3.1) and [5], Theorem 1.5.1, $P(X_0 > x/a_n + b_n) \sim (-\log G(x))/n$, we have that

$$\begin{aligned} &nP((x - 2\varepsilon)/a_n + b_n < X_0 \leq (x + 2\varepsilon)/a_n + b_n) \\ &\quad \to \log G(x + 2\varepsilon) - \log G(x - 2\varepsilon). \end{aligned}$$

It follows from (3.1) that $G(x)$ is an extreme value distribution, and hence continuous, and thus, since $G(x) > 0$, $\log G(x + 2\varepsilon) - \log G(x - 2\varepsilon) \to 0$ as $\varepsilon \to 0$. Hence $\sup_{\mathbf{i},\mathbf{j}} \Delta_n \to 0$ as $n \to \infty$, and since $\nu > 0$ is arbitrary, this shows that the hypothesis in Lemma 3.2.1(ii) of [5] is satisfied, and thus that $D(u_n)$ holds.

(ii) Let $\mu = EZ$, $\sigma^2 = V(Z)$. The assumptions on a_n and $\{c_\lambda\}_{\lambda=-\infty}^{\infty}$ show that $a_n \Sigma_{n\nu}^{\infty} |c_\lambda| \to 0$ as $n \to \infty$, and hence for large n Chebycheff's inequality gives that

$$nP\left(a_n \left| \sum_{n\nu}^{\infty} c_\lambda Z_\lambda \right| > \varepsilon\right) \le nP\left(a_n \left| \sum_{n\nu}^{\infty} c_\lambda (Z_\lambda - \mu) \right| > \varepsilon - \mu a_n \sum_{n\nu}^{\infty} |c_\lambda|\right)$$
$$\le n \frac{\sigma^2 a_n^2 \Sigma_{n\nu}^{\infty} c_\lambda^2}{\left(\varepsilon - \mu a_n \Sigma_{n\nu}^{\infty} |c_\lambda|\right)^2}.$$

The assumptions on a_n, $\{c_\lambda\}$ are again readily seen to imply that this tends to zero. The proof of the second part of (3.2) is identical. □

The next result shows how $D'(u_n)$ may be checked for moving average processes, and combining this with the previous lemma gives conditions for convergence of N_n. To avoid the (trivial) complication which arises when G has a finite left endpoint, we only state it for $G(x) = \exp\{-e^{-x}\}$.

LEMMA 3.2. *Suppose that for some constant $\gamma \in (0,1]$, and writing $n' = [n^\gamma]$, it holds for $u_n = x/a_n + b_n$ for any x, that*

$$n \sum_{t=1}^{2n'} P(X_0 + X_t > 2u_n) \to 0 \quad \text{as } n \to \infty, \tag{3.4}$$

$$n^2 P\left(a_n \sum_{n'+1}^{\infty} c_\lambda Z_\lambda > 1\right) \to 0, \qquad n^2 P\left(a_n \sum_{-\infty}^{-n'-1} c_\lambda Z_\lambda > 1\right) \to \infty \tag{3.5}$$

as $n \to \infty$, and that

$$P\left(\sum_{-\infty}^{n'} c_\lambda Z_\lambda > u_n\right) = O(1/n), \qquad P\left(\sum_{-n'}^{\infty} c_\lambda Z_\lambda > u_n\right) = O(1/n). \tag{3.6}$$

Then $D'(u_n)$ holds for $u_n = x/a_n + b_n$ for any x. If in addition the hypothesis of Lemma 3.1(i) *or* (ii) *is satisfied, with $G(x) = \exp\{-e^{-x}\}$, then* (3.6) *may be replaced by*

$$a_n \sum_{n'+1}^{\infty} c_\lambda Z_\lambda \to_P 0, \qquad a_n \sum_{-\infty}^{-n'-1} c_\lambda Z_\lambda \to_P 0, \tag{3.6'}$$

and for N_n as defined in Section 2, $N_n \to_d N$ in $[0, \infty) \times \mathbb{R}$, where N is a Poisson process with intensity measure $dt \times e^{-x}\,dx$.

PROOF. By [5], Theorems 5.7.2 and 3.5.2, the second part of the conclusion is immediate from Lemma 3.1 and the first part, and hence we only have to prove $D'(u_n)$, i.e., that

$$\limsup_{n \to \infty} n \sum_{t=1}^{[n/k]} P(X_0 > u_n, X_t > u_n) \to 0 \quad \text{as } k \to \infty, \tag{3.7}$$

for any $u_n = x/a_n + b_n$. Since $P(X_0 > u_n, X_t > u_n) \le P(X_0 + X_t > 2u_n)$ it follows at once from (3.4) that

$$n \sum_{t=1}^{2n'} P(X_0 > u_n, X_t > u_n) \to 0 \quad \text{as } n \to \infty. \tag{3.8}$$

Next write $X_0' = \sum_{-\infty}^{n'} c_\lambda Z_\lambda$, $X_t'' = \sum_{-n'}^{\infty} c_\lambda Z_{\lambda+t}$ so that X_0' and X_t'' are independent for $t > 2n'$. By similar reasoning as in Lemma 3.1(i) for $t > 2n'$

$$\begin{aligned} P(X_0 > u_n, X_t > u_n) &\le P(X_0' > u_n - 1/a_n)P(X_t'' > u_n - 1/a_n) \\ &\quad + P\left(\sum_{n'+1}^{\infty} c_\lambda Z_\lambda > 1/a_n \right) + P\left(\sum_{-\infty}^{-n'-1} c_\lambda Z_{\lambda+t} > 1/a_n \right), \end{aligned}$$

and hence, using stationarity, and writing $u_n' = (x-1)/a_n + b_n$, we have that

$$\begin{aligned} &n \sum_{t=2n'+1}^{[n/k]} P(X_0 > u_n, X_t > u_n) \\ &\quad \le (n^2/k) P\left(\sum_{-\infty}^{n'} c_\lambda Z_\lambda > u_n' \right) P\left(\sum_{-n'}^{\infty} c Z_\lambda > u_n' \right) \\ &\quad\quad + n^2 P\left(a_n \sum_{n'+1}^{\infty} c_\lambda Z_\lambda > 1 \right) + n^2 P\left(a_n \sum_{-\infty}^{-n'-1} c_\lambda Z_\lambda > 1 \right). \end{aligned}$$

Here the last two terms tend to zero by (3.5), and since (3.6) holds for all x, it also holds with u_n replaced by u_n', so that

$$\limsup_{n\to\infty} n \sum_{t=2n'+1}^{\infty} P(X_0 > u_n, X_t > u_n) \le c/k \to 0 \quad \text{as } k \to \infty$$

for some suitable constant c, which together with (3.8) proves (3.7).

Finally, to see that (3.6) may be replaced by (3.6′) under the stated conditions, first note that (3.1) implies that

$$\begin{aligned} P\left(\sum c_\lambda Z_\lambda > u_n - 1/a_n \right) &\sim n^{-1} \log G(x-1) \\ &= O(1/n), \end{aligned}$$

see [5], Theorem 1.5.1. Hence, since the Z_λ's are independent,

$$\begin{aligned} &P\left(\sum_{-\infty}^{n'} c_\lambda Z_\lambda > u_n \right) P\left(\sum_{n'+1}^{\infty} c_\lambda Z_\lambda > -1/a_n \right) \\ &\quad = P\left(\sum_{-\infty}^{n'} c_\lambda Z_\lambda > u_n, \sum_{n'+1}^{\infty} c_\lambda Z_\lambda > -1/a_n \right) \\ &\quad \le P\left(\sum c_\lambda Z_\lambda > u_n - 1/a_n \right) \\ &\quad = O(1/n), \end{aligned}$$

and as $P(\sum_{n'+1}^{\infty} c_\lambda Z_\lambda > -1/a_n) \to 1$ by (3.6′), the first part of (3.6) follows. The proof of the second part is the same. □

For $0 < p < 1$ we will use the characterization of point process convergence in terms of "finite-dimensional" distributions, viz. that $N_n \to_d N$ in $[0, \infty) \times \mathbb{R} = S$ if and only if

$$(N_n(I_1), \dots, N_n(I_k)) \to_d (N(I_1), \dots, N(I_k)) \quad \text{as } n \to \infty, \tag{3.9}$$

as random vectors in $\mathbb{R}^k$, for any k and finite rectangles $I_1, \dots, I_k$ in S, of the form $[t_1, t_2) \times (x_1, x_2]$, with $P(N(\partial I_j) > 0) = 0$, for $j = 1, \dots, k$, where ∂I_j denotes the boundary of I_j ([4], Theorem 4.2, or [6], Theorem 3.1.7).

LEMMA 3.3. *Let N, $N_n^{(1)}$, and $N_n^{(2)}$ be point processes in $[0, \infty) \times \mathbb{R}$ such that*

$$P\big(N_n^{(1)}(I) \neq N_n^{(2)}(I)\big) \to 0 \quad \text{as } n \to \infty \tag{3.10}$$

for any rectangle I of the form $I = [t_1, t_2) \times (x, \infty)$. Then $N_n^{(1)} \to_d N$ if and only if $N_n^{(2)} \to_d N$.

PROOF. If $I = [t_1, t_2) \times (x_1, x_2]$ is a finite rectangle, then for $I' = [t_1, t_2) \times (x_1, \infty)$, and $I'' = [t_1, t_2) \times (x_2, \infty)$

$$\{N_n^{(1)}(I) \neq N_n^{(2)}(I)\} \subset \{N_n^{(1)}(I') \neq N_n^{(2)}(I')\} \cup \{N_n^{(1)}(I'') \neq N_n^{(2)}(I'')\},$$

since $N_n^{(1)}$ and $N_n^{(2)}$ are measures, and hence additive. Thus, since (3.10) holds for I replaced by I' or by I'', it also holds for $I = [t_1, t_2) \times (x_1, x_2]$. It then follows simply that (3.9) holds with N_n replaced by $N_n^{(1)}$ if and only if it holds with N_n replaced by $N_n^{(2)}$, which in turn proves the lemma. □

To prove convergence of the marked point processes, a slightly more involved description of the sets in (3.9) is needed. Let $I = I^{(1)} \times I^{(2)}$ be the product of a rectangle $I^{(1)} = [t_1, t_2) \times (x_1, x_2]$ in $[0, \infty) \times \mathbb{R} = S$ and a rectangle $I^{(2)} = \cdots \times \mathbb{R} \times J_{-l} \times \cdots \times J_0 \times \cdots \times J_l \times \mathbb{R} \times \cdots$ in $\mathbb{R}^\infty$ with $(2l + 1)$-dimensional base, and with $J_\tau = (u_\tau, v_\tau]$, $\tau = -l, \dots, l$. Further let $\mathscr{I}$ be the class of all sets of this form for $l \geq 0$. With this notation, if N_n', N' are point processes in $S \times \mathbb{R}^\infty$, then $N_n' \to_d N'$ if and only if

$$(N_n'(I_1), \dots, N_n'(I_k)) \to (N'(I_1), \dots, N'(I_k)) \quad \text{as } n \to \infty, \tag{3.11}$$

for any k and $I_1, \dots, I_k \in \mathscr{I}$ with $P(N'(\partial I_j) > 0) = 0$, $j = 1, \dots, k$, by Theorem 3.1.7 of [6], since the class of sets $I \in \mathscr{I}$ with this property satisfies the requirements for the semiring in that theorem.

LEMMA 3.4. *Let N_n, $\{Y_{n,j}'\}$, and N_n' be as defined on page 618. Suppose $N_n \to_d N$ as $n \to \infty$, that N' is obtained by adjoining the mark $y = \{y_\tau\}_{\tau=-\infty}^{\infty}$ to each point of N and that, for any $\varepsilon > 0$ and τ,*

$$P\big(X_0 > u_n, |Y_{n,0}'(\tau) - y_\tau| > \varepsilon\big) = o(1/n) \quad \text{as } n \to \infty \tag{3.12}$$

with $u_n = x/a_n + b_n$ and for any x. Then $N_n' \to_d N'$ as $n \to \infty$ in $S \times \mathbb{R}^\infty$.

PROOF. Let h be the function which maps N into N', and let $\overline{N}_n$ be obtained by adjoining y to each point of N_n, i.e., let $\overline{N}_n = h(N_n)$. Clearly, h is continuous, and hence $N_n \to_d N$ implies $\overline{N}_n = h(N_n) \to_d h(N) = N'$. Thus, reasoning as in the proof of Lemma 3.3, using (3.11) instead of (3.9), the result follows if we prove that

$$P(N_n'(I) \neq \overline{N}_n(I)) \to 0 \quad \text{as } n \to \infty \tag{3.13}$$

for any $I \in \mathscr{I}$ with $P(N'(\partial I) > 0) = 0$.

To prove (3.13), assume first there is a τ_0 with $y_{\tau_0} \notin J_{\tau_0}$. Without loss of generality we may assume that $P(N(I^{(1)}) > 0) > 0$, and it then follows from $P(N'(\partial I) > 0) = 0$ that there is a τ with $y_\tau \notin J_\tau \cup \partial J_\tau$. For that τ, let $\varepsilon > 0$ be the distance between y_τ and J_τ. Then clearly $\overline{N}_n(I) = 0$, and, using stationarity and (3.12), we obtain that for $u_n = x_1/a_n + b_n$

$$\begin{aligned} P(N_n'(I) > 0) &\leq \sum_{nt_1 < t \leq nt_2} P\big(X_t > u_n, |Y_{n,t}(\tau) - y_\tau| > \varepsilon\big) \\ &\leq n(t_2 - t_1) P\big(X_0 > u_n, |Y_{n,0}(\tau) - y_\tau| > \varepsilon\big) \\ &\to 0 \quad \text{as } n \to 0, \end{aligned}$$

so that (3.13) holds in this case. Similarly, as above, if $y_\tau \in J_\tau$ for $\tau = -l, \dots, l$ we may assume that the minimum of the distances between y_τ and the complement of J_τ, for $\tau = -l, \dots, l$ is $\varepsilon > 0$. It is then readily seen, again with $u_n = x_1/a_n + b_n$, that

$$\begin{aligned} P(N_n'(I) \neq \overline{N}_n(I)) &\leq \sum_{nt_1 < t \leq nt_2} \sum_{\tau = -l}^{l} P\big(X_t > u_n, |Y_{n,t}(\tau) - y_\tau| > \varepsilon\big) \\ &\leq n(t_2 - t_1) \sum_{\tau = -l}^{l} P\big(X_0 > u_n, |Y_{n,0}(\tau) - y_\tau| > \varepsilon\big) \\ &\to 0 \quad \text{as } n \to \infty, \end{aligned}$$

proving (3.13) also for this case. □

4. Extremes of the noise sequence. By similar calculations as in [5], Theorem 1.5.3, it is readily seen that (2.2) implies

$$P\big(\tilde{a}_n(\tilde{M}_n - \tilde{b}_n) \leq x\big) \to e^{-e^{-x}} \quad \text{as } n \to \infty \tag{4.1}$$

for

$$\begin{aligned} \tilde{a}_n &= p(\log n)^{1-1/p}, \\ \tilde{b}_n &= (\log n)^{1/p} + p^{-1}((\alpha/p)\log\log n + \log K)(\log n)^{1/p-1}. \end{aligned} \tag{4.2}$$

Alternatively, by Theorem 1.5.1 of the cited reference, this can be obtained by checking that, for $\tilde{a}_n, \tilde{b}_n$ given by (4.2), $P(Z > x/\tilde{a}_n + \tilde{b}_n) \sim e^{-x}/n$. It follows immediately, see [5], Theorem 5.7.2, that $\tilde{N}_n \to_d \tilde{N}$, where $\tilde{N}$ is a Poisson process in $[0, \infty) \times \mathbb{R} = S$ whose intensity measure is the product of Lebesgue measure and the measure with density e^{-x} (i.e., in short notation, the intensity measure is $dt \times e^{-x}\,dx$).

Further, $\tilde{N}_n' \to_d N'$ and $\tilde{N}_n'' \to_d N'$, where N' is the point process in $S \times \mathbb{R}^\infty$ obtained by adjoining the point $y \in \mathbb{R}^\infty$ defined by $y_0 = 1$ and $y_\tau = 0$, $\tau \neq 0$ to each point of N. This of course corresponds to the obvious fact that for independent sequences extreme values have no influence on neighboring values, and it is easily proved (or obtained as a special case) by the same methods as used for $\{X_t\}$.

Similarly, for the $\{\hat{X}_t\}$ sequence, the only problem to be solved is to find $\hat{a}_n > 0$, $\hat{b}_n$ such that $P(\hat{a}_n(\hat{M}_n - \hat{b}_n) \le x) \to \exp\{-e^{-x}\}$, or equivalently such that $P(X_0 > x/\hat{a}_n + \hat{b}_n) \sim e^{-x}/n$ as $n \to \infty$, since the results for $\hat{N}_n$, $\hat{N}_n'$, and $\hat{N}_n''$ then follows trivially, in the same way as above.

5. Extremes of the associated independent sequence for $p > 1$. The main result of the companion paper [8] is that if assumptions B.1 or B.3 from Section 2 are satisfied then

$$\frac{P\left(\Sigma c_\lambda Z_\lambda > z + x/z^{p/q}\right)}{P(\Sigma c_\lambda Z_\lambda > z)} \to \exp\left\{-p\|c\|_q^{-p}x\right\} \quad \text{as } z \to \infty \tag{5.1}$$

for fixed x, and if instead B.2 holds then

$$\frac{P\left(\Sigma c_\lambda Z_\lambda > z + x/z^{p/q}\right)}{P(\Sigma c_\lambda Z_\lambda > z)} \to \exp\left\{-p\|c^+\|_q^{-p}x\right\} \quad \text{as } z \to \infty \tag{5.2}$$

with $\|c^+\|_q = \{\Sigma(c_\lambda^+)^q\}^{1/q}$, as defined in Section 2. By monotonicity, both relations remain valid if in the left-hand sides x is replaced by $x(z)$, with $x(z) \to x$ as $z \to \infty$.

Now define norming constants $\hat{a}_n > 0$, $\hat{b}_n$ by

$$\hat{a}_n = \begin{cases} p\|c\|_q^{-1}(\log n)^{1/q} & \text{if B.1 or B.3 holds,} \\ p\|c^+\|_q^{-1}(\log n)^{1/q} & \text{if B.2 holds,} \end{cases} \tag{5.3}$$

and by requiring that

$$P\left(\sum c_\lambda Z_\lambda > \hat{b}_n\right) \sim n^{-1} \quad \text{as } n \to \infty. \tag{5.4}$$

The $\hat{b}_n$'s are not completely determined by the conditions B.1–B.3, but below it will be seen that

$$\hat{b}_n = \begin{cases} \|c\|_q(\log n)^{1/p} + O\left((\log n)^{1/(\theta q) - 1/q}\right) & \text{if B.1 or B.3 holds,} \\ \|c^+\|_q(\log n)^{1/p} + O\left((\log n)^{\max(1/(\theta q),\, q'/q) - 1/q}\right) & \text{if B.2 holds.} \end{cases} \tag{5.5}$$

The type I limit for maxima of the associated i.i.d. sequence $\{\hat{X}_t\}$ with the same marginal d.f. as $\Sigma c_\lambda Z_\lambda$ now follows readily.

THEOREM 5.1. *Suppose that one of B.1–B.3 is satisfied, and let $\{\hat{a}_n, \hat{b}_n\}$ be as defined above. Then*

$$P\left(\hat{a}_n(\hat{M}_n - \hat{b}_n) \le x\right) \to \exp\{-e^{-x}\} \quad \text{as } n \to \infty. \tag{5.6}$$

PROOF. It is readily seen that (5.6) is equivalent to

$$nP\left(\sum c_\lambda Z_\lambda > x/\hat{a}_n + \hat{b}_n\right) \to e^{-x} \quad \text{as } n \to \infty \tag{5.7}$$

(cf. [5], Theorem 1.5.1). Suppose B.1 holds. Then, according to (5.1), since $x\hat{b}_n^{p/q}/\hat{a}_n \to x\|c\|_q^p/p$ as $n \to \infty$,

$$\begin{aligned} nP\left(\sum c_\lambda Z_\lambda > x/\hat{a}_n + \hat{b}_n\right) &> \frac{P\left(\sum c_\lambda Z_\lambda > \hat{b}_n + \left(x\hat{b}_n^{p/q}/\hat{a}_n\right)/\hat{b}_n^{p/q}\right)}{P(\sum c_\lambda Z_\lambda > \hat{b}_n)} \\ &\to \exp\left\{-p\|c\|_q^{-p}x\|c\|_q^p/p\right\} \\ &= \exp\{-x\} \quad \text{as } n \to \infty, \end{aligned}$$

so that (5.7) holds. The proofs under B.2 or B.3 are the same. □

The estimate in (5.5) is contained in the following lemma, which for later use is stated in a slightly more general form than needed here.

LEMMA 5.2. *Suppose that* B.1 *or* B.3 *holds. Then*

(i) $$P\left(\sum c_\lambda Z_\lambda > z\right) = \exp\left\{-\left(z/\|c\|_q\right)^p + O(z^\gamma)\right\} \quad \text{as } z \to \infty$$

for $\gamma = p/(\theta q)$, *and for any constant* $D > 0$ *this is uniform in all* $\{c_\lambda\}$ *satisfying* $|c_\lambda| \le D|\lambda|^{-\theta}$.

(ii) *If* $\{u_n\}$ *satisfies*

$$P\left(\sum c_\lambda Z_\lambda > u_n\right) \sim \tau/n \quad \text{as } n \to \infty$$

for some $\tau > 0$, *then*

$$u_n = \|c\|_q(\log n)^{1/p} + O\left((\log n)^{\gamma/p - 1/q}\right) \quad \text{as } n \to \infty.$$

(iii) *If instead* B.2 *is satisfied, then the conclusions of* (i) *and* (ii) *are still valid if* $\|c\|_q$ *is replaced by* $\|c^+\|_q$ *and if* γ *is defined as* $\gamma = p\max(1/(\theta q), q'/q)$ *for* $q' = (1 - 1/p')^{-1}$ *with* p' *given by* B.2.

PROOF. (i) and (iii). These are proved in [8], Section 6.

(ii) Suppose B.1 holds. Then, according to the assumption and part (i),

$$\tau/n \sim P\left(\sum c_\lambda Z_\lambda > u_n\right) = \exp\left\{-\left(u_n/\|c\|_q\right)^p + O(u_n^\gamma)\right\},$$

and thus

$$-\log n = -\left(u_n/\|c\|_q\right)^p + O(u_n^\gamma).$$

This shows that $u_n = O((\log n)^{1/p})$, so that

$$\begin{aligned} u_n &= \|c\|_q(\log n)^{1/p}\left(1 + (\log n)^{-1}O\left((\log n)^{\gamma/p}\right)\right) \\ &= \|c\|_q(\log n)^{1/p} + O\left((\log n)^{\gamma/p - 1/q}\right) \quad \text{as } n \to \infty. \end{aligned}$$

The proofs under B.2 and B.3 are similar. □

In concrete situations it would be desirable to have more precise estimates for $\hat{b}_n$ than (5.5), and one might perhaps be tempted to think that the appearance of the "big O" term is due to inaccuracies in the estimates. In a sense this is however not the case, since it can be seen that the assumptions B.1–B.3 only determine $\hat{b}_n$ up to terms of this order.

Nevertheless, there are cases when $\hat{b}_n$ can be explicitly computed. If the Z_λ's are normal with mean one and variance $\frac{1}{2}$, so that $f(z) = \exp\{-z^2\}/\sqrt{\pi}$, then one possible choice is

$$\hat{b}_n = \|c\|_2(\log n)^{1/2} - \|c\|_2(\log\log n + \log 4\pi)/\left(4(\log n)^{1/2}\right). \tag{5.8}$$

Further, in [8], Section 6 it is shown that if only finitely many, say $k > 0$, of the c_λ's are nonzero, and if B.1 holds, then

$$P\left(\sum c_\lambda Z_\lambda > z\right) \sim \hat{K}\left(z/\|c\|_q\right)^{\hat{\alpha}}\exp\left\{-\left(z/\|c\|_q\right)^p\right\} \quad \text{as } z \to \infty, \tag{5.9}$$

with

$$\begin{aligned} \hat{\alpha} &= k\left\{\left(\alpha' + \tfrac{1}{2}\right) - p/(2q)\right\} - p/2, \\ \hat{K} &= (K')^k(2\pi/g_2)^{(k-1)/2} p^{k\{(1-q/p)(\alpha'+1)-p/2\}-p/2} \\ &\quad \times \prod_\lambda \left(c_\lambda/\|c\|_q\right)^{(\alpha'+1/2)q/p-1/2}. \end{aligned} \tag{5.10}$$

As in Section 4 it then follows that $\hat{b}_n$ may be chosen as

$$\hat{b}_n = \|c\|_q(\log n)^{1/p} + \|c\|_q p^{-1}\left((\hat{\alpha}/p)\log\log n + \log \hat{K}\right)/(\log n)^{1-1/p}. \tag{5.11}$$

If instead B.2 or B.3 are satisfied, then (5.9)–(5.11) are replaced by slightly more complicated expressions, which we leave to the reader to derive. However, in the special case of B.2 when $f(z)$ is symmetric, (5.9)–(5.11) remain unchanged and in particular in the normal case discussed above, with $K = \pi^{-1/2}$, $\alpha = 0$, (5.11) reduces to (5.8) as it should.

Finally, even if $\hat{b}_n$ seems to be difficult to compute analytically in general, numerical computation should not be difficult.

6. Extremes of the moving average process for $p > 1$. Using the results of Sections 3 and 5 we in this section show that the maximum M_n of the moving average process $\{X_t = \sum c_{\lambda-t}Z_\lambda\}$ behaves asymptotically in the same way as $\hat{M}_n$. This is proved as a consequence of the more general result that the point process N_n of heights and locations of extremes converges to a Poisson process N in the plane with intensity $dt \times e^{-x}\,dx$ (Theorem 6.1) in the same way as for $\hat{N}_n$. However, of course the sample path behavior of $\{X_t\}$ and of $\{\hat{X}_t\}$ near extremes differ markedly. Let

$$y_\tau = \begin{cases} \sum_\lambda c_{\lambda-\tau}|c_\lambda|^{q/p}\operatorname{sign}(c_\lambda)/\|c\|_q^q & \text{if B.1 or B.3 holds,} \\ \sum_\lambda c_{\lambda-\tau}(c_\lambda^+)^{q/p}/\|c^+\|_q^q & \text{if B.2 holds,} \end{cases} \tag{6.1}$$

with $\operatorname{sign}(c_\lambda)$ equal to one if $c_\lambda \geq 0$, and to minus one otherwise, and let N' be obtained by adjoining the mark y to each point of N. Then for N_n', N_n'' as defined in Section 2, $N_n' \to_d N'$, $N_n'' \to N'$ (Theorem 6.3).

For these results, the norming constants are the same as for the associated independent sequence, i.e., we may use

$$a_n = \hat{a}_n, \qquad b_n = \hat{b}_n \tag{6.2}$$

with $\hat{a}_n$, $\hat{b}_n$ given by (5.3)–(5.5).

THEOREM 6.1. *Suppose that one of* B.1–B.3 *is satisfied, let* a_n, b_n *be as in* (6.2), *and let* N_n *be as defined in Section* 2. *Then* $N_n \to_d N$ *as* $n \to \infty$ *in* $[0, \infty) \times \mathbb{R}$, *where* N *is a Poisson process with intensity measure* $dt \times e^{-x}\,dx$. *In particular*,

$$P(a_n(M_n - b_n) \leq x) \to e^{-e^{-x}} \quad \text{as } n \to \infty. \tag{6.3}$$

PROOF. We will prove (3.4)–(3.6). Since the other assumptions of Lemma 3.2 clearly are satisfied [using Theorem 5.1 for (3.1)], this is sufficient to prove that $N_n \to_d N$.

Suppose now, to fix ideas, that B.1 holds. The proofs under B.2 and B.3 proceed similarly, as in Theorem 5.1 and will be left to the reader. According to Minkowski's inequality, $(\Sigma(c_\lambda + c_{\lambda-t})^q)^{1/q} < 2\|c\|_q$ for $t \neq 0$, since $\leq$ always holds and since equality would mean that $\{c_\lambda\}$ and $\{c_{\lambda-t}\}$ are proportional, which is impossible. Further e.g., by summing over $\lambda \leqq t/2$ and $\lambda > t/2$ separately it can be seen that $(\Sigma(c_\lambda + c_{\lambda-t})^q)^{1/q} \to 2^{1/q}\|c\|^q$ as $t \to \pm\infty$. Thus there exists a $\gamma' > 0$ such that

$$2\|c\|_q \Big/ \left(\sum (c_\lambda + c_{\lambda-t})^q\right)^{1/q} \geq 1 + \gamma' \quad \text{for } t \neq 0. \tag{6.4}$$

Let γ satisfy $0 < \gamma$ and $1 + \gamma < (1 + \gamma')^p$, and, as in Lemma 3.2, write $n' = [n^\gamma]$ and $u_n = x/a_n + b_n$.

By Lemma 5.2(i) and (6.4)

$$\begin{aligned} P(X_0 + X_t > 2u_n) &= P\left(\sum (c_\lambda + c_{\lambda-t}) Z_\lambda > 2u_n\right) \\ &= \exp\left\{-\left(\frac{2u_n}{\left(\Sigma(c_\lambda + c_{\lambda-t})^q\right)^{1/q}}\right)^p (1 + o(1))\right\} \\ &\leq \exp\left\{-(1 + \gamma')^p \left(\frac{u_n}{\|c\|_q}\right)^p (1 + o(1))\right\} \quad \text{as } n \to \infty, \end{aligned}$$

uniformly for $t \neq 0$. Since by (5.3) and (5.5) $u_n/\|c\|_q = (\log n)^{1/p}(1 + o(1))$ and since $1 + \gamma < (1 + \gamma')^p$, it follows that

$$P(X_0 + X_t > 2u_n) = o(n^{1+\gamma}) \quad \text{as } n \to \infty,$$

and hence (3.4) is satisfied.

Let $\Phi_{n'}(h) = E\exp\{h\sum_{n'<\lambda} c_\lambda Z_\lambda\}$ so that by Lemma 5.1(iii) of [8] and (2.9), with C a generic constant,

$$\Phi_{n'}(h) \le \exp\Big\{C \sum_{n'<\lambda} c_\lambda h\Big\} \le \exp\big\{C(n')^{1-\theta} h\big\}$$

for $h \le (n')^\theta$.

To prove the first part of (3.5), we insert this into Bernstein's inequality

$$P\Big(\sum_{n'<\lambda} c_\lambda Z_\lambda > z\Big) \le \Phi_{n'}(h)\exp\{-hz\}$$

for $z = 1/a_n$, $h = (n')^\theta$.

It follows that

$$\begin{aligned} P\Big(\sum_{n'<\lambda} c_\lambda Z_\lambda > 1/a_n\Big) &\le \exp\big\{Cn' - (n')^\theta/a_n\big\} \\ &= o(n^{-2}) \quad \text{as } n \to \infty, \end{aligned}$$

and hence the first part of (3.5) holds. The second part is completely similar.

Next, the first part of (3.6′) follows from $E|a_n\sum_{n'+1}^{\infty} c_\lambda Z_\lambda| \leqq a_n\sum_{n'+1}^{\infty}|c_\lambda|E|Z| \to 0$ as $n \to \infty$, and the second part is the same. This completes the proof of (3.4)–(3.6′), and hence of $N_n \to_d N$ as $n \to \infty$.

Finally, this implies in particular that $N_n((0,1] \times (x,\infty)) \to_d N((0,1] \times (x,\infty))$, and hence

$$\begin{aligned} P(a_n(M_n - b_n) \le x) &= P(N_n((0,1] \times (x,\infty)) = 0) \\ &\to P(N((0,1] \times (x,\infty)) = 0) \\ &= 1 \times \exp\Big\{-\int_x^\infty e^{-z}\,dz\Big\} \\ &= \exp\{-e^{-x}\} \quad \text{as } n \to \infty, \end{aligned}$$

so that (6.3) holds. □

The major step in finding the sample path behavior of $\{X_t\}$ near an extreme value is contained in the following lemma, which makes precise the "geometrical" heuristics in the introduction.

Lemma 6.2. *Let λ_0 be a fixed integer, let $\varepsilon > 0$ be arbitrary, and suppose that $u'/u \to 1$ as $u \to \infty$. If* B.1 *or* B.3 *is satisfied, then*

(6.5) $$P\Big(\big|Z_{\lambda_0} - u'|c_{\lambda_0}|^{q/p}\operatorname{sign}(c_{\lambda_0})/\|c\|_q^q\big| \le \varepsilon u' \,\Big|\, \sum c_\lambda Z_\lambda > u\Big) \to 1 \quad \text{as } u \to \infty,$$

and if B.2 *is satisfied then*

(6.6) $$P\Big(\big|Z_{\lambda_0} - u'(c_{\lambda_0}^+)^{q/p}/\|c^+\|_q^q\big| \le \varepsilon u' \,\Big|\, \sum_\lambda c_\lambda Z_\lambda > u\Big) \to 1 \quad \text{as } u \to \infty.$$

Proof. For notational convenience we will assume $\lambda_0 = 0$. By independence the result is obvious if $c_0 = 0$, so we may further assume that $c_0 \ne 0$. First suppose that B.1 holds, so that in particular $c_0 > 0$. Let

$$\bar{\beta} = \frac{c_0^{q/p}\|c\|_q^{-q} + \varepsilon}{c_0^{q/p}\|c\|_q^{-q}}, \qquad \underline{\beta} = \frac{c_0^{q/p}\|c\|_q^{-q} - \varepsilon}{c_0^{q/p}\|c\|_q^{-q}}.$$

Then (6.5) (for $\lambda_0 = 0$) is equivalent to the two relations

$$\frac{P\left(Z_0 > u'\bar{\beta}c_0^{q/p}/\|c\|_q^q, \Sigma c_\lambda Z_\lambda > u\right)}{P(\Sigma c_\lambda Z_\lambda > u)} \to 0 \tag{6.7}$$

and

$$\frac{P\left(Z_0 < u'\underline{\beta}c_0^{q/p}/\|c\|_q^q, \Sigma c_\lambda Z_\lambda > u\right)}{P(\Sigma c_\lambda Z_\lambda > u)} \to 0 \quad \text{as } u \to \infty. \tag{6.8}$$

Since the proofs of (6.7) and (6.8) are similar, we will only verify (6.7).

The result follows readily if $c_0/\|c\|_q = 1$ [e.g., from (5.1)] and hence we may assume that $0 < c_0/\|c\|_q < 1$, and then without loss of generality that $1 < \bar{\beta} < \|c\|_q^q/c_0^q$. Thus if we let β be a constant with $1 < \beta < \bar{\beta}$, and define

$$\bar{c}_\lambda = \begin{cases} c_0\beta^{p/q} & \text{if } \lambda = 0, \\ c_\lambda\left(\dfrac{1 - \beta c_0^q/\|c\|_q^q}{1 - c_0^q/\|c\|_q^q}\right)^{p/q} & \text{if } \lambda \neq 0, \end{cases}$$

then $\bar{c}_\lambda \geq 0$ for all λ. It is straightforward to check that

$$\left\{Z_0 > u\beta c_0^{q/p}/\|c\|_q^q, \sum c_\lambda Z_\lambda > u\right\} \subset \left\{\sum \bar{c}_\lambda Z_\lambda > u\|\bar{c}\|_q^q/\|c\|_q^q\right\}. \tag{6.9}$$

Since $u'/u \to 1$ by assumption, $u\beta < u'\bar{\beta}$ for all sufficiently large u, and hence for such u, using (6.9) for the second step and Lemma 5.2(i) for the third step,

$$\begin{aligned} &P\left(Z_0 > u'\bar{\beta}c_0^{q/p}/\|c\|_q^q, \sum c_\lambda Z_\lambda > u\right) \\ &\leq P\left(Z_0 > u\beta c_0^q/\|c\|_q^q, \sum c_\lambda Z_\lambda > u\right) \\ &\leq P\left(\sum \bar{c}_\lambda Z_\lambda > u\|\bar{c}\|_q^q/\|c\|_q^q\right) \\ &= \exp\left\{-\left(\frac{u\|\bar{c}\|_q^q}{\|c\|_q^q\|\bar{c}\|_q}\right)^p(1 + o(1))\right\} \quad \text{as } u \to \infty. \end{aligned}$$

Since $P(\Sigma c_\lambda Z_\lambda > u) = \exp\{-(u/\|c\|_q)^p(1 + o(1))\}$ [again by Lemma 5.2(i)], it follows that

$$\begin{aligned} &\frac{P\left(Z_0 > u'\bar{\beta}c_0^{q/p}/\|c\|_q^q, \Sigma c_\lambda Z_\lambda > u\right)}{P(\Sigma c_\lambda Z_\lambda > u)} \\ &= O\left(\exp\left\{-\left(\frac{u}{\|c\|_q}\right)^p\left(\frac{\|\bar{c}\|_q^q}{\|c\|_q^q} - 1\right)(1 + o(1))\right\}\right). \end{aligned} \tag{6.10}$$

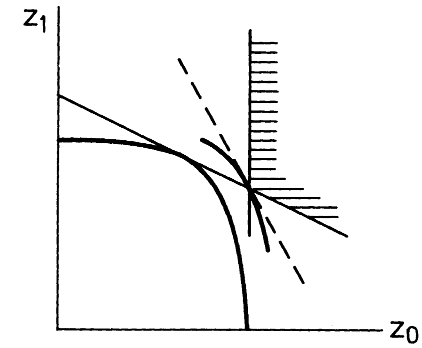

FIG. 2. *Probability of shaded area is approximated by probability of area outside dashed line. The curves are level curves of* $\exp\{-\|z\|_p^p\}$.

Here

$$\frac{\|\bar{c}\|_q^q}{\|c\|_q^q} = \left\{\beta^p c_0^q + \left(\frac{1-\beta c_0^q/\|c\|_q^q}{1-c_0^q/\|c\|_q^q}\right)^p \left(\|c\|_q^q - c_0^q\right)\right\}/\|c\|_q^q$$
$$= \beta^p c_0^q/\|c\|_q^q + \left(1-\beta c_0^q/\|c\|_q^q\right)^p/\left(1-c_0^q/\|c\|_q^q\right)^{p-1},$$

and since elementary calculations show that the function $g(\beta, x) = \beta^p x + (1-\beta x)^p/(1-x)^{p-1}$ is strictly greater than one for $0 < x < 1$ and $1 < \beta < 1/x$, we have that $\|\bar{c}\|_q^q/\|c\|_q^q > 1$, and (6.5) follows at once from (6.10).

(For a geometrical interpretation of this proof, see Figure 2.)

Next, suppose that instead B.3 holds. Then, replacing c_λ by $|c_\lambda|$ and Z_λ by $Z_\lambda \operatorname{sign}(c_\lambda)$ in the previous computations, the same result again ensues.

If B.2 holds, then $P(\sum c_\lambda Z_\lambda > u) = \exp\{-(u/\|c^+\|_q)^p(1+o(1))\}$ by Lemma 5.2(iii), and

$$P(Z_0 < -\varepsilon u') = \exp\{(\varepsilon u')^{p'}(1+o(1))\} = o\left(\exp\left\{-\left(u/\|c^+\|_q\right)^p(1+o(1))\right\}\right)$$

since $p' > p$, and hence

$$P\left(Z_0 < -\varepsilon u' \middle| \sum c_\lambda Z_\lambda > u\right) \to 0 \quad \text{as } u \to \infty. \tag{6.11}$$

Further, if $c_0 < 0$, so that $\{Z_0 > \varepsilon u'\} = \{-c_0 Z > |c_0|\varepsilon u'\}$, then

$$P\left(Z_0 > \varepsilon u', \sum c_\lambda Z_\lambda > u\right) \le P\left(\sum_{\lambda \ne 0} c_\lambda Z_\lambda > u + |c_0|\varepsilon u'\right)$$
$$= \exp\left\{-\left(\frac{u+|c_0|\varepsilon u'}{\|c^+\|_q}\right)^p (1+o(1))\right\},$$

and hence, similarly as above,

$$P\left(Z_0 > \varepsilon u' \middle| \sum c_\lambda Z_\lambda > u\right) \to 0 \quad \text{as } u \to \infty. \tag{6.12}$$

Together (6.11) and (6.12) prove (6.6) for the case $c_0 < 0$. Finally, if $c_0 \geq 0$, (6.6) follows from similar calculations as for hypothesis B.1, after replacing $\|c\|_q$ and $\|\bar{c}\|_q$ by $\|c^+\|_q$ and $\|\bar{c}^+\|_q$ throughout (with obvious notation). $\square$

We will only prove convergence of sample paths near extremes under the hypotheses B.2 and B.3. The corresponding result surely holds also if B.1 is satisfied, but it seems a proof of this would require further complications in an already long proof.

THEOREM 6.3. *Suppose that* B.2 *or* B.3 *holds, and let* N_n' *and* N_n'' *be as defined in Section* 2 *with* a_n, b_n *given by* (6.2). *Then* $N_n' \to_d N'$ *and* $N_n'' \to_d N'$ *as* $n \to \infty$ *in* $S \times \mathbb{R}^\infty$, *where* N' *is the point process obtained by adjoining the mark* y *given by* (6.1) *to each point of the Poison process* N *in* $[0, \infty) \times \mathbb{R} = S$, *with intensity measure* $dt \times e^{-x}\,dx$.

PROOF. According to Lemma 3.4, to prove $N_n' \to_d N'$ it is sufficient to prove (3.11). Let $u_n = x/a_n + b_n$ for fixed x so that $b_n/u_n \to 1$ as $n \to \infty$ by (5.3), (5.5), and $P(X_0 > u_n) = P(\Sigma c_\lambda Z_\lambda > u_n) \sim e^{-x}/n$, as noted above. Suppose B.3 holds. Then by Lemma 6.2 with $u = u_n$, $u' = b_n$ for any $\varepsilon > 0$ and λ_0,

$$P\left(X_0 > u_n, \left|Z_{\lambda_0} - b_n |c_{\lambda_0}|^{q/p} \operatorname{sign}(c_{\lambda_0})/\|c\|_q^q\right| > \varepsilon b_n\right) = o(1/n) \quad \text{as } n \to \infty.$$

It readily follows that, for any $\bar{\lambda} \geq 0$ and $\varepsilon > 0$,

$$P\left(\{X_0 > u_n\} \cap \bigcup_{|\lambda| \leq \bar{\lambda}} \left\{\left|Z_\lambda - b_n |c_\lambda|^{q/p} \operatorname{sign}(c_\lambda)/\|c\|_q^q\right| > \varepsilon b_n\right\}\right)$$
$$= o(1/n) \quad \text{as } n \to \infty,$$

and then that, for fixed τ,

$$P\left(X_0 > u_n, \left|b_n^{-1} \sum_{|\lambda| \leq \bar{\lambda}} c_{\lambda-\tau} Z_\lambda - \sum_{|\lambda| \leq \bar{\lambda}} c_{\lambda-\tau} |c_\lambda|^{q/p} \operatorname{sign}(c_\lambda)/\|c\|_q^q\right| > \varepsilon\right) \tag{6.13}$$
$$= o(1/n).$$

Now for fixed $\varepsilon > 0$, choose $\bar{\lambda}$ large enough to make $|\Sigma_{|\lambda| > \bar{\lambda}} c_{\lambda-\tau} |c_\lambda|^{q/p} \operatorname{sign}(c_\lambda)| < \varepsilon \|c\|_q^q$ and $(\Sigma_{|\lambda| > \bar{\lambda}} |c_\lambda|^q)^{1/q} < \varepsilon \|c\|_q$. Then, using the definitions of $Y_{n,0}'(\tau)$ and y_τ

$$P\left(X_0 > u_n, |Y_{n,0}'(\tau) - y_\tau| > 3\varepsilon\right)$$
$$\leq P\left(X_0 > u_n, \left|b_n^{-1} \Sigma c_{\lambda-\tau} Z_\lambda - \sum_{|\lambda| \leq \bar{\lambda}} c_{\lambda-\tau} |c_\lambda|^{q/p}/\|c\|_q^q\right| > 2\varepsilon\right)$$
$$\leq P\left(X_0 > u_n, \left|b_n^{-1} \sum_{|\lambda| \leq \bar{\lambda}} c_{\lambda-\tau} Z_\lambda - \sum_{|\lambda| \leq \bar{\lambda}} c_{\lambda-\tau} |c_\lambda|^{q/p}/\|c\|_q^q\right| > \varepsilon\right) \tag{6.14}$$
$$+ P\left(\left|\sum_{|\lambda| > \bar{\lambda}} c_{\lambda-\tau} Z_\lambda\right| > b_n \varepsilon\right)$$
$$= o(1/n) + P\left(\left|\sum_{|\lambda| > \bar{\lambda}} c_{\lambda-\tau} Z_\lambda\right| > b_n \varepsilon\right) \quad \text{as } n \to \infty$$

by (6.13). It follows from Lemma 5.2(i) and (5.5) that

$$P\Big(\sum_{|\lambda|>\bar{\lambda}} c_{\lambda-\tau}Z_\lambda > b_n\varepsilon\Big) = \exp\left\{-\left(\frac{b_n\varepsilon}{\left(\sum_{|\lambda|>\bar{\lambda}}|c_\lambda|^q\right)^{1/q}}\right)^p(1+o(1))\right\}$$

$$(6.15)\qquad = \exp\left\{-\log n\left(\frac{\|c\|_q\varepsilon}{\left(\sum_{|\lambda|>\bar{\lambda}}|c_\lambda|^q\right)^{1/q}}\right)^p(1+o(1))\right\}$$

$$= o(1/n) \quad \text{as } n\to\infty,$$

since $\bar{\lambda}$ was chosen to make $\|c\|_q\varepsilon/(\sum_{|\lambda|>\bar{\lambda}}|c_\lambda|^q)^{1/q}) > 1$. Similarly

$$(6.16)\qquad P\Big(\sum c_{\lambda-\tau}Z_\lambda < -b_n\varepsilon\Big) = o(1/n) \quad \text{as } n\to\infty.$$

Thus, by (6.14)–(6.16) for any $\varepsilon > 0$ and τ

$$(6.17)\qquad P\big(X_0 > u_n, |Y'_{n,0}(\tau) - y_\tau| > 3\varepsilon\big) = o(1/n) \quad \text{as } n\to\infty,$$

which proves (3.11) and hence that $N'_n \to_d N'$ if B.3 holds.

Further, since $y_0 = 1$ and $Y'_{n,0}(0) = X_0/b_n$, it follows from (6.17), replacing 3ε by ε, that

$$P\big(X_0 > u_n, |X_0/b_n - 1| > \varepsilon\big) = o(1/n) \quad \text{as } n\to\infty,$$

and then by easy arguments that, for $\varepsilon > 0$ and τ fixed,

$$P\big(X_0 > u_n, |Y''_{n,0}(\tau) - y_\tau| > \varepsilon\big) = o(1/n) \quad \text{as } n\to\infty.$$

By Lemma 3.4, with N'_n replaced by N''_n, and $Y'_{n,0}$ by $Y''_{n,0}$, this shows that $N''_n \to_d N'$ as $n\to\infty$ if B.3 holds.

The proof under assumption B.2 is similar and is left to the reader. □

7. Extremes for $p = 1$. The extremal behavior and the technique needed to study it is less complex for $p = 1$ than for $p > 1$, although there is an interesting extra diversity of behavior when the weights $\{c_\lambda\}$ assume their maximum for more than one value of λ. We will therefore be briefer than in the previous sections, leaving arguments to the reader and excluding some cases which could be treated by similar methods, but at the cost of further complications.

In each of the cases A.1–A.3 we will find the appropriate norming constants $\hat{a}_n$, $\hat{b}_n$ for the maximum $\hat{M}_n$ of the associated independent sequence $\{X_t\}$ (Theorem 7.3). The corresponding results for the maximum M_n of the moving average process, and for the point process N_n will, for $\alpha > -1$ also be proved in all three cases, but for $\alpha < -1$ only when $k_+ = 1$ and in the cases A.1 and A.2 of positive weights and of a dominating right tail, respectively (Theorem 7.4). In those cases, as for $p > 1$, the norming constants and limits are the same as for $\{\hat{X}_t\}$.

Similarly, proofs concerning sample paths near extremes are only given for cases A.1 and A.2 with $k_+ = 1$ (Theorem 7.5). Some of the remaining cases, which more resemble $0 < p < 1$, are discussed at the end of the section, without proofs.

The first lemma of this section contains some straightforward estimates of convolution integrals and will, again quite straightforwardly, lead to the tail behavior of $\Sigma c_\lambda Z_\lambda$ for $p = 1$.

LEMMA 7.1. (i) *Suppose the random variable* Y_1 *satisfies* (2.2), *with* $p = 1$, *and is independent of* Y_2 *which satisfies*

$$Ee^{\beta Y_2} < \infty \quad \textit{for some } \beta > 1.$$

Then

$$P(Y_1 + Y_2 > z) \sim KEe^{Y_2} z^\alpha e^{-z} \quad \textit{as } z \to \infty. \tag{7.1}$$

Furthermore, for fixed Y_1, $C > 0$ *and* $\beta > 1$ *the relation* (7.1) *is uniform in* $Y_2 \in \{Y;\, Ee^{\beta Y} \le C\}$.

(ii) *Let* Y_1 *and* Y_2 *be as in* (i). *Then*

$$\limsup_{z \to \infty} P\big(|Y_1 - z| > A \,|\, Y_1 + Y_2 > z\big) \to 0 \quad \textit{as } A \to \infty. \tag{7.2}$$

(iii) *Suppose that* Y_1 *and* Y_2 *are independent and satisfy* (2.2) *with* $p = 1$, *but with* K, α *replaced by* K_1, α_1 *and* K_2, α_2, *respectively. Then if* $-1 > \alpha_1 = \alpha_2 = \alpha$, *say,*

$$P(Y_1 + Y_2 > z) \sim \big(K_1 Ee^{Y_2} + K_2 Ee^{Y_1}\big) z^\alpha e^{-z} \quad \textit{as } z \to \infty, \tag{7.3}$$

and if $\alpha_1 > -1$, $\alpha_2 > -1$, *then*

$$P(Y_1 + Y_2 > z) \sim K_1 K_2 \Gamma(\alpha_1 + 1)\Gamma(\alpha_2 + 1)\Gamma(\alpha_1 + \alpha_2 + 2)^{-1} z^{\alpha_1 + \alpha_2 + 1} e^{-z}$$
$$\textit{as } z \to \infty, \quad \textit{for } \Gamma(\alpha) = \textstyle\int_0^\infty y^{\alpha - 1} e^{-y}\, dy \ (\alpha > 0). \tag{7.4}$$

PROOF. (i) Let γ be a fixed number, with $1/\beta < \gamma < 1$, and let F_1 and F_2 be the distribution functions of Y_1 and Y_2. Then

$$P(Y_1 + Y_2 > z) = \int P(Y_1 > z - x) F_2(dx)$$
$$= Kz^\alpha e^{-z} \int_{-\infty}^{\gamma z} \frac{P(Y_1 > z - x)}{Kz^\alpha e^{-z}} F_2(dx) + \int_{\gamma z}^{\infty} P(Y_1 > z - x) F_2(dx).$$

Here, since $P(Y_2 > \gamma z) \le Ee^{\beta Y_2} e^{-\beta \gamma z}$ by Bernstein's inequality,

$$\int_{\gamma z}^{\infty} P(Y_1 > z - x) F_2(dx) \le P(Y_2 > \gamma z) = O(e^{-\beta \gamma z})$$
$$= o(z^\alpha e^{-z}) \quad \text{as } z \to \infty,$$

since $\beta\gamma > 1$. Further, by (2.2) and dominated convergence

$$\int_{-\infty}^{\gamma z} \frac{P(Y_1 > z - x)}{K z^{\alpha} e^{-z}} F_2(dx) \sim \int_{-\infty}^{\gamma z} (1 - x/z)^{\alpha} e^x F_2(dx)$$

$$\to \int e^x \, dF_2(x) \quad \text{as } z \to \infty,$$

since the integrand tends pointwise to e^x, and, for $z \geq 1$, is bounded by a constant times $(1 + |x|^{\alpha})e^x$, which is integrable [since $P(Y_2 > x) = O(e^{-\beta x})$]. This proves (7.1), and the uniformity is then obtained by inspection of the proof.

(ii) Clearly

$$\begin{aligned} &P(|Y_1 - z| > A, Y_1 + Y_2 > z) \\ &\quad = \int P(|Y_1 - z| > A, Y_1 > z - x) F_2(dx) \\ &\quad \leq P(Y_1 > z + A) P(Y_2 > -A) + \int_{-\infty}^{-A} P(Y_1 > z - x) F_2(dx) \\ &\qquad + \int_A^{\infty} P(Y_1 > z - x) F_2(dx). \end{aligned}$$

Reasoning as in (i), we have that

$$\int_{|x| \geq A} P(Y_1 > z - x) F_2(dx) \sim K z^{\alpha} e^{-z} \int_{|x| \geq A} e^x F_2(dx),$$

and hence, using (2.2) to estimate $P(Y_1 > z + A)$ and part (i) to estimate $P(Y_1 + Y_2 > z)$, that

$$\limsup_{z \to \infty} P\big(|Y_1 - z| > A \,|\, Y_1 + Y_2 > z\big) \leq \left\{ e^{-A} + \int_{|x| \geq A} e^x F_2(dx) \right\} \Big/ \int e^x F_2(dx).$$

Clearly the right-hand side tends to zero as $A \to \infty$ which proves (ii).

(iii) It is readily seen that

$$\begin{aligned} P(Y_1 + Y_2 > z) &= \int_{-\infty}^{z/2} P(Y_1 > z - x) F_2(dx) + \int_{-\infty}^{z/2} P(Y_2 > z - x) F_1(dx) \\ &\quad + P(Y_1 > z/2) P(Y_2 > z/2) \\ &= \int_{-\infty}^{z/2} P(Y_1 > z - x) F_2(dx) + \int_{-\infty}^{z/2} P(Y_2 > z - x) F_1(dx) \\ &\quad + O(z^{\alpha_1 + \alpha_2} e^{-z}) \quad \text{as } z \to \infty \end{aligned} \tag{7.5}$$

by (2.2). Here

$$\int_{-\infty}^{z/2} P(Y_1 > z - x) F_2(dx) \sim K_1 z^{\alpha_1} e^{-z} \int_{-\infty}^{z/2} (1 - x/z)^{\alpha_1} e^x F_2(dx), \tag{7.6}$$

and if $\alpha_2 < -1$,

$$\int_{-\infty}^{z/2} (1 - x/z)^{\alpha_1} e^x F_2(dx) \to \int e^x F_2(dx) \quad \text{as } z \to \infty \tag{7.7}$$

by dominated convergence. Together with the same computations for the last integral in (7.5), the relations (7.5)–(7.7) prove (7.3).

If $\alpha_2 > -1$ then $\int_{-\infty}^{z/2}(1 - x/z)^{\alpha_1} e^x F_2(dx)$ tends to infinity, while $\int_{-\infty}^{0}(1 - x/z)^{\alpha_1} e^x F_2(dx)$ is bounded, and thus, using partial integration in the second step, and (2.2) in the third one, we have that,

$$\begin{aligned}
&\int_{-\infty}^{z/2} (1 - x/z)^{\alpha_1} e^x F_2(dx) \\
&\quad \sim \int_0^{z/2} (1 - x/z)^{\alpha_1} e^x F_2(dx) \\
&\quad = 1 - F_2(0) - 2^{-\alpha_1} e^{-z/2}(1 - F_2(z/2)) \\
&\qquad + \int_0^{z/2} \left\{ (1 - x/z)^{\alpha_1} - (\alpha_1/z)(1 - x/z)^{\alpha_1 - 1} e^x (1 - F_2(x))\, dx \right\} \\
&\quad \sim K_2 \int_0^{z/2} (1 - x/z)^{\alpha_1} x^{\alpha_2}\, dx \\
&\quad = K_2 z^{\alpha_2 + 1} \int_0^{1/2} (1 - y)^{\alpha_1} y^{\alpha_2}\, dy \quad \text{as } z \to \infty.
\end{aligned}$$

Now, insert this into (7.6), and then the result into (7.5), together with the corresponding formula for the last integral in (7.5) to yield that

$$P(Y_1 + Y_2 > z) \sim K_1 K_2 \int_0^1 (1 - y)^{\alpha_1} y^{\alpha_2}\, dy\, z^{\alpha_1 + \alpha_2 + 1} e^{-z} \quad \text{as } z \to \infty,$$

and since $\int_0^1 (1 - y)^{\alpha_1} y^{\alpha_2}\, dy = \Gamma(\alpha_1 + 1)\Gamma(\alpha_2 + 1)/\Gamma(\alpha_1 + \alpha_2 + 2)$, this is the same as (7.4). □

Here, in part (iii) we have for simplicity not included the case $\alpha_1 = \alpha_2 = -1$, which could be treated similarly, but with further complications involving logarithmic terms. Below we will accordingly exclude such cases.

To state the next lemma, on the tail behavior of $\Sigma c_\lambda Z_\lambda$, some further notation is needed. With c_+, c_-, Λ_+, Λ_-, k_+, and k_- as defined in Section 2, let

$$k = \begin{cases} k_+ & \text{if A.1 or A.2 holds,} \\ k_+ + k_- & \text{if A.3 holds,} \end{cases}$$

and let

$$\Lambda = \begin{cases} \Lambda_+ & \text{if A.1 or A.2 holds,} \\ \Lambda_+ \cup \Lambda_- & \text{if A.3 holds.} \end{cases}$$

With this notation, define

$$\hat{\alpha} = \begin{cases} k\alpha + k - 1 & \text{if } \alpha > -1 \\ \alpha & \text{if } \alpha < -1, \end{cases} \tag{7.8}$$

and

$$
(7.9)\quad \hat{K} = \begin{cases} K^k\Gamma(\alpha+1)^k\Gamma(k(\alpha+1))^{-1}E\exp\left\{\sum_{\lambda\notin\Lambda} c_\lambda Z_\lambda/c_+\right\} \\ \qquad\qquad \text{if A.1 or A.2 holds and } \alpha > -1, \\ K^{k_+}\left(K_-\gamma^{\alpha/p}\right)^{k_-}\Gamma(\alpha+1)^k\Gamma(k(\alpha+1))^{-1}E\exp\left\{\sum_{\lambda\notin\Lambda} c_\lambda Z_\lambda/c_+\right\} \\ \qquad\qquad \text{if A.3 holds and } \alpha > -1, \\ kK(Ee^Z)^{k-1}E\exp\left\{\sum_{\lambda\notin\Lambda} c_\lambda Z_\lambda/c_+\right\} \\ \qquad\qquad \text{if A.1 or A.2 holds and } \alpha < -1, \\ \left\{k_+K(Ee^Z)^{k_+-1}\left(Ee^{-Z/c_-}\right)^{k_-} \right. \\ \quad \left. + k_-K_-\gamma^{\alpha/p}(Ee^Z)^{k_+}\left(Ee^{-Z/c_-}\right)^{K_--1}\right\} \\ \quad \times E\exp\left\{\sum_{\lambda\notin\Lambda} c_\lambda Z_\lambda/c_+\right\} \qquad \text{if A.3 holds and } \alpha < -1. \end{cases}
$$

LEMMA 7.2. *Suppose that one of the assumptions* A.1–A.3 *is satisfied with* $p = 1$ *and* $\alpha \neq -1$. *Then, with* $\hat{\alpha}$, $\hat{K}$ *given by* (7.8), (7.9)

$$
(7.10)\qquad P\left(\sum c_\lambda Z_\lambda > z\right) \sim \hat{K}(z/c_+)^{\hat{\alpha}}e^{-z/c_+} \quad \text{as } z \to \infty.
$$

PROOF. Since $P(\Sigma c_\lambda Z_\lambda > z) = P(\Sigma(c_\lambda/c_+)Z_\lambda > z/c_+)$, we may without loss of generality assume that $c_+ = 1$. Suppose first A.1 holds with $\alpha \neq -1$. Let $\bar{c} = \max\{c_\lambda;\ \lambda \notin \Lambda_+\} < 1$. Clearly $\psi(h) = E\exp\{hZ\}$ is finite for $0 \le h < 1$, and $\psi(h) = 1 + hEZ(1 + o(1))$ as $h \to 0$, and for any $\beta < 1/\bar{c}$ it follows from (2.8) that $\Pi_{\lambda\notin\Lambda_+}E\exp\{\beta c_\lambda Z_\lambda\} = \Pi_{\lambda\notin\Lambda_+}(1 + \beta c_\lambda EZ(1 + o(1))$ is convergent, and hence

$$
(7.11)\qquad E\exp\left\{\beta\sum_{\lambda\notin\Lambda_+} c_\lambda Z_\lambda\right\} = \prod_{\lambda\notin\Lambda_+} E\exp\{\beta c_\lambda Z_\lambda\} < \infty.
$$

The result then follows immediately by writing $\Sigma c_\lambda Z_\lambda = Z_{\lambda_1} + \cdots + Z_{\lambda_{k_+}} + \Sigma_{\lambda\notin\Lambda_+}c_\lambda Z_\lambda$, and first applying Lemma 7.1(iii) repeatedly to evaluate the tail of the distribution of $Z_{\lambda_1} + \cdots + Z_{\lambda_{k_+}}$, and then Lemma 7.1(i), using (7.11) for some $\beta \in (1, 1/\bar{c})$, to establish (7.10) (remember that in this case $k = k_+$, $\Lambda = \Lambda_+$).

If instead A.2 holds, (7.10) again follows by the same argument, but with $\bar{c}$ defined as $\bar{c} = \max\{c_\lambda^+, c_\lambda^-/c_-;\ \lambda \notin \Lambda\}$.

Finally, the case A.3 follows similarly, after writing $c_\lambda Z_\lambda$ as $c_\lambda^-\gamma^{1/p}(-\gamma^{-1/p}Z_\lambda)$ for negative c_λ's, after noting that, by A.3, $P(-\gamma^{1/p}Z_\lambda > z) \sim K_-\gamma^{\alpha/p}z^\alpha e^{-z^p}$ as $z \to \infty$. □

The type I limit for $\hat{M}_n$, the maximum of the associated independent sequence is an immediate consequence of (7.10), by the same argument as for (4.1). Let

$$\begin{aligned} \hat{a}_n &= 1/c_+, \\ \hat{b}_n &= c_+ \log n + c_+(\hat{\alpha} \log\log n + \log \hat{K}). \end{aligned} \tag{7.12}$$

Theorem 7.3. *Suppose that one of* A.1–A.3 *is satisfied with* $p = 1$ *and* $\alpha \neq -1$ *and let* $\hat{a}_n$, $\hat{b}_n$ *be given by* (7.12). *Then*

$$P(\hat{a}_n(\hat{M}_n - \hat{b}_n) \leq x) \to e^{-e^{-x}} \quad \text{as } n \to \infty.$$

The behavior of extremes of the moving average process $\{X_t = \sum c_{\lambda-t} Z_\lambda\}$ is qualitatively different when $\alpha > -1$ and $\alpha < -1$. Here we will only treat the cases $\alpha > -1$ and $\alpha < -1$, $k = 1$ formally, with k as defined above. The remaining case, $\alpha < -1$, $k > 1$ is similar to the case $p < 1$, but with some added complexity. It will be treated separately in a later paper, as an example of a general convergence theorem, and will only be commented on briefly here.

Theorem 7.4. *Suppose that one of* A.1–A.3 *holds with* $p = 1$, *and that in addition either* $\alpha > -1$ *or* $\alpha < -1$ *and* $k = 1$. *Further let* $a_n = \hat{a}_n$, $b_n = \hat{b}_n$ *be given by* (7.12) *and let* N_n *and* M_n *be as defined in Section* 2. *Then* $N_n \to_d N$ *as* $n \to \infty$ *in* $[0, \infty) \times \mathbb{R}$, *where* N *is a Poisson process with intensity measure* $dt \times e^{-x}\,dx$. *In particular*

$$P(a_n(M_n - b_n) \leq x) \to e^{-e^{-x}} \quad \text{as } n \to \infty.$$

Proof. By Lemma 3.2 we only have to establish (3.4)–(3.6), similarly as for Theorem 6.1. Furthermore as before we will, without loss of generality, assume that $c_+ = 1$ so that also $a_n \equiv 1$.

Suppose now that A.1 holds with $k = k_+ = 1$. Then $\bar{c} = \max_{t \geq 1} \max_\lambda (c_\lambda + c_{\lambda-t}) < 2$, and we may choose $\beta > \frac{1}{2}$ with $\bar{c}\beta < 1$ and hence with $E\exp\{\beta(c_\lambda + c_{\lambda-t})Z_\lambda\} = \psi(\beta(c_\lambda + c_{\lambda-t}))$ well defined for $t \geq 1$ and all λ. For such t,

$$\begin{aligned} &E\exp\{\beta(X_0 + X_t)\} \\ &\quad = E\exp\Big\{\sum \beta(c_\lambda + c_{\lambda-t})Z_\lambda\Big\} \\ &\quad = \left\{\prod_{-\infty}^{[t/2]} \psi(\beta(c_\lambda + c_{\lambda-t}))\right\}\left\{\prod_{[t/2]+1}^{\infty} \psi(\beta(c_\lambda + c_{\lambda-t}))\right\}. \end{aligned} \tag{7.13}$$

Here $\psi'(h) = EZ\exp\{hZ\}$ is bounded, and $\psi(h)$ is bounded away from zero for $0 \leq h \leq \bar{c}\beta$ so that $C = \sup\{|\psi'(h + x)/\psi(h)|; 0 \leq h + x \leq \bar{c}\beta, h > 0, x > 0\} < \infty$. Hence, by the mean value theorem, $\psi(h_1 + h_2) \leq \psi(h_1)(1 + Ch_2)$ for $0 \leq h_1, h_2$ and $h_1 + h_2 \leq \bar{c}\beta$. Thus

$$\begin{aligned} \prod_{-\infty}^{[t/2]} \psi(\beta(c_\lambda + c_{\lambda-t})) &\leq \left\{\prod_{-\infty}^{[t/2]} \psi(\beta c_\lambda)\right\}\left\{\prod_{-\infty}^{[t/2]} (1 + C\beta c_{\lambda-t})\right\} \\ &\leq \left\{\prod_{-\infty}^{[t/2]} \psi(\beta c_\lambda)\right\}\left\{\prod_{-\infty}^{-[t/2]} (1 + C\beta c_\lambda)\right\}, \end{aligned}$$

which is bounded, uniformly in t by (2.8). Together with a similar computation for the second product in (7.13) this shows that $E\exp\{\beta(X_0 + X_t)\}$ is bounded, uniformly in $t \geq 1$. Choose $\gamma > 0$ with $1 + \gamma < 2\beta$, and for fixed x let $u_n = x/a_n + b_n$ so that $u_n \sim \log n$ as $n \to \infty$. Then, by Bernstein's inequality

$$\begin{aligned} P(X_0 + X_t > 2u_n) &\leq E\exp\{\beta(X_0 + X_t)\}e^{-2\beta u_n} \\ &= O(e^{-2\beta u_n}) \\ &= o(n^{-(1+\gamma)}) \quad \text{as } n \to \infty, \end{aligned}$$

uniformly in $t \geq 1$, which proves (3.4).

To prove (3.5) it is by the same inequality sufficient to show that e.g., $E\exp\{(\log n)^3\sum_{n'+1}^{\infty} c_\lambda Z_\lambda\}$ and $E\exp\{(\log n)^3\sum_{-\infty}^{-n'-1} c_\lambda Z_\lambda\}$ are bounded as $n \to \infty$ for $n' = [n^\gamma]$. However, this follows readily from (2.2) and (2.8) since $\psi(h) - 1 \sim hEZ$ as $h \to 0$.

Finally, by the same arguments as in Lemma 7.2, using the uniformity in Lemma 7.1(i), it follows that $P(\sum_{-\infty}^{n'} c_\lambda Z_\lambda > u_n) \sim \hat{K}u_n^{\hat{\alpha}}e^{-u_n}$ as $n \to \infty$, which, by the choice of u_n, proves the first part of (3.6). The second part is the same, so this concludes the proof for the case when A.1 holds and $k = 1$ $(= k_+)$.

The proof when A.2 holds and $k = 1$ is similar, while A.3 and A.1 and A.2 for $\alpha > -1$, $k > 1$ leads to an additional complication in the estimation of $P(X_0 + X_t > 2u_n)$ for small t. However, we omit the details of this. □

The behavior of sample paths near extremes is simplest if A.1 or A.2 holds with $k = k_+ = 1$. For these cases, let the limiting marks $y = \{y_\tau\}_{\tau=-\infty}^{\infty}$ be defined by

$$y_\tau = c_{\lambda_1 - \tau}/c_+, \qquad \tau = 0, \pm 1, \dots. \tag{7.14}$$

Theorem 7.5. *Suppose that* A.1 *or* A.2 *holds with* $k = 1$, *and let* N_n' *and* N_n'' *be as defined in Section* 2 *with* $a_n = \hat{a}_n$, $b_n = \hat{b}_n$ *given by* (7.12). *Then* $N_n' \to_d N'$ *and* $N_n'' \to_d N'$ *as* $n \to \infty$ *in* $S \times \mathbb{R}^\infty$, *where* N' *is the point process obtained by adjoining the mark* y *given by* (7.14) *to each point of the Poisson process* N *in* $[0, \infty) \times \mathbb{R} = S$ *with intensity measure* $dt \times e^{-x}\,dx$.

Proof. To establish that $N_n' \to_d N$ it is by Lemma 3.4 and Theorem 7.4 sufficient to prove (3.11). Suppose that A.1 holds and $k = k_+ = 1$. As usual we may assume that $c_+ = 1$ so that $c_\lambda < 1$ for $\lambda \neq \lambda_1$. Let $u_n = x/a_n + b_n$ for x fixed, and let $\varepsilon > 0$ be given. For $A > 0$, using independence in the second inequality, we have that

$$\begin{aligned} &P\big(X_0 > u_n, |Y_{n,0}'(\tau) - y_\tau| > \varepsilon\big) \\ &\quad\leq P\big(X_0 > u_n, |Z_{\lambda_1} - u_n| > A\big) \\ &\qquad + P\Big(|Z_{\lambda_1} - u_n| \leq A, \Big|\sum c_{\lambda-\tau}Z_\lambda - b_n c_{\lambda_1-\tau}\Big| > \varepsilon b_n\Big) \\ &\quad\leq P\big(X_0 > u_n, |Z_{\lambda_1} - u_n| > A\big) + P\big(|Z_{\lambda_1} - u_n| \leq A\big) \\ &\qquad \times P\Bigg(\Bigg|\sum_{\lambda \notin \lambda_1} c_{\lambda-\tau}Z_\lambda\Bigg| > \varepsilon b_n - A - |x|\Bigg). \end{aligned} \tag{7.15}$$

Here, by (2.2) and the choice of u_n, $nP(|Z_{\lambda_1} - u_n| \le A)$ tends to a finite constant as $n \to \infty$ and $P(|\sum_{\lambda \notin \lambda_1} c_{\lambda-\tau} Z_\lambda| > \varepsilon b_n - A - |x|) \to 0$ since b_n tends to infinity, so that the last term in (7.15) is $o(1/n)$ as $n \to \infty$. Furthermore, writing $X_0 = Z_{\lambda_1} + \sum_{\lambda \neq \lambda_1} c_\lambda Z_\lambda$, the assumptions of Lemma 7.1(ii) are satisfied for $Y_1 = Z_{\lambda_1}$ and $Y_2 = \sum_{\lambda \neq \lambda_1} c_\lambda Z_\lambda$ by Lemma 7.2. Thus, since $P(X_0 > u_n) \sim e^{-x}/n$ as $n \to \infty$ (by Theorem 7.3 and [5], Theorem 1.5.1),

$$\limsup_{n\to\infty} nP\big(X_0 > u_n, |Z_{\lambda_1} - u_n| > A\big) = e^{-x} \limsup_{n\to\infty} P\big(|Z_{\lambda_1} - u_n| > A | X_0 > u_n\big)$$
$$\to 0 \quad \text{as } A \to \infty.$$

By (7.15) this proves that

$$nP\big(X_0 > u_n, |Y'_{n,0}(\tau) - y_\tau| > \varepsilon\big) \to 0 \quad \text{as } n \to \infty,$$

i.e., (3.11) holds and hence $N'_n \to_d N'$ The proof that $N''_n \to_d N'$ then is the same as for Theorem 6.3, which proves the result when A.1 holds and $k = 1$.

The proof when A.2 holds with $k = 1$ consists of a minor variation of the same argument. □

The cases when A.3 holds or A.1 or A.2 holds with $k > 1$ and when $\alpha > -1$ are more complicated since then large values of $\{X_t\}$ are caused not by one but by k large Z_λ-values. As an example we will, omitting proofs, briefly discuss what happens when A.3 holds and $\gamma > -1$, in the particular case of a symmetric underlying distribution, i.e., when $P(Z > z) = P(Z < -z)$ for $z \ge 0$. Let $U_1, \dots, U_{k-1}$ be random variables in $[0,1]$ with joint density function

$$f(u_1, \dots, u_{k-1}) = \frac{\Gamma(k(\alpha+1))}{\Gamma(\alpha+1)^k} u_1^\alpha \dots u_{k-1}^\alpha,$$

for $0 \le u_i \le 1$; $i = 1, \dots, k-1$ and $\sum_1^{k-1} u_i \le 1$, and define $\bar{Z}_{\lambda_1} = U_1, \dots \bar{Z}_{\lambda_{k_+}} = U_{k_+}$, $\bar{Z}_{\lambda_1^-} = U_{k_++1}, \dots \bar{Z}_{\lambda_{k_--1}^-} = U_{k-1}$, $\bar{Z}_{\lambda_k^-} = 1 - \sum_1^{k-1} U_i$, and let $\bar{Z}_\lambda = 0$ for $\lambda \notin \Lambda = \Lambda_+ \cup \Lambda_-$. Now define a stochastic process $Y = \{Y_\tau\}_{\tau=-\infty}^{\infty}$ by

$$Y_\tau = \sum c_{\lambda-\tau} \bar{Z}_\lambda, \qquad \tau = 0, \pm 1, \dots,$$

and let $Y^{(1)}, Y^{(2)}, \dots$ be independent copies of Y, which are also independent of the Poisson process N with intensity measure $dt \times e^{-x}\,dx$. Let the point process N' in $S \times \mathbb{R}^\infty$ be defined by "adjoining independent marks Y to each point of N'', i.e., if N has the points $\{(t_i, x_i), i = 1, 2, \dots\}$, then let N' have the points $\{((t_i, x_i), Y^{(i)}); i = 1, 2, \dots\}$. Then with N'_n and N''_n as defined in Section 2, $N'_n \to_d N'$ and $N''_n \to_d N'$ as $n \to \infty$ in $S \times \mathbb{R}^\infty$.

Finally, as mentioned above, the sample path behavior for $\alpha < -1$ is similar to that for $p < 1$, but with some interesting extra complications.

8. Extremes for $0 < p < 1$. For $0 < p < 1$, as for $p = 1$, $\alpha < -1$, extreme values of weighted sums are caused by just one of the summands being large. However, in this case the scale of extremes increases instead of being constant as for $p = 1$ or decreasing as for $p > 1$, which allows for some further simplification. In the proofs we will use a direct approach, similar to the methods of [7].

Thus in the present case it is fairly straightforward to find the tail behavior of the distribution of $\Sigma c_\lambda Z_\lambda$ by estimating convolution integrals and then the limiting distribution of the maximum $\hat{M}_n$ of the associated independent sequence (Theorem 8.3). For $0 < p < 1$ the limit of the point process N_n of heights and locations of extreme values of $\{X_t = \Sigma c_{\lambda - t} Z_\lambda\}$ is not a simple Poisson process but, if A.1 or A.2 holds, obtained from a Poisson process by replacing each point by k_+ points at the same location (Theorem 8.5). If instead A.3 holds, then each point is replaced randomly by either k_+ or k_- points. This is just as expected: e.g., in cases A.1 or A.2, if an extreme value of $\{X_t\}$ is caused by just one big Z_λ, say $Z_{\bar{\lambda}}$, then X_t should be large at the k_+ time instants $\bar{\lambda} - \lambda_1, \dots, \bar{\lambda} - \lambda_{k_+}$ when the factor before $Z_{\bar{\lambda}}$ in $\Sigma c_{\lambda - t} Z_\lambda$ equals c_+. This behavior is further described in the limit results for the marked point processes N_n' and N_n'' (Theorem 8.6).

We start by proving a counterpart of Lemma 7.1, estimating convolutions of two random variables.

Lemma 8.1. (i) *Suppose the random variables Y_1 and Y_2 are independent and satisfy* (2.2) *with the same α and p* $(0 < p < 1)$, *but with K replaced by K_1 and K_2 for Z replaced by Y_1 and Y_2, respectively. Then*

$$P(Y_1 + Y_2 > z) \sim (K_1 + K_2) z^\alpha e^{-z^p} \quad \text{as } z \to \infty.$$

(ii) *Suppose that Y_1 satisfies* (2.2) *with $p \in (0,1)$ for Z replaced by Y_1 and is independent of Y_2 which satisfies $P(Y_2 > z) = o(z^\alpha e^{-z^p})$ as $z \to \infty$. Then*

$$P(Y_1 + Y_2 > z) \sim K z^\alpha e^{-z^p} \quad \text{as } z \to \infty.$$

(ii) *Suppose $\{Z_\lambda\}_{-\infty}^{\infty}$ are independent random variables such that for some C, $z_0 > 0$ and $p \in (0,1)$,*

$$P(Z_\lambda > z) \le C z^\alpha e^{-z^p} \quad \text{for } z > z_0, \qquad \lambda = 0, \pm 1, \dots,$$

and that $\{c_\lambda\}_{-\infty}^{\infty}$ are constants with $0 < c_\lambda < 1$ and $|\log c_\lambda| > 8$ for all λ and $\Sigma c_\lambda |\log c_\lambda|^{1/p} < 1$. Then

$$P\Big(\sum c_\lambda Z_\lambda > z\Big) = O\big(e^{-2z^p}\big) \quad \text{as } z \to \infty.$$

Proof. (i) We will use (7.5). By (2.2)

$$(8.1) \quad \int_{-\infty}^{z/2} P(Y_1 > z - x) F_2(dx) \sim K_1 z^\alpha e^{-z^p} \int_{-\infty}^{z/2} (1 - x/z)^\alpha e^{z^p - (z-x)^p} F_2(dx)$$

as $z \to \infty$. Here the last integrand tends pointwise to one and is bounded for $-\infty < x \le z^{1-p}$ and $z \ge 1$, since $z^p - (z - x)^p \le \text{constant} \times x/z^{1-p}$ for $0 \le x \le z/2$, and hence

$$(8.2) \qquad \int_{-\infty}^{z^{1-p}} (1 - x/z)^\alpha e^{z^p - (z-x)^p} F_2(dx) \to \int F_2(dx) = 1 \quad \text{as } z \to \infty.$$

As before, let C be a generic constant, whose value may change from one appearance to the next. It then follows from partial integration and (2.2) that

$$
\begin{aligned}
&\int_{z^{1-p}}^{z/2}(1-x/z)^{\alpha}e^{z^p-(z-x)^p}F_2(dx)\\
&\quad\le C\int_{z^{1-p}}^{z/2}e^{z^p-(z-x)^p}F_2(dx)\\
&\quad\le C\Big\{e^{z^p-(z-z^{1-p})^p}\big(1-F_2(z^{1-p})\big)\\
&\qquad+\int_{z^{1-p}}^{z/2}\big(p/(z-x)^{1-p}\big)e^{z^p-(z-x)^p}\big(1-F_2(x)\big)\,dx\Big\}\\
&\quad\le C\Big\{z^{\alpha(1-p)}e^{z^p-(z-z^{1-p})^p-z^{p-p^2}}+z^{\alpha}\int_{z^{1-p}}^{z/2}e^{z^p-(z-x)^p-x^p}\,dx\Big\}.
\end{aligned}
\tag{8.3}
$$

As a function of x, $z^p-(z-x)^p-x^p$ is decreasing for $0<x<z/2$, so replacing the last integrand by its maximum value and using that $z^p-(z-z^{1-p})^p-z^{p(1-p)}=-z^{p-p^2}(1+o(1))$, it follows from (8.3) that

$$
\int_{z^{1-p}}^{z/2}(1-x/z)^{\alpha}e^{z^p-(z-x)^p}F_2(dx)\le C\big(z^{\alpha(1-p)}+z^{\alpha+1}\big)e^{-z^{p-p^2}(1+o(1))}
\tag{8.4}
$$

$$\to 0 \quad \text{as } z\to\infty.$$

Hence, from (8.1), (8.2), and (8.4)

$$\int_{-\infty}^{z/2}P(Y_1>z-x)F_2(dx)\sim K_1z^{\alpha}e^{-z^p}\quad \text{as } z\to\infty,$$

and similarly

$$\int_{-\infty}^{z/2}P(Y_2>z-x)F_1(dx)\sim K_2z^{\alpha}e^{-z^p}\quad \text{as } z\to\infty.$$

Since furthermore $P(Y_1>z/2)P(Y_2>z/2)=O(z^{2\alpha}\exp\{-z^p(2^{-p}+2^{-p})\})=o(z^{\alpha}\exp\{-z^p\})$, part (i) now follows by insertion into (7.5). (ii) follows by similar arguments as in part (i).

(iii) By the assumption $\Sigma c_\lambda|\log c_\lambda|^{1/p}<1$ and Boole's inequality

$$
\begin{aligned}
P\Big(\sum c_\lambda Z_\lambda>z\Big)&\le P\Big(\sum c_\lambda Z_\lambda>\sum c_\lambda|\log c_\lambda|^{1/p}z\Big)\\
&\le\sum P\big(Z_\lambda>|\log c_\lambda|^{1/p}z\big).
\end{aligned}
\tag{8.5}
$$

Here, for $z>z_0$ also $|\log c_\lambda|^{1/p}z>z_0$ so that

$$P\big(Z_\lambda>|\log c_\lambda|^{1/p}z\big)\le C\big(|\log c_\lambda|^{1/p}z\big)^{\alpha}e^{-(|\log c_\lambda|^{1/p}z)^p}$$

and hence, since $x^{\alpha}\exp\{-x^p\}\le\text{constant}\times\exp\{-x^p/2\}$ for $x>z_0$, and using that $|\log c_\lambda|z^p/2=|\log c_\lambda|z^p/4+|\log c_\lambda|z^p/4\ge|\log c_\lambda|+2z^p$ for $z^p>4$ and $|\log c_\lambda|>8$, we have that, for some $C_1>0$ and such z,

$$
\begin{aligned}
P\big(Z_\lambda>|\log c_\lambda|^{1/p}z\big)&\le C_1e^{-|\log c_\lambda|z^p/2}\\
&\le C_1e^{-2z^p}e^{-|\log c_\lambda|}\\
&=C_1e^{-2z^p}c_\lambda.
\end{aligned}
$$

Now, insert this into (8.5) to show that for $z > \max(z_0, 4^{1/p})$,

$$P\left(\sum c_\lambda Z_\lambda > z\right) \le \sum C_1 e^{-2z^p c_\lambda}$$
$$= O(e^{-2z^p}) \quad \text{as } z \to \infty. \quad \square$$

It is now easy to find the asymptotic form of the tail of the distribution of $\Sigma c_\lambda Z_\lambda$.

LEMMA 8.2. *Suppose that one of* A.1–A.3 *holds with* $0 < p < 1$. *Then*

$$P\left(\sum c_\lambda Z_\lambda > z\right) \sim \hat{K}(z/c_+)^\alpha e^{-(z/c_+)^p} \quad \text{as } z \to \infty,$$

where $\hat{K} = k_+ K$ *if* A.1 *or* A.2 *holds and* $\hat{K} = k_+ K + k_- K_- \gamma^{\alpha/p}$ *if* A.3 *holds.*

PROOF. Assume that A.1 holds and, as usual without loss of generality, that $c_+ = 1$. From (2.2) and Lemma 8.1(i) used $k_+ - 1$ times it follows that

$$P\left(Z_{\lambda_1} + \cdots + Z_{\lambda_{k_+}} > z\right) \sim k_+ K z^\alpha e^{-z^p} \quad \text{as } z \to \infty.$$

Similarly it then follows from repeated use of Lemma 8.1(ii) that if $\bar{\lambda}$ is large enough to make $|\lambda_i| \le \bar{\lambda}$ for $i = 1, \dots, k_+$ then

$$P\left(\sum_{|\lambda| \le \bar{\lambda}} c_\lambda Z_\lambda > z\right) = P\left(Z_{\lambda_1} + \cdots + Z_{\lambda_{k_+}} + \sum_{\substack{\lambda \notin \Lambda_+ \\ |\lambda| \le \bar{\lambda}}} c_\lambda Z_\lambda > z\right) \tag{8.6}$$
$$\sim k_+ K z^\alpha e^{-z^p} \quad \text{as } z \to \infty.$$

Now, let $\bar{\lambda}$ be large enough to make $|\log c_\lambda| > 8$ for $|\lambda| > \bar{\lambda}$ and $\Sigma_{|\lambda| > \bar{\lambda}} c_\lambda |\log c_\lambda|^{1/p} < 1$, which is possible since (2.8) is assumed to hold. It then follows from Lemma 8.1(iii) and (2.2) that

$$P\left(\sum_{|\lambda| > \bar{\lambda}} c_\lambda Z_\lambda > z\right) = o(ze^{-z^p}) \quad \text{as } z \to \infty,$$

and this together with (8.6) is by Lemma 8.1(ii) sufficient to establish that

$$P\left(\sum c_\lambda Z_\lambda > z\right) = P\left(\sum_{|\lambda| \le \bar{\lambda}} c_\lambda Z_\lambda + \sum_{|\lambda| > \bar{\lambda}} c_\lambda Z_\lambda > z\right)$$
$$\sim k_+ K z^\alpha e^{-z^p} \quad \text{as } z \to \infty.$$

The result follows similarly under hypothesis A.2 and also under A.3 after writing $c_\lambda Z_\lambda$ as $c_\lambda^- \gamma^{1/p}(-\gamma^{-1/p} Z_\lambda)$ for negative c_λ's in $\Sigma c_\lambda Z_\lambda$ again noting that $P(-\gamma^{-1/p} Z_\lambda > z) \sim K_- \gamma^{\alpha/p} z^\alpha e^{-z^p}$ as $z \to \infty$ by A.3. $\square$

Hence, the appropriate norming constants for the maximum of the associated independent sequence are

$$\hat{a}_n = c_+^{-1} p (\log n)^{1-1/p}, \tag{8.7}$$
$$\hat{b}_n = c_+ (\log n)^{1/p} + (c_+/p)((\alpha/p)\log\log n + \log \hat{K})/(\log n)^{1-1/p},$$

with

$$\hat{K} = \begin{cases} k_+ K & \text{if A.1 or A.2 holds,} \\ k_+ K + k_- K_- \gamma^{\alpha/p} & \text{if A.3 holds.} \end{cases}$$

THEOREM 8.3. *Suppose that one of* A.1–A.3 *is satisfied, with* $0 < p < 1$ *and let* $\hat{a}_n$, $\hat{b}_n$ *be given by* (8.7). *Then*

$$P\big(\hat{a}_n(\hat{M}_n - \hat{b}_n) \le x\big) \to e^{-e^{-x}} \quad \text{as } n \to \infty.$$

However, for $0 < p < 1$ the norming constants a_n, b_n for the moving average process $\{X_t = \sum c_{\lambda - t} Z_\lambda\}$ are the same as for the noise variables (provided $c_+ = 1$ and if A.1 or A.2 holds) and not as for $p > 1$, those of the associated independent sequence. Thus let

$$a_n = c_+^{-1} p (\log n)^{1-1/p} \tag{8.8}$$

and

$$b_n = \begin{cases} c_+(\log n)^{1/p} + (c_+/p)\big((\alpha/p)\log\log n + \log K\big)/(\log n)^{1-1/p} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if A.1 or A.2 holds,} \\ c_+(\log n)^{1/p} \\ \quad + (c_+/p)\big((\alpha/p)\log\log n + \log\big(K + K_-\gamma^{\alpha/p}\big)\big)/(\log n)^{1-1/p} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if A.3 holds.} \end{cases} \tag{8.9}$$

[However, it may be noted that the difference between the various norming constants is not large, e.g., if A.1 or A.2 holds and $c_+ = 1$ then

$$\begin{aligned} a_n(M_n - b_n) &= \tilde{a}_n(M_n - \tilde{b}_n) \\ &= \hat{a}_n(M_n - \hat{b}_n) + \log \hat{K}/K \\ &= \hat{a}_n(M_n - \hat{b}_n) + \log k_+ .] \end{aligned}$$

The next lemma is the first step in making precise the notion that large values of $X_t = \sum c_{\lambda - t} Z_\lambda$ are caused by just one large Z_λ.

LEMMA 8.4. *Let* a_n *and* b_n *be given by* (8.8) *and the first part of* (8.9), *with* $c_+ = 1$ [*or equivalently, let* $a_n = \tilde{a}_n$, $b_n = \tilde{b}_n$ *with* $\tilde{a}_n$, $\tilde{b}_n$ *given by* (4.2) *with* $K > 0$ *a fixed arbitrary constant*]. *Let* $\varepsilon > 0$ *and* x *be fixed and write* $\varepsilon_n = \varepsilon / a_n$ *and* $u_n = x/a_n + b_n$.

(i) *Suppose* Y_1 *and* Y_2 *are as in Lemma* 8.1(i). *Then*

$$nP(Y_1 \le u_n - \varepsilon_n, Y_2 \le u_n - \varepsilon_n, Y_1 + Y_2 > u_n) \to 0 \quad \text{as } n \to \infty. \tag{8.10}$$

(ii) *If* Y_1 *is as in part* (i) *and is independent of* Y_2 *with* $P(Y_2 > z) = o(z^\alpha e^{-z^p})$ *as* $z \to \infty$, *then*

$$nP(Y_1 \le u_n - \varepsilon_n, Y_1 + Y_2 > u_n) \to 0 \quad \text{as } n \to \infty.$$

(iii) *Let* $Y_1, \dots, Y_k$ *be independent and satisfy* (2.2) *with the same* α *and* $p \in (0,1)$ *but possibly with different* K*'s for* Z *replaced by* Y_i, $i = 1, \dots, k$. *Then*

$$nP\left(Y_1 \le u_n - \varepsilon_n, \dots, Y_k \le u_n - \varepsilon_n, \sum_{i=1}^{k} Y_i > u_n\right) \to 0 \quad \text{as } n \to \infty.$$

PROOF. (i) Similarly as for (7.5) we have that

$$\begin{aligned} P(Y_1 &\le u_n - \varepsilon_n, Y_2 \le u_n - \varepsilon_n, Y_1 + Y_2 > u_n) \\ &\le \int_{\varepsilon_n}^{u_n/2} P(Y_1 > u_n - x) F_2(dx) \\ &\quad + \int_{\varepsilon_n}^{u_n/2} P(Y_2 > u_n - x) F_1(dx) + P(Y_1 > u_n/2) P(Y_2 > u_n/2). \end{aligned} \tag{8.11}$$

By the choice of a_n, b_n, it holds that $u_n^{\alpha} \exp\{-u_n^p\} = O(1/n)$. Hence, using in turn (2.2), this and $\varepsilon_n \to \infty$ as $n \to \infty$,and estimating

$$\int_{u_n^{1-p}}^{u_n/2} (1 - x/u_n)^{\alpha} \exp\{u_n^p - (u_n - x)^p\} F_2(dx)$$

as in Lemma 8.1(i), it follows that

$$\begin{aligned} \int_{\varepsilon_n}^{u_n/2} P(Y_1 > u_n - x) F_2(dx) &\sim K_1 u_n^{\alpha} e^{-u_n^p} \int_{\varepsilon_n}^{u_n/2} (1 - x/u_n)^{\alpha} e^{u_n^p - (u_n - x)^p} F_2(dx) \\ &= O(1/n) \int_{\varepsilon_n}^{u_n^{1-p}} F_2(dx) + o(1/n) \\ &= o(1/n) \quad \text{as } n \to \infty. \end{aligned}$$

Similarly, the second integral in (8.11) is $o(n^{-1})$, and since $P(Y_1 > u_N/2)P(Y_2 > u_n/2) = o(n^{-1})$ as in Lemma 8.1(i), it follows from (8.11) that (8.10) holds.

(ii) This follows similarly [cf. Lemma 8.1(ii)] after replacing (8.10) by

$$\begin{aligned} P(Y_1 &\le u_n - \varepsilon_n, Y_1 + Y_2 > u_n) \\ &\le \int_{\varepsilon_n}^{u_n/2} P(Y_1 > u_n - x) F_2(dx) \\ &\quad + \int_{-\infty}^{u_n/2} P(Y_2 > u_n - x) F_1(dx) + P(Y_1 > u_n/2) P(Y_2 > u_n/2). \end{aligned}$$

(iii) It is readily seen that

$$\begin{aligned} &\left\{ \sum_{i=1}^{k} Y_i > u_n, Y_1 \le u_n - \varepsilon_n, \dots, Y_k \le u_n - \varepsilon_n \right\} \\ &\quad \subset \left\{ \sum_{i=1}^{k-1} Y_i > u_n - \varepsilon_n/k, Y_1 \le u_n - \varepsilon_n, \dots, Y_{k-1} \le u_n - \varepsilon_n \right\} \\ &\qquad \cup \left\{ \sum_{i=1}^{k} Y_i > u_n, \sum_{i=1}^{k-1} Y_1 \le u_n - \varepsilon_n/k, Y_k \le u_n - \varepsilon_n \right\}, \end{aligned}$$

and repeating the procedure shows that

$$\left\{\sum_{i=1}^{k} Y_i > u_n, Y_1 \le u_n - \varepsilon_n, \dots, Y_k \le u_n - \varepsilon_n\right\}$$
$$\subset \bigcup_{l=2}^{k} \left\{\sum_{i=1}^{l} Y_i > u_n - ((k-l)/k)\varepsilon_n, \right.$$
$$\left. \sum_{i=1}^{l-1} Y_i \le u_n - ((k+1-l)/k)\varepsilon_n, Y_l \le u_n - \varepsilon_n\right\}.$$

Hence

$$P\left(\sum_{i=1}^{k} Y_i > u_n, Y_1 \le u_n - \varepsilon_n, \dots, Y_k \le u_n - \varepsilon_n\right)$$
$$\le \sum_{l=2}^{k} P\left(\sum_{i=1}^{l} Y_i > u_n - ((k-l)/k)\varepsilon_n, \right.$$
$$\left. \sum_{i=1}^{l-1} Y_i \le u_n - ((k+1-l)/k)\varepsilon_n, Y_l \le u_n - ((k+1-l)/k)\varepsilon_n\right)$$

and the result follows from applying part (i) to each term in the sum with the obvious identifications, since $\Sigma_1^{l-1} Y_i$ satisfies the requirements put on Y_1 in part (i) by Lemma 8.2. □

As discussed above, it will presently be shown that if A.1 or A.2 holds, then each large Z_λ-value, say $Z_{\bar{\lambda}}$, leads to precisely k_+ large X_t values at fixed distances from $\bar{\lambda}$ and with heights approximately equal to $c_+ Z_{\bar{\lambda}}$. Similarly if A.3 holds, a large (positive) Z_λ causes k_+ large (positive) X_t-values, and a large negative Z_λ causes k_- large (positive) X_t-values. Thus, taking into account the effect of time and height scaling in N_n, its limit is of the following form. Let $\tilde{N}$, $\tilde{N}_+$, and $\tilde{N}_-$ be Poisson processes in $[0, \infty) \times \mathbb{R}$ with intensities $dt \times e^{-x}\,dx$, $dt \times K(K + K_-\gamma^{\alpha/p})^{-1}e^{-x}\,dx$, and $dt \times K_-\gamma^{\alpha/p}(K + K_-\gamma^{\alpha/p})^{-1}e^{-x}\,dx$, respectively, with $\tilde{N}_+$ and $\tilde{N}_-$ mutually independent, and define the point process N by

$$N(B) = \begin{cases} k_+\tilde{N}(B) & \text{if A.1 or A.2 holds,} \\ k_+\tilde{N}_+(B) + k_-\tilde{N}_-(B) & \text{if A.3 holds.} \end{cases} \tag{8.12}$$

For the proof that $N_n \to_d N$ we will directly use the structure of extremes discussed above. The basic idea of the proof is quite simple and the calculations are elementary, but involve some long expressions.

THEOREM 8.5. *Suppose that one of* A.1–A.3 *is satisfied, and let* N_n *be as defined in Section* 2, *with* a_n, b_n *given by* (8.8), (8.9). *Then* $N_n \to_d N$ *as* $n \to \infty$ *in* $[0, \infty) \times \mathbb{R}$ *with* N *given by* (8.12). *In particular,*

$$P(a_n(M_n - b_n) \le x) \to e^{-e^{-x}} \quad \text{as } n \to \infty.$$

PROOF. Assume A.1 holds, and as usual without loss of generality, that $c_+ = 1$. Let $I = [s, t) \times (x, \infty)$ be a fixed rectangle in $[0, \infty) \times \mathbb{R}$, write $u_n = x/a_n + b_n$, and define

$$\bar{X}_t = \sum_{\lambda \in \Lambda_+} c_\lambda Z_{\lambda+t} = \sum_{\lambda \in \Lambda_+} Z_{\lambda+t},$$

$$\bar{\bar{X}}_t = \sum_{\lambda \in \Lambda_+} Z_{\lambda+t} 1\{Z_{\lambda+t} > u_n\},$$

and let $\bar{N}_n$ and $\bar{\bar{N}}_n$ be defined from $\{\bar{X}_t\}$ and $\{\bar{\bar{X}}_t\}$ in the same way as N_n is defined from $\{X_t\}$, and let $\tilde{N}_n$ be similarly defined from $\{Z_t\}$. We will prove that

$$P\big(\bar{\bar{N}}_n(I) \neq k_+ \tilde{N}_n(I)\big) \to 0 \quad \text{as } n \to \infty, \tag{8.13}$$

$$P\big(\bar{N}_n(I) \neq \bar{\bar{N}}_n(I)\big) \to 0 \quad \text{as } n \to \infty, \tag{8.14}$$

and that

$$P(N_n(I) \neq \bar{N}_n(I)) \to 0 \quad \text{as } n \to \infty. \tag{8.15}$$

As noted in Section 4, $\tilde{N}_n \to_d \tilde{N}$ as $n \to \infty$, and hence obviously $k_+ \tilde{N}_n \to_d k_+ \tilde{N} = N$ and $N_n \to_d N$ then follows from (8.13)–(8.15) by applying Lemma 3.3 three times.

It is readily seen, that for

$$A = \max\{|\lambda|; \lambda \in \Lambda_+\},$$

it holds that

$$\begin{aligned} &\big\{\bar{\bar{N}}_n(I) \neq k_+ \tilde{N}_n(I)\big\} \\ &\quad \subset \{Z_\lambda > u_n \text{ for some } \lambda \in [ns - A, ns + A] \cup [nt - A, nt + A]\} \\ &\quad\quad \cup \big\{Z_\lambda > u_n, Z_{\lambda+\mu} > u_n \text{ for some } \lambda \in [ns, nt) \\ &\quad\quad\quad \text{and } \mu \text{ with } \mu \neq 0 \text{ and } |\mu| \le A\big\}. \end{aligned} \tag{8.16}$$

Here, by Boole's inequality and stationarity

$$\begin{aligned} &P(Z_\lambda > u_n \text{ for some } \lambda \in [ns - A, ns + A] \cup [nt - A, nt + A]) \\ &\quad \le 2(2A + 1) P(Z > u_n) \\ &\quad \to 0 \quad \text{as } n \to \infty, \end{aligned} \tag{8.17}$$

and similarly

$$\begin{aligned} &P\big(Z_\lambda > u_n, Z_{\lambda+\mu} > u_n \text{ for some } \lambda \in [ns, nt) \\ &\quad \text{and } \mu \text{ with } \mu \neq 0 \text{ and } |\mu| \le A\big) \\ &\quad \le n(t - s) 2A P(Z > u_n)^2 \\ &\quad \to 0 \quad \text{as } n \to \infty, \end{aligned} \tag{8.18}$$

since $P(Z > u_n) \sim K u_n^\alpha \exp\{-u_n^p\} = O(1/n)$ by the choice of a_n, b_n. Now (8.13) is an immediate consequence of (8.16)–(8.18).

Next, fix $\varepsilon > 0$, define $I_\varepsilon = [s-\varepsilon, t+\varepsilon) \times [x-\varepsilon, x+\varepsilon]$, and write $\varepsilon_n = \varepsilon/a_n$. It can be seen that for large n

$$\begin{aligned}&\{\overline{N}_n(I) < \overline{\overline{N}}_n(I)\}\\ &\subset \{\tilde{N}_n(I_\varepsilon) > 0\} \cup \{Z_\lambda > u_n, Z_{\lambda+\mu} \le -\varepsilon_n/k_+,\\ &\qquad \text{for some } \lambda \in [ns-A, nt+A) \text{ and } \mu \ne 0 \text{ with } |\mu| \le A\},\end{aligned} \tag{8.19}$$

and that

$$\begin{aligned}\{\overline{N}_n(I) \ge \overline{\overline{N}}_n(I)\} \subset \{\tilde{N}_n(I_\varepsilon) > 0\} \cup \Bigg\{ &\sum_{\mu\in\Lambda_+} Z_{\mu+\lambda} > u_n,\\ &Z_{\lambda_1+\lambda} \le u_n - \varepsilon_n, \dots, Z_{\lambda_{k_+}+\lambda} \le u_n - \varepsilon_n \text{ for some } \lambda \in [ns, nt)\Bigg\}.\end{aligned} \tag{8.20}$$

Since the Z_λ's are independent, it follows from Boole's inequality and stationarity that

$$\begin{aligned}&P\big(Z_\lambda > u_n, Z_{\lambda+\mu} \le -\varepsilon_n/k_+ \text{ for some } \lambda \in [ns-A, nt+A)\\ &\quad \text{and } \mu \ne 0 \text{ with } |\mu| \le A\big)\\ &\quad \le (n(t-s) + 2A)P(Z > u_n)P(Z < -\varepsilon_n/k_+)\\ &\quad \to 0 \quad \text{as } n \to \infty,\end{aligned} \tag{8.21}$$

since $P(Z > u_n) = O(1/n)$ as noted above, and since $\varepsilon_n \to \infty$, and hence $P(Z < -\varepsilon_n/k_+) \to 0$ as $n \to \infty$. Moreover, a similar argument, together with Lemma 8.4(iii), shows that

$$\begin{aligned}&P\Bigg(\sum_{\mu\in\Lambda_+} Z_{\mu+\lambda} > u_n, Z_{\lambda_1+\lambda} \le u_n - \varepsilon_n, \dots, Z_{\lambda_{k_+}+\lambda} \le u_n - \varepsilon_n\\ &\qquad \text{for some } \lambda \in [ns, nt)\Bigg)\\ &\quad \le n(t-s)P\Bigg(\sum_{\mu\in\Lambda_+} Z_\mu > u_n, Z_{\lambda_1} \le u_n - \varepsilon_n, \dots, Z_{\lambda_{k_+}} \le u_n - \varepsilon_n\Bigg)\\ &\quad \to 0 \quad \text{as } n \to \infty.\end{aligned} \tag{8.22}$$

Since $\tilde{N}_n \to_d \tilde{N}$ it follows from (8.19)–(8.22) that

$$\limsup_{n\to\infty} P\big(\overline{N}_n(I) \ne \overline{\overline{N}}_n(I)\big) \le \limsup_{n\to\infty} P\big(\tilde{N}_n(I_\varepsilon) > 0\big) = P\big(\tilde{N}(I_\varepsilon) > 0\big),$$

and since the latter quantity tends to zero as $\varepsilon \to 0$, this proves (8.14).

Finally, (8.15) is proved in a similar manner. In fact, with the same notation,

$$\begin{aligned}\{N_n(I) \ne \overline{N}_n(I)\} &\subset \{\overline{N}_n(I_\varepsilon) > 0\}\\ &\quad \cup \{X_\lambda > u_n, \overline{X}_\lambda \le u_n - \varepsilon_n \text{ for some } \lambda \in [ns, nt)\}\\ &\quad \cup \{X_\lambda \le u_n, \overline{X}_\lambda > u_n + \varepsilon_n \text{ for some } \lambda \in [ns, nt)\}.\end{aligned} \tag{8.23}$$

Lemma 3.3 together with the already proved relations show that $\overline{N}_n \to_d k_+\tilde{N} =$

N, and thus in particular that

$$P(\overline{N}_n(I_\varepsilon) > 0) \to P(\tilde{N}(I_\varepsilon) > 0),$$

which as before tends to zero as $\varepsilon \to 0$. Since $X_t = \overline{X}_t + \Sigma_{\lambda \notin \Lambda_+} c_\lambda Z_\lambda$, where the two terms are independent and satisfy the hypothesis of Lemma 8.4(ii), according to Lemma 8.2, it follows as in (8.22) that the probability of the next to last event in (8.23) tends to zero. Further,

$$\{X_0 \le u_n, \overline{X}_0 > u_n + \varepsilon_n\} \subset \{\Sigma_{\lambda \notin \Lambda_+} c_\lambda Z_\lambda < -\varepsilon_n, \overline{X}_0 > u_n + \varepsilon_n\},$$

and since $\Sigma_{\lambda \notin \Lambda_+} c_\lambda Z_\lambda$ and $\overline{X}_0$ are independent, this shows, as in (8.21), that also the probability of the last set in (8.23) tends to zero. Now, (8.15) follows in the same way as (8.14), which completes the proof of the theorem for the case when A.1 holds. The proofs for hypotheses A.2 and A.3 follow similar lines. □

Next, define points $y^{(i)} = \{y^{(i)}(\tau)\}_{\tau=-\infty}^{\infty}$ and $y_-^{(i)} = \{y_-^{(i)}(\tau)\}_{\tau=-\infty}^{\infty}$ in $\mathbb{R}^\infty$ by

$$y^{(i)}(\tau) = c_{\lambda_i - \tau}/c_+ \quad \text{for } \tau = 0, \pm 1, \dots \text{ and } i = 1, \dots, k_+,$$

$$y_-^{(i)}(\tau) = -c_{\lambda_i^- - \tau}/c_- \quad \text{for } \tau = 0, \pm 1, \dots \text{ and } i = 1, \dots, k_-.$$

Further, let $\tilde{N}$, $\tilde{N}_+$ and $\tilde{N}_-$ be as defined just before Theorem 8.5. The limit N' of the marked point processes N_n' and N_n'' is then, if A.1 or A.2 holds, defined by requiring that to each point (t, x) of $\tilde{N}$ there correspond k_+ points

$$((t, x), y^{(1)}), \dots, ((t, x), y^{(k_+)})$$

of N'. If instead A.3 holds, then N' is defined from the independent Poisson processes $\tilde{N}_+$ and $\tilde{N}_-$ by requiring that to each point (t_+, x_+) of $\tilde{N}_+$ there correspond k_+ points

$$((t_+, x_+), y^{(1)}), \dots, ((t_+, x_+), y^{(k_+)})$$

of N' and to each point (t_-, x_-) of $\tilde{N}_-$ there correspond the k_- points

$$((t_-, x_-), y_-^{(1)}), \dots, ((t_-, x_-), y_-^{(k_-)})$$

of N'. The convergence of N_n', N_n'' can now be obtained by direct approximation by similar arguments as for Theorem 8.5. Since no new ideas are involved in this, we omit the proof.

THEOREM 8.6. *Suppose that one of* A.1–A.3 *is satisfied, let N' be as defined above, and let N_n', N_n'' be as defined in Section* 2, *with a_n, b_n given by* (8.8), (8.9). *Then $N_n' \to_d N'$ and $N_n'' \to_d N'$ as $n \to \infty$ in $S \times \mathbb{R}^\infty$.*

9. Remarks on polynomial tails, autoregressive processes and the conditions. This section contains some comments on (i) noise variables with polynomially decreasing tails, (ii) how the results apply to autoregressive (AR) and autoregressive-moving average (ARMA) processes, (iii) the conditions on the weights $\{c_\lambda\}$, and (iv) the conditions on the distribution of the noise variables $\{Z_\lambda\}$.

(i) *Polynomial tails.* Formally, this is the case when $p = 0$ in (2.2), i.e., when

$$P(Z > z) \sim Kz^{\alpha} \quad \text{as } z \to \infty \tag{9.1}$$

for some $\alpha \in (-\infty, 0)$. Special classes of moving averages $X_t = \Sigma c_{\lambda - t} Z_\lambda$ which satisfy (9.1) are studied in [7] and [3]. As for $0 < p < 1$, an extreme value of the moving average process for $p = 0$ is caused by just one large noise variable Z_λ. In particular, if Z satisfies (9.1), and if the same relation holds, but with K replaced by K_-, if Z is replaced by $-Z$, this leads to a type II limit for the maximum,

$$P(a_n(M_n - b_n) \le x) \to e^{-x^{-|\alpha|}} \tag{9.2}$$

for $x \ge 0$ if a_n, b_n e.g., are chosen as

$$a_n = \left(Kc_+^{|\alpha|} + K_- c_-^{|\alpha|}\right)^{-1/|\alpha|} n^{-1/|\alpha|},$$
$$b_n = 0.$$

Thus extremes increase much faster for $p = 0$ than for $p > 0$, and in addition scale and location are of the same order, so that it is possible to choose $b_n = 0$. In contrast to $0 < p < 1$ this also introduces a random amplitude into the behavior of sample paths near extremes. Specifically, for the case when the Z_λ's have a (nonnormal) stable distribution—which then satisfies (9.1) with $|\alpha| \in (0, 2)$—this is discussed at length in [7], in a somewhat different point process formulation. Rather loosely described, it is shown there that e.g., for positive c_λ's the normalized sample path $a_n X_\tau$ near an extreme value at, say, zero has the same distribution as a random translate of the function

$$y'_\tau = U' c_{-\tau}/c_+, \qquad \tau = 0, \pm 1, \dots,$$

where U' is a certain Pareto distributed random variable. Furthermore, sample paths near different separated extreme values are asymptotically independent. It then follows that X_τ / X_0 has a similar form, i.e., it approaches a translate of

$$y''_\tau = U'' c_{-\tau}, \qquad \tau = 0, \pm 1, \dots,$$

where the random variable U'' only assumes the values $\dots \pm 1/c_{-1}$, $\pm 1/c_0, \pm 1/c_1 \dots$. Thus for $p = 0$, the limits of N'_n and N''_n are not the same, but have a similar, deterministic form, except for a random amplitude and time translation.

In [3] the limit (9.2) is obtained for general Z's which satisfy (9.1) [and indeed also for a slightly more general case when the Z's belong to the domain of attraction of the type II extreme value distribution, or equivalently when the right-hand side of (9.1) may include a further slowly varying factor]. The conditions include $c_0 > |c_\lambda|$ for $\lambda \ne 0$. As noted in [7] the methods in that paper work also for such general Z's, the only supplementary fact needed is a bound for the tail of the d.f. of $\Sigma c_\lambda Z_\lambda$, which in turn can be obtained in the same way as for $0 < p < 1$.

(ii) *Autoregressive and autoregressive-moving average processes.* A stationary process $\{X_t\}$ is an infinite ARMA process if it satisfies the difference equation

$$X_t + d_1 X_{t+1} + d_2 X_{t+2} + \cdots = Z_t + e_1 Z_{t+1} + e_2 Z_{t+2} + \cdots \quad \text{for } t = 0, \pm 1, \dots \tag{9.3}$$

for some constants $\{d_\lambda\}_1^\infty$ and $\{e_\lambda\}_1^\infty$. If all the e_λ's are zero, then X_t is an AR-process. Here we only consider the case when the noise variables $\{Z_\lambda\}$ are independent and identically distributed. Rather generally, under weak conditions on $\{d_\lambda\}$, such processes can be "inverted," i.e., written as infinite moving averages. Let z be a complex variable and introduce the generating functions $D(z) = 1 + d_1 z + d_2 z^2 + \cdots$ and $E(z) = 1 + e_1 z + e_2 z^2 + \cdots$. If the coefficients $\{c_\lambda\}$ defined by $E(z)/D(z) = c_0 + c_1 z + c_2 z^2 + \cdots$ make $\sum c_\lambda Z_\lambda$ convergent then inversion to $X_t = \sum_{\lambda=0}^\infty c_\lambda Z_{\lambda+t}$ is possible, and if in addition the c_λ's satisfy (2.8) or (2.9), as required, the results of Sections 5–8 also give the extremal behavior of the ARMA-process (9.3).

From complex function theory it follows that if $D(z)$ and $E(z)$ converge for $|z| \le 1 + \varepsilon$ for some $\varepsilon > 0$ and $D(z)$ has no zeros in $|Z| \le 1 + \varepsilon$, then the c_λ's decrease exponentially and (2.8) and (2.9) are trivially satisfied, but of course these conditions are by no means necessary. In particular if $\{X_t\}$ is a finite ARMA-process (i.e., if only finitely many of $\{d_\lambda, e_\lambda\}$ are nonzero) and if $D(z) \ne 0$ for $|z| \le 1$, as is usually assumed, then (2.8) and (2.9) hold [since $D(z)$ only has finitely many zeros so that $D(z) \ne 0$ for $|z| \le 1 + \varepsilon$ for some $\varepsilon > 0$].

The results of [3] on exponential and polynomial tails are proved for infinite AR-processes subject to $\sum_{\lambda=1}^\infty |d_\lambda| < 1$. Since $\{c_\lambda\}$ then can be obtained recursively from $c_0 = 1$ and $c_n = -(d_1 c_{n-1} + \cdots + d_n c_0)$, it is easy to see that this implies that $|c_\lambda| < 1$ for $\lambda \ne 0$, and that $\sum_1^\infty |c_\lambda| \le \sum_1^\infty |d_\lambda| / (1 - \sum_1^\infty |d_\lambda|)$. Thus $|c_\lambda| < c_0$ for $\lambda \ne 0$ and $\sum_1^\infty |c_\lambda| < \infty$, but the c_λ's do not have to satisfy any condition of the type $|c_\lambda| = O(|\lambda|^{-\theta})$ for any $\theta > 0$.

(iii) *The conditions on the weights* $\{c_\lambda\}$. In a sense the main restriction (2.8) on the c_λ's [which is the same as (2.9) for $1 < p \le 2$] that $|c_\lambda| = O(|\lambda|^{-\theta})$ as $\lambda \to \pm\infty$ for some $\theta > 1$ is quite weak, being close to the requirement that $\sum c_\lambda$ is convergent, which in turn is necessary for convergence of $\sum c_\lambda Z_\lambda$ if $EZ \ne 0$. However, if $EZ = 0$ and $EZ^2 < \infty$, then

$$\sum c_\lambda^2 < \infty \tag{9.4}$$

is sufficient for convergence, and there is more room for weaker conditions. It is known that, at least in the normal case, some further condition beyond (9.4) is needed for the extremal results of this paper to hold, since if the noise variables are normally distributed and e.g., $\lim_{t\to\infty} \log t \sum_\lambda c_{\lambda-t} c_\lambda = \gamma > 0$ then the limit distribution of M_n is different from the one in Corollary 6.5 (see e.g., [5], Section 6.5).

However, Berman (1984) shows that if the Z_λ's are normal and

$$\log n \sum_{|n|<\lambda}^{\infty} c_\lambda^2 \to 0 \quad \text{as } n \to \infty, \tag{9.5}$$

then the conclusion of Corollary 6.5 is still valid, and thus (2.9) can be substantially weakened in this case. In fact, it follows easily from Lemmas 3.1 and 3.2 that if $(\log n)^2 \Sigma_{|n| < \lambda} c_\lambda^2 \to 0$, then the result of Theorem 6.4 holds, and some further work shows that this also is true under the weaker condition (9.5).

(iv) *The conditions on the noise variables.* The condition (2.2) defines the scope of the present investigation. However, of course all the results trivially extend to the case where instead of Z some location-scale transformation $a(Z - b)$ of it satisfies (2.2) (for $0 < p \leq 1$) or (2.3), (2.4), and (2.7) (for $1 < p$). Further the methods probably also work if z^p in (2.2) [or (2.3)] is replaced by some suitable polynomial $d_1 z^{p_1} + \cdots + d_k z^{p_k}$, and for $0 < p \leq 1$ the factor z^α can be replaced by $L(z)$, where L is regularly varying with index α. For $p > 1$ in addition to (2.2) we have imposed the smoothness restrictions (2.3), (2.4), and (2.7). These conditions were introduced in the proofs for technical reasons and certainly should be possible to relax to some extent. Nevertheless, it does not seem likely that the results for $p > 1$ hold in general without any further restrictions beyond (2.2).

Acknowledgments. Much of the work leading to the present paper has been done during two visits to the Department of Statistics and the Center for Stochastic Processes at the University of North Carolina at Chapel Hill. It is a pleasure to thank the department, and in particular Ross Leadbetter and Stamatis Cambanis, for the hospitality shown to me during these visits. Further I want to thank Jane Wille for her swift typing.

REFERENCES

[1] BERMAN, S. M. (1984). Limiting distribution of sums of nonnegative stationary random variables. *Ann. Inst. Statist. Math.* **36** 301–321.

[2] CHERNICK, M. R. (1981). A limit theorem for the maximum of autoregressive processes with uniform marginal distributions. *Ann. Probab.* **9** 145–149.

[3] FINSTER, M. (1982). The maximum term and first passage time for autoregressions. *Ann. Probab.* **10** 737–744.

[4] KALLENBERG, O. (1976). *Random Measures.* Academic, London.

[5] LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes.* Springer, New York.

[6] MATTHES, K., KERSTAN, J. and MECKE, J. (1978). *Infinitely Divisible Point Processes.* Wiley, Chichester.

[7] ROOTZÉN, H. (1978). Extremes of moving averages of stable processes. *Ann. Probab.* **6** 847–869.

[8] ROOTZÉN, H. (1985). A ratio limit theorem for the tails of weighted sums. Report, Inst. of Math. Statist., Univ. of Copenhagen.

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN
5 UNIVERSITETSPARKEN
DK-2100 COPENHAGEN Ø
DENMARK