Bernoulli **9**(3), 2003, 497–515

Compound Poisson limit theorems for high-level exceedances of some non-stationary processes

LISE BELLANGER[1] and GONZALO PERERA[2]

[1]*Université de Nantes, Laboratoire de Mathématiques-UMR 6629, 2, rue de la Houssinière BP 92208, 44322 Nantes Cedex 03, France. E-mail: lise.bellanger@math.univ-nantes.fr*
[2]*Laboratorio de Probabilidad y Estadística IMERL, Facultad de Ingeniería, Julio Herrera y Reissig 565, 11200 Montevideo, Uruguay. E-mail: gperera@fing.edu.uy*

We show the convergence to compound Poisson process of the high-level exceedances point process $N_n(B = \sum_{j/n \in B} 1_{\{X_j > u_n\}}$, where $X_n = \varphi(\xi_n, Y_n)$ φ is a (regular) regression function, u_n grows to infinity with n in some suitable way, ξ and Y are mutually independent, ξ is stationary and weakly dependent, and Y is non-stationary, satisfying some ergodic conditions. The basic technique is the study of high-level exceedances of stationary processes over suitable collections of random sets.

Keywords: asymptotically ponderable collections of sets; compound Poisson process; convergence; exceedances; level sets; mean occupation measures; point processes

1. Introduction

In many meteorological or hydrological problems, relevant features are related to exceedances of high levels by some time series. In particular, current standards for ozone regulation involve exceedances of high levels. It is clear that the time series of ozone level depends on non-stationary, meteorological variables such as temperature and wind speed. Therefore, a reasonable model for the ozone level at time t (say, X_t) should be of the form

$$X_t = \varphi(\xi_t, Y_t), \tag{1}$$

where ξ_t is 'pure noise' corresponding to local fluctuations of measurements systems, Y_t is a vector that contains the values of all the 'explanatory' variables at time t (e.g., temperatures for the previous q days) and φ is some suitable regression function. We may also think of t as a d-dimensional parameter corresponding to space and time; all the models and results in this paper are valid in this context, but, for the sake of simplicity, we will only present the case $d = 1$. One important point to make is that we may assume that ξ is stationary and weakly dependent (say, mixing), but Y may not be stationary: for instance, in the case of temperature, besides seasonal effects that affect the time scale, spatial variations due to differences between urban and rural areas do not make it reasonable to assume stationarity. Furthermore, even if Y may satisfy some ergodic properties (some law of large numbers) it is not reasonable to expect mixing, association or any particular weak-dependence structure.

1350–7265 © 2003 ISI/BS

We will deal with discrete-time observations, so we will observe whether $X_1, \ldots, X_n$ exceed a level u_n which grows to infinity with n in a suitable way. When $X = (X_t\colon t \in \mathbb{N})$ consists of independent and identically distributed (i.i.d.) random variables and u_n satisfies $\lim_n nP(\{X_0 > u_n\}) = \lambda$ for some $\lambda > 0$, then a very simple computation shows that the point process

$$N_n(B) = \sum_{t/n \in B} 1_{\{X_t > u_n\}}, \qquad B \in \mathcal{B} \tag{2}$$

(where $\mathcal{B}$ stands for the Borel σ-algebra of [0,1]), converges to a Poisson process of intensity λ.

If X is stationary and weakly dependent, clustering of exceedances may occur and one obtains a compound Poisson process. In this paper, if X is a random process such that (for some sequence u_n) the point process in (2) converges to a compound Poisson process, we shall say that X satisfies a compound Poisson limit theorem (CPLT). CPLTs are known for stationary processes satisfying some mixing conditions (Cohen 1989; Dziubdziela 1988; Ferreira 1993; Hsing *et al*. 1988; Leadbetter 1991; 1995; Leadbetter and Hsing 1990; Leadbetter and Nandagopalan 1989) as well as for Markov chains (Hsiau 1997). Some results are also available for weakly dependent but non-stationary X (Alpuim *et al*. 1995; Dziubdziela 1995; Hudson *et al*. 1989; Hüsler 1993). For a nice summary of many related results, see Falk *et al*. (1994). The authoritative text by Leadbetter *et al*. (1983) is a basic reference for exceedances, extremes and related topics, as are Leadbetter and Rootzén (1988) and Embrechts *et al*. (1997). For continuous-time results, see Volkonskii and Rozanov (1959; 1961), and Wschebor (1985) and the very nice monograph by Berman (1992). In some cases, rates of convergence can also be obtained by means of the Stein–Chein method: an extensive account is given in Barbour *et al*. (1992); see also Brown and Xia (1995). For the application of point-process exceedances to practical modelling of ozone data, see Smith and Shively (1995).

However, models like (1) can fail to satisfy the weak-dependence hypotheses required for those results. The aim of this paper is to prove that for model (1) the point process defined in (2) still has a compound Poisson limit. Our result generalizes the preceeding ones; firstly, our assumptions do not imply that X has a particular weak-dependence structure (such as mixing, association, Markov, etc.). Secondly, without additional effort we also obtain the limit distribution of N_n when Y (hence X) exhibits long-range dependence: in this case the limit distribution is no longer compound Poisson but a mixture of several compound Poisson distributions. Finally, we consider our approach interesting in itself, because the technique is based on the study of the high-level exceedances that belong to an 'irregular' set, and it is found that the geometry of this set plays a key role.

Roughly speaking, what we show here is that the addition of a component Y whose mean occupation measure has a limit in a weakly dependent model just averages the limits that are obtained for the weakly dependent case over irregular sets; if Y is ergodic, averaging will be non-random and a CPLT will hold; if Y is non-ergodic, a mixture of compound Poisson processes will be obtained. If we only look at the ergodic case, it is clear that we require that the regression model really depends on the weakly dependent component which makes a CPLT possible. If this weakly dependent component is negligible, our results fail to hold because the asymptotic behaviour is driven just by Y. This is, of course, a limitation of our approach, but we must also emphasize that we are only requiring that 'noise' (ξ) is not negligible, which seems reasonable in many situations.

The results presented here concerning the asymptotic distribution of the high-level exceedances over a collection of sets of irregular shape are, to the best of our knowledge, new; we know of no previous results determining the role played by the geometry of the collection. We extend the results of Perera (1997a; 1997b) for central limit theorems to the context of CPLTs.

This paper is organized as follows. Section 2 presents some basic notation and definitions and the statement of the main result (Theorem 1). The proof of Theorem 1 is based on the CPLT for weakly dependent stationary processes over irregular sets that is given in Section 3. The proof of the main result is given at the end of Section 3.

2. Definitions and statement of the main result

In this paper we will consider $\mathbb{R}^d$ equipped with the supremum norm, and C will denote a generic constant that may change from line to line. We shall also denote by C_s^d the combinatorial coefficients $C_s^d = d!/(d-s)!s!$, $0 \leqslant s \leqslant d$. An important role will be played by the coefficients

$$\Theta(s;\, d) = (-1)^{d-1} \sum_{j=0}^{s-1} C_j^d (-1)^j, \qquad 1 \leqslant s \leqslant d.$$

Recall that a point process N is a compound Poisson process with intensity measure ν (denoted by $CP(\nu)$), where ν is a positive finite measure on $\mathbb{N}$, if the following conditions hold:

- For any $h \in \mathbb{N}$, if $B_1, \ldots, B_h$ are disjoint Borel sets, then $N(B_1, \ldots, N(B_h)$ are independent.
- For any Borel set B, the Laplace transform of $N(B)$ is $L(B;\, s) = \exp(m(B)\sum_{j=1}^{\infty} \nu_j(\exp(-sj) - 1))$, where m denotes Lebesgue measure, and $\nu_j = \nu(\{j\})$ for all $j \in \mathbb{N}$.

For $A \subset \mathbb{N}$ and $n \in \mathbb{N}$, we set $A_n = A \cap [1, n]$. If $B \subset \mathbb{N}$, $\vec{r} \in \mathbb{N}^d$, $d \geqslant 1$, and $\mathcal{A} = (A^1, \ldots, A^d)$ is any (ordered) finite collection of subsets of $\mathbb{N}$, then let, for any $n \in \mathbb{N}$,

$$T_n(\vec{r};\, B) = \bigcap_{i=1}^{d} (B_n - r_i), \qquad G_n(\vec{r};\, \mathcal{A}) = \bigcap_{i=1}^{d} (A_n^i - r_i).$$

Definition 1. *Let A be a subset of $\mathbb{N}$. A is an* asymptotically ponderable set (APC) *if for any $d \geqslant 1$, $\vec{r} \in \mathbb{N}^d$, the following limit exists:*

$$\lim_n \frac{\mathrm{card}(T_n(\vec{r};\, A))}{n} = \tau(\vec{r};\, A).$$

If $(A^1, \ldots, A^h)$ *is a collection of subsets of* $\mathbb{N}$*, we will say that* $A, \ldots, A^h)$ *is an* asymptotically ponderable collection (APC) *if, for any* $d \geqslant 1$, $\vec{r} \in \mathbb{N}^d$, $\{i_1, \ldots, i_d\} \subset \{1, \ldots, h\}^d$ *and any subcollection* $\mathcal{A} = (A^{i_1}, \ldots, A^{i_d})$*, the following limit exists:*

$$\lim_n \frac{\mathrm{card}(G_n(\vec{r};\, \mathcal{A}))}{n} = \gamma(\vec{r};\, \mathcal{A}).$$

Remark 1. (a) A is an APS if and only if $\mathcal{A} = (A)$ is an APC. Asymptotic ponderability is hereditary: if a collection is an APC, so is any subcollection. A set A is an asymptotically measurable set (AMS) in the sense of Perera (1997a) if the convergence of Definition 1 holds for $d = 1, 2$. Therefore, the non-AMSs of Perera (1997a) are non-APSs. The following is an example of an AMS that is not an APS. Consider $U = (U_t)_{t\in\mathbb{N}}$ i.i.d. with common Bernoulli $(\frac{1}{2})$ law. It is easy to see that we can construct a 3-dependent stationary process $V = (V_t)_{t\in\mathbb{N}}$, independent of U, whose two-dimensional laws are identical to those of U but such that (U_1, U_2, U_3) and (V, V_2, V_3) have different laws. For any $n \in \mathbb{N}$, define $I(n) = [100^{2^{n-1}}, 100^{2^n})$ and set $B = \cup_{r=0}^{\infty} I(2r)$; finally, define

$$A(\omega) = \{n \in B\colon U_n(\omega) = 1\} \bigcup \{n \in B^c\colon V_{n(\omega)} = 1\}.$$

Straightforward computations show that A is, with probability one, an AMS, with $\tau(r;\, A) = \frac{1}{4}$ for all $r \geqslant 1$. On the other hand, for $\bar{r} = (0, 1, 2) \in \mathbb{N}^3$, it is easy to show that

$$\limsup_n \frac{\mathrm{card}\{T_n(\bar{r};\, A)\}}{n} \geqslant \frac{13}{96} \geqslant \frac{1}{8} \geqslant \liminf_n \frac{\mathrm{card}\{T_n(\bar{r};\, A)\}}{n} \text{almost surely.}$$

Therefore A is a non-APS with probability one.

(b) Consider $\{Y_t : t \in \mathbb{N}\}$ is a stationary and ergodic random process, such that Y_0 takes values in $\{1, \ldots, k\}$ and let $A^j(\omega) = \{t \in \mathbb{N} : Y_t(\omega) = j\}$. Then, by the ergodic theorem, $\mathcal{A} = (A^1, \ldots, A^k)$ is, almost surely, an APC and $\gamma(\bar{r};\, \mathcal{A}) = \mathrm{E}(\Pi_{i=1}^{i=m} 1_{\{y_{r_i}=j\}})$, $\bar{r} \in \mathbb{N}^m$, $m \in \mathbb{N}$.

Definition 2. *We will say that a real-valued process* Y *is* ponderable *if for every* $d \in \mathbb{N}$, $\bar{r} \in \mathbb{N}^d$*, there exists a (random) probability measure* $\mu^{\bar{r}}(\cdot)(\omega))$ *defined on the Borel sets of* $\mathbb{R}^d$ *such that if* $B_1, \ldots, B_d$ *are Borel real sets then the (random) collection* $\mathcal{A}((B_1, \ldots, B_d))(\omega) = (A^1(\omega), \ldots, A^d(\omega))$ *defined by*

$$A^j(\omega) = \{t \in \mathbb{N} : Y_t(\omega) \in B_j\}$$

is an APC almost surely with $\gamma(\vec{r}, \mathcal{A}((B_1, \ldots, B_d))(\omega)) = \mu^{\bar{r}}(B_1 \times B_2 \times \cdots \times B_d)(\omega)$*. If, in addition, the measures* $\mu^{\bar{r}}$ *are non-random, we will say that* Y *is* regular.

Remark 2. Let Y be a real-valued process and fix $k \in \mathbb{N}$ and $\bar{r} \in \mathbb{N}^k$, $h \in \mathbb{N}$; it is clear from its definition that $\gamma(\vec{r}, \mathcal{A}((B_1, \ldots, B_k))(\omega))$ does not depend on $Y_t : \|t\| \leqslant h)$ (a finite set of coordinates does not affect limits of averages), and hence we deduce that $\mu^{\vec{r}}$ is measurable with respect to the σ-algebra

$$\sigma_{\infty}^{Y} = \bigcap_{h=1}^{\infty} \sigma(Y_t : \|t\| \geqslant h).$$

Therefore, if σ_{∞}^{Y} is trivial, then Y is regular.

Remark 3. Y ponderable means that for any $B_1, \ldots, B_d$ Borel real sets their (mean) asymptotic occupation measure is defined almost surely as

$$\mu^{\vec{r}}(B_1 \times B_2 \times \cdots \times B_k)(\omega) = \lim_n \frac{1}{n} \sum_{t=0}^{n} 1_{\{Y_{t+r_j}(\omega) \in B_j, j=1,\ldots,h\}}.$$

In this way, a process is regular when a deterministic mean occupation measure exists.

Example 1. By the ergodic theorem (see Guyon 1995, p. 108), if Y is stationary, then Y is ponderable. We have already seen that if, in addition, σ_{∞}^{Y} is trivial, then Y is regular. In particular, this is the case if Y satisfies a Marcinkiewicz–Zygmund inequality. More precisely, we say that a centred random process $Y = \{Y_t : t \in \mathbb{N}\}$ satisfies a Marcinkiewicz–Zygmund inequality of order $q > 2$ if for any $d \geqslant 1$, $\vec{r} \in \mathbb{N}^d$, there exists a constant $C(\vec{r}, q)$ such that, for any function $f : \mathbb{R}^d \to \mathbb{R}$ bounded by 1, the inequality

$$\mathrm{E}\left\{ \left(\sum_{t=1}^{N} [f(Y_t(\vec{r})) - \mathrm{E}\{f(Y_t(\vec{r}))\}] \right)^q \right\} \leqslant C(\vec{r}, q) N^{q/2}$$

holds, where $Y_t(\vec{r}) = (Y_{t+r_1}, Y_{t+r_2}, \ldots, Y_{t+r_d})$. We refer to Doukhan (1995) for an overview of different contexts where these inequalities apply; see also Bryc & Smolenski (1993). Further, if Y is non-stationary, but it satisfies a Marcinkiewicz–Zygmund inequality of order $q > 2$ and there exists a probability measure $\mu^{\vec{r}}$ such that for any Borel sets $B_1, \ldots, B_k$ we have

$$\lim_n \frac{1}{n} \sum_{m=1}^{n} \mathrm{E}\left(\prod_{i=1}^{m} 1_{\{Y_{m+r_i} \in B_j\}} \right) = \mu^{\vec{r}}(B_1 \times B_2 \times \cdots \times B_k),$$

then a simple Borel–Cantelli argument proves that

$$\gamma(\vec{r}; \mathcal{A}((B_1, \ldots, B_k))) = \mu^{\vec{r}}(B_1 \times B_2 \times \cdots \times B_k) \qquad \forall \vec{r} \in \mathbb{N}^m, \forall m,$$

and Y is regular.

Now we turn our attention to the component ξ of model (1).

Definition 3. *Let ξ be a real-valued random process and $\varphi : \mathbb{R}^2 \to \mathbb{R}$ a measurable function. We will say that ξ is* φ-noise *if for every finite $h \in \mathbb{N}$, any vector $(y_1, \ldots, y_h) \in \mathbb{R}^h$ and any* APC $\mathcal{A} = (A^1, \ldots, A^h)$*, the random process*

$$X_t = \sum_{j=1}^{h} \varphi(\xi_t, y^j) 1_{A^j}(t) \tag{3}$$

satisfies the CPLT.

Introduce the notation $\varphi(\xi(\vec{r}), \vec{y}) = (\varphi(\xi_{r_1}, y_1), \ldots, \varphi(\xi_{r_d}, y_d))$ for $\vec{r} \in \mathbb{N}_L^d$, $\vec{y} \in \mathbb{R}^d$; if $J \subset \mathbb{R}$ and $d \geqslant 1$, write $J_L^d = \{(j_1, \ldots, j_d) : j_i \in J, j_i < j_{i+1}, i = 1, \ldots, d-1\}$. If $\vec{r} \in \mathbb{N}^d$, $J \subset \mathbb{N}$, for any random process V and $u > 0$, let

$$\{V(\vec{r}) > u\} := \bigcap_{i=1}^{d} \{V_{r_i} > u\}, \qquad \{V(J) > u\} := \bigcap_{t \in J} \{V_t > u\}.$$

Definition 4. *Let X be a real-valued random process. We will say that X admits an* I-decomposable regression *if there exists a ponderable process Y, a measurable function φ and a φ-noise ξ independent of Y such that $X_t = \varphi(\xi_t, Y_t)$ for all $t \in \mathbb{N}$, and the following conditions are fulfilled:*

(a) *For all $K > 0$,*

$$\limsup_{\delta \to 0} \left(\limsup_{n} \left(\sup_{|x-z| \leqslant \delta, |x| \leqslant K} nP(\{\varphi(\xi_0, x) > u_n\} \nabla \{\varphi(\xi_0, z) > U_n\}) \right) \right) = 0,$$

where $A \nabla B = (A^c \cap B) \cup (A \cap B)$; and

$$\limsup_{K} \left(\limsup_{n} \left(\sup_{|x| > K} nP(\{\varphi(\xi_0, x) > u_n\}) \right) \right) < \infty.$$

(b) *For all $x \in \mathbb{R}$, the following limit exists:*

$$\lim_{n} nP(\{\varphi(\xi_0, x) > u_n\}) = \lambda(x).$$

(c) *For all $d \in \mathbb{N}$, $\vec{y} \in \mathbb{R}^d$, $\vec{r} \in \mathbb{N}_L^d$, the limit*

$$\lim_{n} P(\{\varphi(\xi(\vec{r}), \vec{y}) > u_n\} | \{\varphi(\xi_0, y_0\} > u_n\}) = a(\vec{r}, \vec{y})$$

exists, and $a(\vec{r}, \cdot)$ is continuous for all $\vec{r}$.

(d) *For any $s \in \mathbb{N}$,*

$$\sum_{d=s}^{\infty} |\Theta(s; d)| \sum_{\vec{r} \in \mathbb{N}_L^d} \sup_{y \in \mathbb{R}^d} a(\vec{r}, \vec{y}) < \infty, \qquad \lim_{s \to \infty} \sum_{d=s}^{\infty} |\Theta(s; d)| \sum_{\vec{r} \in \mathbb{N}_L^d} \sup_{y \in \mathbb{R}^d} a(\vec{r}, \vec{y}) = 0.$$

Remark 4. A straightforward computation shows that condition (a) implies that the function λ defined in (b) is uniformly continuous and bounded.

Remark 5. The reader may ask where the 'I' in 'I-decomposable' comes from. It comes from 'Independent': keep in mind that X is a process that we can decompose into two *independent* random components, one mainly 'local' and 'weakly dependent' (ξ) and another which we can control 'in the mean' (Y).

Example 2. Consider $U = (U_t)_{t\in\mathbb{N}}$ i.i.d. with common absolutely continuous law μ. Let V be a random variable assuming a finite number of values ($V \in \{1, \ldots, S\}$) and independent of U, and let $(a_t)_{t\in\mathbb{N}}$ be a sequence of real numbers satisfying $\lim_t a_{kt+h} = a(h)$, $h = 0, 1, \ldots, k-1$. Define $Y_t = U_t + a_t V$. Then a straightforward computation shows that Y is ponderable, and Y is regular if and only if $a(h) = 0$ for any h. Now consider $\varphi(\xi, y) = \xi g(y)$ with g a real bounded and positive function and ξ a moving average of i.i.d. Cauchy variables. It is easy to check that $X_t = \varphi(\xi_t, Y_t)$ satisfies Definition 1.

The main result of this paper is the following:

Theorem 1. *Suppose that X admits an I-decomposable regression.*

(a) *If Y is regular, X satisfies the CPLT. More precisely, if* $X = \varphi(\xi, Y)$, *and*

$$N_n(B) = \sum_{t/n\in B} 1_{\{X_t > u_n\}}, \qquad B \in \mathcal{B},$$

then N_n *converges in law to N, a compound Poisson with Laplace transform*

$$L(B; x) = \exp\left(m(B) \sum_{j=1}^{\infty} \nu_j (\mathrm{e}^{-xj} - 1) \right),$$

$$\nu_j = \sum_{d=j}^{\infty} (-1)^{j+d} C_j^d \int_{\mathbb{R}^d} \sum_{\bar{r}\in\mathbb{N}_L^{d-1}} a(\vec{r}, \vec{y}) \lambda(y_0) \mu^{\bar{r}}(\mathrm{d}y), \qquad \forall j \in \mathbb{N}.$$

(b) *If Y is not regular, then* $N_n | Y$ *satisfies the CPLT and* N_n *converges weakly to a mixture of compound Poisson processes.*

3. Compound Poisson limit theorems over asymptotically ponderable collections, and proof of the main result

Consider first a stationary process $\xi = \{\xi_n : n \in \mathbb{N}\}$, an APS A, and let

$$X_t = \xi_t 1_A(t). \tag{4}$$

Remember that the sequence u_n satisfies

$$\lim_n nP(\{\xi_0 > u_n\}) = \lambda > 0. \tag{5}$$

If B is any set and $k \in \mathbb{N}$, then $C_k(B)$ will denote the collection of all the subsets of B with k elements, i.e., $C_k(B) = \{D \subset B : \mathrm{card}(D) = k\}$.

Lemma 1. *Let B be a subset of* $\mathbb{N}$, *let* X *be as in* (4), $u > 0$, *and define*

$$N_n^*(B) = \sum_{t \in B_n} 1_{\{X_t > u\}}. \tag{6}$$

Then

$$P(\{N_n^*(B) \geqslant s\}) = \sum_{d=s}^{\mathrm{card}(B_n \cap A)} \Theta(s;\, d) \sum_{\vec{r} \in [1, \mathrm{card}(B_n \cap A)]_L^{d-1}} P(\{\xi(\vec{r}) > u\})\, \mathrm{card}(T_n(\vec{r};\, B \cap A)), \tag{7}$$

where

$$\Theta(s;\, d) = \sum_{k=1}^{\infty} (-1)^{k-1} \theta(k,\, s;\, d), \qquad \theta(k,\, s;\, d) = \mathrm{card}(\mathcal{C}_k)\mathcal{C}_s(\{1,\, \ldots,\, d\}))). \tag{8}$$

Remark 6. In fact the sum in (8) is finite: if $\theta(k,\, s;\, d) > 0$, then there exists a decomposition $\{1,\, \ldots,\, d\} = \cup_{j=1}^{k} I_j$, with $\mathrm{card}(I_j) = s$, $I_j \neq I_h$ if $j \neq h$; this implies that $sk \geqslant d$, $k \leqslant C_s^d$. Therefore, $\theta(k,\, s;\, d) = 0$ if $k > d/s$ or $k > C_s^d$. In particular, if $s > d$, then $\Theta(s;\, d) = 0$. On the other hand, if $d = s = 1$, the corresponding term on the sum (7) nust be interpreted as $P(\{\varepsilon_0 > u\})\, \mathrm{card}(B_n \cap A)$.

Proof of Lemma 1. First, observe that with our notation the elementary inclusion–exclusion formula is

$$P\left(\bigcup_{\gamma \in \Gamma} A_\gamma\right) = \sum_{k=1}^{\mathrm{card}(\Gamma)} (-1)^{k=1} \sum_{C \in C_k(\Gamma)} P\left(\bigcap_{\gamma \in C} A_\gamma\right),$$

for any finite Γ. Therefore, we have

$$\begin{aligned}
P(\{N_n^*(B) \geqslant s\}) &= P\left(\bigcup_{I \in C_s(B_n \cap A)} \{\xi(I) > u\}\right) \\
&= \sum_{k=1}^{\mathrm{card}(\mathcal{C}_s(B_n \cap A))} (-1)^{k-1} \sum_{C \in C_k(\mathcal{C}_s(B_n \cap A))} P\left(\bigcap_{I \in C} \{\xi(I) > u\}\right) \\
&= \sum_{k=1}^{\mathrm{card}(\mathcal{C}_s(B_n \cap A))} (-1)^{k-1} \sum_{C \in \mathcal{C}_s(B_n \cap A))} P\left(\left\{\xi\left(\bigcup_{I \in C} I\right) > u\right\}\right) \\
&= \sum_{k=1}^{\mathrm{card}(\mathcal{C}_s(B_n \cap A))} (-1)^{k-1} \sum_{d=1}^{\infty} \sum_{H \in \mathcal{C}_d(B_n \cap A)} P(\{\xi(H) > u\}) \\
&\quad \times \mathrm{card}\left(\left\{C \in \mathcal{C}_k(\mathcal{C}_s(B_n \cap A)) : \bigcup_{I \in C} I = H\right\}\right).
\end{aligned}$$

But if $H \in \mathcal{C}_d(B_n \cap A)$, then

$$\operatorname{card}\left(\left\{C \in \mathcal{C}_k(\mathcal{C}_s(B_n \cap A)) : \bigcup_{I \in C} I = H\right\}\right) = \theta(k, s; d)$$

and, on the other hand, by the stationary of ξ, we obtain

$$P(\{\xi(H) > u\}) = P(\{\xi(H)) > u_n\}),$$

where if $H = \{h_1, \ldots, h_d\}$ is such that $h_i < h_{i+1}$ for all i, then

$$\vec{r}(H) := (h_2 - h_1, \ldots, h_3 - h_1, h_d - h_1) \in [1, \operatorname{card}(B_n \cap A)]_L^{d-1}.$$

It follows that

$$\begin{aligned}
&P(\{N_n^*(B) \geqslant s\}) \\
&= \sum_{k=1}^{\operatorname{card}(\mathcal{C}_s(B_n \cap A))} (-1)^{k-1} \sum_{d=1}^{\infty} \theta(k, s; d) \sum_{\vec{r} \in [1, \operatorname{card}(B_n \cap A)]_L^{d-1}} P(\{\xi(\vec{r}) > u\}) \\
&\quad \operatorname{card}(\{H \in \mathcal{C}_d(B_n \cap A) : \vec{r}(H) = \vec{r}\}),
\end{aligned}$$

but an elementary argument proves that

$$\operatorname{card}(\{H \in \mathcal{C}_d(B_n \cap A); \vec{r}(H) = \vec{r}\}) = \operatorname{card}(T_n(\vec{r}; B \cap A))$$

and the lemma follows. □

Lemma 2. *If* $s \leqslant d$, *then* $\Theta(s; d) = (-1)^{d-1} \sum_{j=0}^{s-1} C_j^d (-1)^j$.

Proof. Pick an arbitrary $\rho \in (0, 1)$ and take $B = A = \mathbb{N}$, ξ i.i.d. and $u > 0$ such that $P(\{\xi_0 > u\}) = \rho$. Applying Lemma 1, we obtain

$$P(\{N_n^*(\mathbb{N}) \geqslant s\}) = \sum_{d=s}^{n} \Theta(s; d) \sum_{\vec{r} \in [1, n]_L^{d-1}} \rho^d \operatorname{card}(T_n(\vec{r}; \mathbb{N})). \tag{9}$$

On the other hand, using the fact that $N_n^*(\mathbb{N}) \sim Bin(n, p)$ and the binomial expansion for $(1 - \rho)^k$, we have that

$$P(\{N_n^*(\mathbb{N}) \geqslant s\}) = \sum_{m=s}^{n} C_m^n \rho^m (1 - \rho)^{n-m} = \sum_{m=s}^{n} C_m^n (-1)^m \left(\sum_{j=0}^{n-m} C_j^{n-m} (-\rho)^j\right) \rho^m. \tag{10}$$

Equating coefficients of both polynomial expressions, by an elementary computation, we arrive at

$$\Theta(s; d) = (-1)^d \sum_{j=s}^{d} C_j^d (-1)^j = (-1)^{d-1} \sum_{j=0}^{s-1} C_j^d (-1)^j. \qquad \square$$

Remark 7. Lemma 2 and a trivial computation show that, for s fixed and d tending to infinity, $\Theta(s;\, d) \approx (-1)^{d+s} d^{s-1}/(s-1)!$, and that, for any s, d, $|\Theta(s;\, d)| \leqslant d^{s-1}/(s-2)!$

We now introduce a weak-dependence hypothesis; we denote by $\mathrm{int}(x)$ the integer part of $x \in \mathbb{R}$.

(H1) There exist two non-decreasing sequences $(p_n)_{n\in\mathbb{N}}$, $(q_n)_{n\in\mathbb{N}}$, such that $\lim_{n\to\infty} p_n = \lim_{n\to\infty} q_n = \infty$, $\lim_{n\to\infty} p_n/n = \lim_{n\to\infty} q_n/p_n = 0$, and if $B_n^i = [(i-1)(p_n+q_n), (i-1)(p_n, q_n)+p_n)$ then, for any $r \in \mathbb{N}$ and $D_1, \ldots, D_r \subset \mathbb{N}$, $(N_n^*(D_1), \ldots, N_n^*(D_r))$ has the same asymptotic distribution as $(\hat{N}_n(D_1), \ldots, \hat{N}_n(D_r))$, where $\hat{N}_n(D_j) = \sum_{i=1}^{k_n} Z_n^i(D_j)$, $j = 1, \ldots, r$, with, $k_n = \mathrm{int}(n/p_n + q_n)$ and $(Z_n^i(D_1), \ldots, Z_n^i(D_r))_{1\leqslant i\leqslant k_n}$ independent copies of $(N_n^*(B_n^i \cap D_1), \ldots, N_n^*(B_n^i \cap D_r))_{1\leqslant i\leqslant k_n}$.

Remark 8. (a) It is very easy to check that mixing assumptions guarantee assumtion (H1). More precisely, let $s, t \in \mathbb{N}$ and $\Sigma_n(s,\, t) = \sigma(\{\{\xi_i > u_n\} : s \leqslant i \leqslant t\})$; define

$$\alpha_{n,l} = \sup\{|P(A\cap B) - P(A)P(B)| : A \in \Sigma_n(h,\, s+h),\, B \in \Sigma_n(s+h+l,\, n),$$
$$h \geqslant 0,\, s+l+h < n\}.$$

We will say that $N^* = \{N_n^*\}_{n\in\mathbb{N}}$ is strongly mixing if there exists a non-decreasing sequence $(q_n)_{n\in\mathbb{N}}$ such that $\lim_n q_n = \infty$ and $\lim_n \alpha_{n,q_n} = \lim q_n/n = 0$. It is straightforward to check that we can choose $(k_n)_{n\in\mathbb{N}}$ such that $\lim_n k_n = \lim n/(k_n+q_n) = \infty$ and $\lim_n k_n\alpha_{n,q_n} = 0$. Then, taking $p_n = n/k_n$, it follows that (H1) is satisfied. See Hsing *et al*. (1988, Lemma 2.2) for a detailed proof; see Bradley (1986) and Doukhan (1995) for mixing conditions, examples and covariance inequalities.

(b) Association is a weak-dependence structure that has been widely used as an alternative to mixing; see Newman (1980) and Roussas (1994) for the definition and basic properties of associated processes. It is easy to check that if ξ is associated and

$$\lim_n \sum_{m=q_n}^{\infty} n|\mathrm{cov}(1\{\xi_0 > u_n\},\, 1\{\xi_m > u_n\})| = 0,$$

then (H1) holds.

Similar conditions can be obtained for other weak-dependence structures. Doukhan and Louichi (1997) give a summary of different weak-dependence conditions under which (H1) can be proved in a similar way.

In what follows, we will make the following assumption

(H2) $\limsup_n \sum_{m=1}^{\infty} n|\mathrm{cov}(1\{\xi_0 > u_n\},\, 1\{\xi_m > u_n\})| < \infty$.

Proposition 1. *Let X be as in* (4)*, and assume that* (5)*,* (H1) *and* (H2) *hold. For any* $s \in \mathbb{N}$*, define*

$$Q_n(s) = \sum_{d=s}^{p_n} \Theta(s;\, d) \sum_{\vec{r} \in [1,\, p_n]_L^{d-1}} P(\{\xi(\vec{r}) > u_n\} | \{\xi_0 > u_n\}) F_n(\vec{r};\, A),$$

where

$$F_n(\vec{r};\, A) = \frac{1}{n} \sum_{\{1 \leqslant i \leqslant k_n : \mathrm{card}(B_n^i \cap A) \geqslant d\}} \mathrm{card}(T_n(\vec{r},\, B_n^i \cap A)), \qquad \vec{r} \in \mathbb{N}^{d-1}.$$

Assume further that, for any $s \in \mathbb{N}$,

$$\lim_n Q_n(s) = Q(s). \tag{11}$$

Then N_n *converges in law (as a process) to a* $CP(\nu)$ *process, where* $\nu_s = \lambda(Q(s) - Q(s+1))$ *for all* $s \in \mathbb{N}$.

Proof. By Theorem 4.2 of Kallenberg (1983), it suffices to show that for any k-tuple of semiclosed intervals $I_1, \ldots, I_k$, the random vector $N_n(I_1), \ldots, N_n(I_k))$ converges in law to $N(I_1), \ldots, N(I_k))$, where N is a $CP(\nu)$ process. Without loss of generality, we may assume that $I_1, \ldots, I_k$ are disjoint. In this case, by (H1), the coordinates of the random vector $N_n(I_1), \ldots, N_n(I_k))$ are asymptotically independent, and therefore it suffices to show that for any semiclosed interval I we have that $N_n(I)$ converges in law to $N(I)$. But an elementary argument shows that it is enough to consider the case $I = (0, a]$, $0 < a < 1$. Finally, this can be reduced, by a scale change, to checking that $N_n((0, 1]) \xrightarrow[n]{w} N((0, 1])$, where N is a random variable with Laplace transform $L(s) = \exp(\sum_{j=1}^{\infty} \nu_j(\mathrm{e}^{-sj} - 1))$. For the rest of this proof, $N_n((0, 1])$ will be denoted simply by N_n.

By (H1), N_n is asymptotically equivalent to $\hat{N}_n = \sum_{i=1}^{k_n} Z_n^i$, with $(Z_n^i)_{1 \leqslant i \leqslant k_n}$ independent copies of $(N_n^*(B_n^i))_{1 \leqslant i \leqslant k_n}$. We will first prove that $(\hat{N}_n)_{n \in \mathbb{N}}$ is tight. For this we will show that $\mathrm{E}(\hat{N}_n)$ and $\mathrm{var}(\hat{N}_n)$ are bounded, which implies the uniform integrability of $(\hat{N}_n)_{n \in \mathbb{N}}$, hence its tightness. $\mathrm{E}(\hat{N}_n)$ is obviously bounded by the convergent sequence $nP(\{\xi_0 > u_n\})$; therefore $\mathrm{E}(\hat{N}_n)$ is bounded. For the variance, we have

$$\begin{aligned}
\mathrm{var}(\hat{N}_n) &= \sum_{i=1}^{k_n} \mathrm{var}(N_n^*(B_n^i)) = \sum_{i=1}^{k_n} \sum_{s,t \in B_n^i \cap A} \mathrm{cov}(1_{\{\xi_0 > u_n\}},\, 1_{\{\xi_t - s > u_n\}}) \\
&= \sum_{i=1}^{k_n} \mathrm{card}(B_n^i \cap A) P(\{\xi_0 > u_n\})(1 - P(\{\xi_0 > u_n\})) \\
&\quad + 2 \sum_{i=1}^{k_n} \sum_{m=1}^{p_n} \mathrm{cov}(1_{\{\xi_0 > u_n\}}, 1_{\{\xi_m > u_n\}})\, \mathrm{card}(B_n^i \cap A \cap (B_n^i \cap A - m)).
\end{aligned}$$

The first term in the last expression is bounded by $nP(\{\xi_0 > u_n\})$ again, so it is bounded, and by (H2) and the fact that $\mathrm{card}(B_n^i) = p_n$, the last term is bounded. Therefore, the variance is bounded and tightness follows.

Given any subsequence $(\hat{N}_{n_h})_{h \in \mathbb{N}}$ pick a sub-subsequence that is weakly convergent to some random variable W. To simplify the notation, we will also call this second subsequence $(\hat{N}_{n_h})_{h\in\mathbb{N}}$. Since the law of W must be infinitely divisible and concentrated on $\mathbb{N}$, its Laplace transform is $H(s) = \exp(\sum_{j=1}^{\infty}\pi_j(\mathrm{e}^{-sj} - 1))$. Thus, it suffices to show that $\pi_j = \nu_j$ for all $j \in \mathbb{N}$. Observe that

$$P(\{N_n^*(B_n^i) \neq 0\}) \leqslant \sum_{k\in A\cap B_n^i} P(\{\xi_k > u_n\}) \leqslant p_n P(\{\xi_0 > u_n\}),$$

and therefore $\lim_n P(\{N_n^*(B_n^i) \neq 0\}) = 0$. But, form Theorem 3 of Rényi (1951) (see also Dziubdziela 1995) we deduce that $\pi_j = \lim_h \sum_{i=1}^{k_{n_h}} P(\{N_{n_h}^*(B_{n_h}^i) = j\})$ for all j. But (11) and Lemma 1 imply that $\nu_j = \lim_n \sum_{i=1}^{k_n} P(\{N_n^*(B_n^i) = j\})$. Therefore $\nu_j = \pi_j$ for all j, and the lemma is proved. □

Corollary 1. *Assume that A is an APS and that X is as in Proposition 1. Also make the following assumptions:*

(H3) For all $d \in \mathbb{N}$, for all $\vec{r} \in \mathbb{N}_L^{d-1}$, $\lim_n P(\{\xi(\vec{r}) > u_n\}|\{\xi_0 > u_n\}) = a(\vec{r})$.
(H4) For all $d \in \mathbb{N}$, for all $\forall \vec{r} \in \mathbb{N}_L^{d-1}$, $P(\{\xi(\vec{r}) > u_n\}|\{\xi_0 > u_n\}) = a_n(\vec{r}) + b_n(\vec{r})$.

Here

$$a_n(\vec{r}) \leqslant c(\vec{r}) \quad \forall n \in \mathbb{N}, \quad \vec{r} \in \mathbb{N}_L^{d-1},$$

$$\sum_{d=s}^{\infty} |\Theta(s;\, d)| \sum_{\vec{r}\in\mathbb{N}_L^{d-1}} c(\vec{r}) < \infty, \quad \lim_n \sum_{d=s}^{p_n} |\Theta(s;\, d)| \sum_{\vec{r}\in[1,\, p_n]_L^{d-1}} |b_n(\vec{r})| = 0.$$

Then N_n converges in law (as a process) to a CP(ν) process, where

$$\nu_s = \lambda \sum_{d=s}^{\infty} (-1)^{s+d} C_s^d \sum_{\vec{r}\in\mathbb{N}_L^{d-1}} a(\vec{r})\tau(\vec{r};\, A).$$

Proof. We have

$$Q_n(s) = \sum_{d=s}^{p_n} \Theta(s;\, d) \sum_{\vec{r}\in[1,\, p_n]_L^{d-1}} a_n(\vec{r})F_n(\vec{r};\, A) + \sum_{d=s}^{p_n} \Theta(s;\, d) \sum_{\vec{r}\in[1,\, p_n]_L^{d-1}} b_n(\vec{r})F_n(\vec{r},\, A).$$

since F_n is bounded by 1, the second term on the right-side of this expression converges, by (H4), to zero. Therefore, it suffices to show that

$$\lim_n \sum_{d=s}^{p_n} \Theta(s;\, d) \sum_{\vec{r}\in[1,\, p_n]_L^{d-1}} a_n(\vec{r})F_n(\vec{r};\, A) = \sum_{d=s}^{\infty} \Theta(s;\, d) \sum_{\vec{r}\in\mathbb{N}_L^{d-1}} a(\vec{r})\tau(\vec{r};\, A) \tag{12}$$

and that

$$\sum_{d=s}^{\infty}\Theta(s;\, d)\sum_{\vec{r}\in\mathbb{N}_L^{d-1}} a(\vec{r})\tau(\vec{r};\, A) - \sum_{d=s+1}^{\infty}\Theta(s+1;\, d)\sum_{\vec{r}\in\mathbb{N}_L^{d-1}} a(\vec{r})\tau(\vec{r};\, A)$$
$$= \sum_{d=s}^{\infty}(-1)^{s+d}C_s^d \sum_{\vec{r}\in\mathbb{N}_L^{d-1}} a(\vec{r})\tau(\vec{r};\, A). \quad (13)$$

Let us first consider (12). Assume that we have shown that

$$\lim_n F_n(\vec{r};\, A) = \tau(\vec{r},\, A), \qquad \forall \vec{r}\in\mathbb{N}^m,\ \forall m. \quad (14)$$

Then, by (H3), (H4) and dominated convergence, (12) follows.

Turning to (14), fix $d\in\mathbb{N}$ and $\vec{r}\in\mathbb{N}^{d-1}$ and let n be big enough so that $q_n > 2\|\vec{r}\|$. Then, if $i\neq h$, $(B_n^i - r_j)\cap(B_n^h - r_k) = \varnothing$, for all $j,\, k$. Thus,

$$\mathrm{card}(T_n(\vec{r};\, A)) = \sum_{i=1}^{k_n}\mathrm{card}(T_n(\vec{r};\, B_n^i\cap A)) + \Delta_n(\vec{r}), \quad (15)$$

where $\Delta_n(\vec{r})$ is the sum of the cardinals of all the sets of the form

$$\bigcap_{i=1}^{i=d-1}(K_i - r_i)\cap K_0 \quad (16)$$

where there are two alternatives for each K_i

- $B_n^j\cap A$ for some j, and if for some i, $K_i = B_n^j\cap A$, then there is no h such that $K_h + B_n^k\cap A$, with $j\neq k$.
- $H_n^j\cap A$ for some j, where H_n^j is one of the 'holes' $[(j-1)(p_n+q_n)+p_n,\, j(p_n+q_n)\wedge n)$. Furthermore, this choice is made for at least one i.

Since the cardinality of the 'holes' is at most q_n the cardinality of a set of the form (16) is bounded by q_n. Taking into account that $\Delta_n(\vec{r})$ is the sum of at most $(k_n+1)(d-1)$ cardinals of sets of the form (16), we conclude that

$$\max_{\|\vec{r}\|<q_n/2}\Delta_n(\vec{r}) \leqslant q_n(k_n+1)(d-1). \quad (17)$$

Applying (17) and Definition 1 in (15), we obtain

$$\lim_n \frac{1}{n}\sum_{i=1}^{k_n}\mathrm{card}(T_n(\vec{r};\, B_n^i\cap A)) = \tau(\vec{r};\, A).$$

But

$$\left|\frac{1}{n}\sum_{i=1}^{k_n}\mathrm{card}(T_n(\vec{r};\, B_n^i\cap A)) - F_n(\vec{r};\, A)\right| \leqslant \frac{1}{n}\sum_{1\leqslant i\leqslant k_n\ :\mathrm{card}(B_n^i\cap A)\leqslant d}\mathrm{card}(T_n(\vec{r},\, B_n^i\cap A)) \leqslant \frac{dk_n}{n}\underset{n}{\rightarrow} 0,$$

and (14) follows.

Finally, we address (13). We have that

$$\sum_{d=s}^{\infty}\Theta(s;\,d)\sum_{\vec{r}\in\mathbb{N}_L^{d-1}}a(\vec{r}\,)\tau(\vec{r};\,A)-\sum_{d=s+1}^{\infty}\Theta(s+1;\,d)\sum_{\vec{r}\in\mathbb{N}_L^{d-1}}a(\vec{r}\,)\tau(\vec{r};\,A)$$

$$=\Theta(s,\,s)\sum_{\vec{r}\in\mathbb{N}_L^{s-1}}a(\vec{r}\,)\tau(\vec{r};\,A)+\sum_{d=s+1}^{\infty}(\Theta(s;\,d)-\Theta(s+1;\,d))\sum_{\vec{r}\in\mathbb{N}_L^{d-1}}a(\vec{r}\,)\tau(\vec{r};\,A).$$

But, by Lemma 2,

$$\Theta(s;\,d)-\Theta(s+1;\,d)=\sum_{j=0}^{s-1}C_j^d(-1)^{j+d-1}-\sum_{j=0}^{s}C_j^d(-1)^{j+d-1}=-C_s^d(-1)^{s+d-1}=C_s^d(-1)^{s+d}$$

and, taking into account that $\sum_{j=0}^{s}C_j^s(-1)j=0$, we obtain

$$\Theta(s,\,s)=\sum_{j=0}^{j=s-1}C_j^s(-1)^{j+s-1}=(-1)^{s+s}C_s^s=1$$

and deduce (13). □

We now present some examples of this result.

Example 3. Assume that ξ is a stationary m-dependent process and that (5) holds. Then ξ satisfies (H1) for any $(p_n)_{n\in\mathbb{N}}, (q_n)_{n\in\mathbb{N}}$, such that $\lim_n p_n=\lim_n q_n=\infty$, $\lim_n p_n/n=\lim_n q_n/p_n=0$. In particular, we will choose $p_n=\sqrt{\log n}$. the Cauchy–Schwarz inequality gives (H2). Let A be and APS and take X as in (4). Assume that (H3) holds. We will prove that (H4) holds, and therefore Corollary 1 applies.

Assume that $\vec{r}\in\mathbb{N}_L^{d-1}$ is such that $\|\vec{r}\,\|>m$. Then $r_{d-1}>m$ and

$$P(\{\xi(\vec{r}\,)>u_n\}|\{\xi_0>u_n\})\leqslant P(\{\xi_0>u_n,\,\xi_{r_{d-1}}>u_n\}|\{\xi_0>u_n\})=P(\{\xi_0>u_n\}),$$

which goes to zero with n.

Now define

$$a_n(\vec{r}\,)=P(\{\xi(\vec{r}\,)>u_n\}|\{\xi_0>u_n\})1_{\{\|\vec{r}\,\|\leqslant m\}},$$

$$b_n(\vec{r}\,)=P(\{\xi(\vec{r}\,)>u_n\}|\{\xi_0>u_n\})1_{\{\|\vec{r}\,\|>m\}},\quad c(\vec{r}\,)=1_{\{\|\vec{r}\,\|\leqslant m\}}.$$

Observe first that if $\vec{r}\in\mathbb{N}_L^{d-1}$ and $d-1>m$, then $\|\vec{r}\,\|>m$ and

$$\sum_{d=s}^{\infty}\Theta(s;\,d)\sum_{\vec{r}\in\mathbb{N}_L^{d-1}}c(\vec{r}\,)=\sum_{d=s}^{m+1}\Theta(s;\,d)\mathrm{card}([1,\,m]_L^{d-1})<\infty.$$

On the other hand, we have, by Remark 3 and (5),

$$\sum_{d=s}^{p_n} |\Theta(s;\, d)| \sum_{\vec{r}\in[1,\, p_n]_L^{d-1}} b_n(\vec{r}) \leqslant C \sum_{d=s}^{p_n} d^{s-1}\mathrm{card}([1,\ p_n]_L^{d-1}) P(\{\xi_0 > u_n\})$$

$$\leqslant \frac{C}{n}\sum_{d=s}^{p_n} d^{s-1}\frac{p_n^{d-1}}{(d-1)!} \leqslant \frac{C}{n} p_n^{s-1}\exp(p_n) \underset{n}{\to} 0.$$

Hence, if ξ is stationary and m-dependent, A and APS, and (4), (5) and (H3) hold then the CPLT holds, and the limit is described by

$$\nu_s = \sum_{d=s}^{m+1} (-1)^{s+d} C_s^d \sum_{\vec{r}\in[1,m]L^{d-1}} a(\vec{r})F(\vec{r};\, A).$$

Example 4. Let us present a very simple case where the limit is a compound Poisson process, but not Poisson. Let A be an APS such that $F(1;\, A) > 0$. Consider an i.i.d. sequence $(\zeta_k)_{k\in\mathbb{N}}$ which follows the Cauchy distribution. Set $\xi_k = \zeta_k + \zeta_{k+1}$. It is clear that ξ is stationary and 1-dependent. In addition, ξ_0 has the same distribution as $2C$, where C stands for a Cauchy variable. Indeed, as we have seen at the beginning of this section, we only need to show (H3) for $\|\vec{r}\| \leqslant m$. Setting X as in (4) and u_n as in (5), it suffices to show that

$$\lim_n P\{\xi_1 > u_n\} | \{\xi_0 > u_n\}) = a(1) \in \left[0, \tfrac{1}{2}\right]. \tag{17}$$

Indeed, if (17) holds, then $\nu_1 = \lambda(F(0;\, A) - 2a(1)F(1;\, A)) > 0$, $\nu_2 = \lambda a(1)F(1;\, A) > 0$, $\nu_s = 0$ for all $s \geqslant 3$. Let us now prove (17). An elementary computation shows that

$$\frac{\mathrm{d}}{\mathrm{d}u} P(\{\xi_0 > u\}) = \frac{-2}{\pi(4+u^2)} := g(u),$$

$$\frac{\mathrm{d}}{\mathrm{d}u} P(\{\xi_0 > u,\, \xi_1 > u\})$$

$$= (-1)\left(\int_{-\infty}^{\infty} f(x)f(u-x)q(x)\mathrm{d}x + \int_{-\infty}^{\infty} f(x)\int_{u-x}^{\infty} f(y)f(u-y)\mathrm{d}y\,\mathrm{d}x\right) := G(u),$$

where $f(x) = 1/(\pi(1+x^2))$ and $q(x) = \int_x^\infty f(y)\mathrm{d}y$. Using the fact that if $x \geqslant u/2$, then $(4+u^2)f(u) \leqslant 4/\pi$ and dominated convergence, it follows that $\lim_{u\to\infty} G(u)/g(u) = \frac{1}{2}$. This implies (17) for $a(1) = \frac{1}{2}$.

Example 5. Consider now a stationary process ξ, an APS A, and assume that (4), (5) and (H3) hold. With the notation of Remark 8(a), define

$$\phi(l) = \sup\{|P(A_i|A_{1-i}) - P(A_i)| i = 0, 1;$$
$$A_0 \in \Sigma_n(h,\, s+h),\ A_1 \in \Sigma_h(s+h+l,\, n),\ h \geqslant 0,\ s+l+h < n,\ n \geqslant 1\}.$$

Assume that

(H5) $\sum_{l=1}^{\infty}\phi(l) = \Phi < 1$.

For any $d > 2$ and $\vec{r} \in \mathbb{N}^{d-1}$, set $r_0 = 0$ and $\pi_i(\vec{r}) = (r_0, r_1, \ldots, r_i)$ $i = 0, 1, \ldots, d-1$. Under (H5), (H1) is obvious and (H2) follows from standard covariance inequalities for mixing processes (see Bradley 1986; Doukhan, 1995). Condition (H5) is very strong, but it can be checked, for instance, for some Markov Chains (Doukhan, 1995). If we also assume

(H6) $\lim_n P(\{\xi(\pi_i(\vec{r})) > u_n\} | \{\xi(\pi_{i-1}(\vec{r})) > u_n\}) = a_i(\vec{r}) \qquad \forall i, \vec{r}, d-1,$

then it can be easily shown that (H3) and (H4) hold, and hence Corollary 1 applies.

Remark 9. Consider now an $\mathbb{R}^h$-valued stationary process $\vec{\xi} = (\vec{\xi}^1, \ldots, \vec{\xi}^h)$, an APC $A^1, \ldots, A^h$, and define

$$X_t = \sum_{j=1}^{h} \xi_t^j 1_{A^j}(t). \tag{18}$$

Without major changes, Proposition 1 and Corollary 1 can be extended to this more general context.

We now give the proof of the main result of this paper.

Proof of Theorem 1. (a) Consider first the case where Y is regular. Once again using Theorem 4.2 of Kallenberg (1983), we have to prove that for any $I_1, \ldots, I_k$ disjoint semiclosed intervals we have that $(N_n(I_1), \ldots, N_n(I_k))$ converges in law to $(N(I_1), \ldots, N(I_k))$, where N is a $CP(\nu)$ process. For the sake of simplicity, we will present here the case $k = 1$, but the general case is obtained by very similar arguments. Further, without loss of generality, we will set $I_1 = (0, 1]$ and prove that $N_n((0, 1]) \underset{n}{\overset{w}{\rightarrow}} N((0, 1])$, where N is a random variable with Laplace transform $L(s) = \exp(\sum_{j=1}^{\infty} \nu_j (e^{-sj} - 1))$.

Assume first that Y takes values on a finite set $y_1, \ldots, y_k$. Restrict the probability space to a set of probability one where the almost sure convergence of Definition 2 holds. Conditioning with respect to Y, since ξ and Y are independent, the law of the exceedances point process for X (N_n) is the same as the law of the point process $\overline{N}_n$ corresponding to the exceedances of a process defined by (18) for $\xi^i = \varphi(\xi_0, y_i)$, $A^i(\omega) = \{t \in \mathbb{N} : Y_t(\omega) = y_i\}$. Then, by Definition 3, the result follows.

Assume now that the result holds for Y bounded. Consider an unbounded Y and let Y^K be the truncation of Y by K, $Y_n^K = c_K(Y_n)$, where $c_K(x) = x$ for $|x| < K$, $c_K(x) = K \operatorname{sgn}(x)$ for $|x| > K$; it is obvious that Y^K itself is ponderable. Now denote by N_n^K the exceedance point process for $X^K = \varphi(\xi, Y^K)$. Since Y^K is finite-valued, N_n^K converges to N^K, whose Laplace transform will be denoted by L^K.

$$\text{But } \mathrm{E}(|N_n - N_n^K|) \leqslant \sum_{i=1}^{n} P(\{\varphi(\xi_i, Y_i) > u_n\} \nabla \{\varphi(\xi_i, Y_i^K) > u_n\})$$

$$\leqslant \sum_{i=1}^{n} P([\{\varphi(\xi_i, Y_i) > u_n\} \nabla \{\varphi(\xi_i, Y_i^K) > u_n\}] \cap \{|Y_i| > K\}) \tag{19}$$

$$\leqslant \sum_{i=1}^{n} \int_{|y|>K} P([\{\varphi(\xi_i, y) > u_n\} \nabla \{\varphi(\xi_i, \mathrm{sgn}(y)K) > u_n\}] \mathrm{d}P^{Y_i}(y)$$

$$\leqslant n2 \sup_n \sup_{|y|>K} nP(\{\varphi(\xi_0, y) > u_n\}) \frac{1}{n} \sum_{i=1}^{n} P(|Y_i| > K).$$

where in (19) we are conditioning with respect to Y. Then, by condition (a) of Definition 4,

$$\lim_K \limsup_n \mathrm{E}(|N_n - N_n^K|) \leqslant C \lim_K \mu^0([-K, K]^c) = 0. \tag{20}$$

On the other hand, $L^K(x) = \exp(\sum_{j=1}^{\infty} \nu_j^K (\mathrm{e}^{-xj} - 1))$, with $\nu_s^K = \tau(s)^K - \tau(s+1)^K$ for all $s \in \mathbb{N}$, where

$$\tau(s)^K = \sum_{d=s}^{\infty} \Theta(s; d) \int_{\mathbb{R}^{d-1}} \sum_{\vec{r} \in \mathbb{N}_L^{d-1}} a(\vec{r}, \vec{y}) \lambda(y_0) \mu^{\vec{r}} \circ c_K^{-1}(\mathrm{d}y);$$

after some elementary computations, it follows from Definition 4 and dominated convergence that $\lim_{K\to\infty} L^K = L$. From this and (20) we obtain the theorem for Y unbounded.

It now suffices to show the result for Y bounded. Without loss of generality, we may assume that Y takes values on $[0, 1)$. For $H \in \mathbb{N}$, define

$$d_H(x) = \sum_{i=1}^{H} \frac{i}{H} 1_{[(i-1)/H, i/H]}(x),$$

and let $Y_n(H) = d_H(Y_n)$. Then Y^H it also ponderable and takes values on a finite set. So the result applies to $X_t(H) = \varphi(\xi_n, Y_n(H))$. Again using Definition 4 and dominated convergence, we show that the result applies for Y bounded.

(b) if Y is non-regular, we can easily see that the limit of $N_n|Y$ can be obtained in the same way, but since this limit depends on Y, the asymptotic distribution of N_n is a mixture of compound Poisson laws. □

Acknowledgements

The first author was partially supported by ADEME and AIRPARIF. The authors wish to express their gratitude to Didier Dacunha-Castelle, who introduced them to the problem discussed in this paper, to the Probability and Statistics team at Orsay for their warm support, to Joaquín Ortega and Omar Gil for their careful revision and to an anonymous referee for his/her very detailed and insightful comments.

References

Alpuim, M., Catkan, N. and Hüsler, J. (1995) Extremes and clustering of nonstationary max-AR(1) sequences. *Stochastic Process. Appl.*, **56**, 171–184.

Barbour, A., Holst, L. and Janson, S. (1992) *Poisson Approximation*. Oxford: Oxford University Press.

Berman, S. (1992) *Sojourns and Extremes of Stochastic Processes*. Pacific Grove, CA: Wadsworth & Brooks/Cole.

Bradley, R. (1986) Basic properties of strong mixing conditions. In E. Eberlein and M.S. Taqqu (eds), *Dependence in Probability and Statistics: A Survey of Recent Results*, pp. 165–192. Boston: Birkhäuser.

Brown, T. and Xia, A. (1995) On Stein–Chein factors for Poisson approximation. *Statist. Probab. Lett.*, **23**, 327–332.

Bryc, W. and Smolenski, W. (1993) Moment conditions for almost sure convergence of weakly correlated random variables. *Proc. Amer. Math. Soc.*, **119**, 355–373.

Cohen, J. (1989) On the compound Poisson limit theorem for high level exceedances. *J. Appl. Probab.*, **26**, 458–465.

Doukhan, P. (1995) *Mixing Properties and Examples*, Lectures Notes in Statist. 85. Berlin: Springer-Verlag.

Doukhan, P. and Louichi, S. (1997) Weak dependence and moment inequalities. Preprint 8, Université de Paris-Sud.

Dziubdziela, W. (1988) A compound Poisson limit theorem for stationary mixing sequences. *Rev. Roumaine Math. Pures Appl.*, **33**(1–2), 39–45.

Dziubdziela, W. (1995) on the limit distribution of the sums of mixing Bernoulli random variables. *Statist. Probab. Lett.*, **23**, 179–182.

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance*. Berlin: Springer-Verlag.

Falk, M., Hüsler, J. and Reiss, R.-D. (1994) *Laws of Small Numbers: Extremes and Rare Events*. DMV Sem. 23. Basel: Birkhäuser-Verlag.

Ferreira, H. (1993) Joint exceedances of high levels under a local dependence condition. *J. Appl. Probab.*, **30**, 112–120.

Guyon, X. (1995) *Random Fields on a Network. Modeling, Statistics and Applications*. New York: Springer-Verlag.

Hudson, W., Tucker, H. and Veeh, J. (1989) Limit distributions of sums of *m*-dependent Bernoulli random variables. *Probab. Theory Related Fields*, **82**, 9–17.

Hüsler, J. (1993) A note on exceedances and rare events of non-stationary sequences. *J. Appl. Probab.*, **30**, 877–888.

Hsiau, S. (1997) Compound Poisson limit theorems for Markov chains. *J. Appl. Probab.*, **34**, 24–34.

Hsing, T., Hüsler, J. and Leadbetter, M. (1988) On the exceedance point process for a stationary sequence. *Probab. Theory Related Fields*, **78**, 97–112.

Kallenberg, O. (1983) *Random Measures*, 3rd. edn. London: Academic Press.

Leadbetter, M. (1991) On a basis for 'peaks over treshold' modeling. *Statist. Probab. Lett.*, **12**, 357–362.

Leadbetter, M. (1995) On high level exceedace modeling and tail inference. *J. Statist. Plann. Inference*, **45**, 247–260.

Leadbetter, M. and Hsing, S. (1990) Limit theorems for strongly mixing stationary random measures. *Stochastic Process. Appl.*, **36**, 231–243.

Leadbetter, M. and Nandagopalan, S. (1989) On exceedance point processes for stationary sequences under mild oscillation restrictions. In J. Hüsler and R.-D. Reiss (eds), *Extreme Value Theory*, Lectures Notes in Statist. 51, pp. 69–80. Berlin: Springer-Verlag.

Leadbetter, M. and Rootzén, H. (1988) Extremal theory for stochastic processes. *Ann. Probab.*, **16**, 431–478.

Leadbetter, M., Lindgren, G. and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes*. Berlin: Springer-Verlag.

Newman, C. (1980) Normal fluctuations and the FKG inequalities. *Commun. Math. Phys.*, **74**, 119–128.

Perera, G. (1997a) Geometry of $\mathbb{Z}^d$ and the central limit theorem for weakly dependent random fields. *J. Theoret. Probab.*, **10**, 581–603.

Perera, G. (1997b) Applications of central limit theorems over asymptotically measurable sets: regression models. *C. R. Acad. Sci. Paris Sér. I*, **324**, 1275–1280.

Rényi, A. (1951) On composed Poisson distribution II. *Acta Math. Acad. Sci. Hungar.*, **2**, 83–98.

Roussas, G. (1994) Asymptotic normality of random fields of positively or negatively associated processes. *J. Multivariate Anal.*, **50**, 152–173.

Smith, R. and Shively, T. (1995) A point process approach to modeling trends in tropospheric ozone based on exceedances of a high treshold. *Atmospheric Environment*, **29**, 3489–3499.

Volkonskii, V. and Rozanov, Y. (1959) Some limit theorems for random functions, I. *Theory Probab. Appl.*, **4**, 178–197.

Volkonskii, V. and Rozanov, Y. (1961) Some limit theorems for random functions, II. *Theory Probab. Appl.*, **6**, 186–198.

Wschebor, M. (1985) *Surfaces Aléatoires: Mesure Géométrique des Ensembles de Niveau*. Lecture Notes in Math. 1147. Berlin: Springer-Verlag.

Received October 1997 and revised November 2002