# The Behavior of Multivariate Maxima of Moving Maxima Processes

Zhengjun Zhang
Department of Mathematics
Washington University
Saint Louis, MO 63130-4899
USA

Richard L. Smith
Department of Statistics
University of North Carolina
Chapel Hill, NC 27599-3260
USA

January 12, 2004

## Abstract

In the characterization of multivariate extreme indices of multivariate stationary processes, multivariate maxima of moving maxima processes, or M4 processes for short, have been introduced by Smith and Weissman. Central to the introduction of M4 processes is that the extreme observations of multivariate stationary processes may be characterized in terms of a limiting max-stable process under quite general conditions and a max-stable process can be arbitrarily closely approximated by a M4 process. In this paper, we derive some additional basic probabilistic properties for a finite class of M4 processes of which each contains finite range clustered moving patterns, called signature patterns, when extreme events occur. We use these properties to construct statistical estimation schemes for model parameters.

**Keywords**: multivariate extremes, extreme value theory, max-stable process, multivariate extremal index.

# 1  Introduction

Univariate extreme value theory studies the limiting distribution of the maxima or the minima of a sequence of random variables. In many circumstances, extreme observations appear to be clustered in time. For example, large price movements in stock market, large insurance claims after a disaster event, heavy rainfalls etc., are lasting a time period. Although the extreme events are short relative to the duration of the whole time series, they do potentially cover several time points, so we need a model for this type of clustering. However, neither univariate nor multivariate extreme value theory is adequate to describe this kind of clustering of extreme events in a time series.

Max-stable processes, introduced by de Haan (1984), are an infinite-dimensional generalization of extreme value theory which does have the potential to describe clustering behavior. The limiting distributions of univariate and multivariate extreme value theory are max-stable, as shown by Leadbetter, Lindgren and Rootzén (1983) in the univariate case and Resnick (1987) in the multivariate case. One of the most important features of max-stable processes is that it does not only model the cross-sectional dependence, but also models the dependence across time. Parametric models for max-stable processes have been considered since the 1980s. Deheuvels (1983) defines the moving minimum process. Pickands' (1981) representation theorem is for multivariate extreme value distribution with unit exponential margins, but is translatable to multivariate extreme value distribution with unit Fréchet margins by the reciprocal transformation $g(x) = 1/x,\ x > 0$. Similarly, a moving minimum process can easily be translated into a moving maximum process, and vice versa. For this reason, a moving minimum process or a moving maximum process can simply be regarded as an MM process. Davis and Resnick (1989) study what they call the max-autoregressive moving average (MARMA) process of a stationary process. For prediction, see also Davis and Resnick (1993).

In the study of characterization and estimation of the multivariate extremal index introduced by Nandagopalan (1990), Smith and Weissman (1996) extend Deheuvels' definition to the so called multivariate maxima of moving maxima (henceforth M4) process. Smith and Weissman (1996) argue that under quite general conditions, the extreme values of a multivariate stationary time series may be characterized in terms of a limiting max-stable process. They also show that a very large class of max-stable processes may be approximated by M4 processes mainly because those processes have the same multivariate extremal indexes as the M4 processes have (Theorem 2.3 in Smith and Weissman 1996).

Since an M4 process involves many parameters, and models cross-sectional dependence and time dependence simultaneously, it is very important to have a good understanding of the behavior and probabilistic properties of M4 processes when we apply M4 processes to real data modeling. The paper aims to provide some insightful understanding for a class of finite M4 processes.

The paper is organized as follows. In Section 2, we re-define M4 process introduced by Smith and Weissman (1996). The main results for deriving an explicit extremal index are quoted. The finite M4 process is defined and some basic distribution properties are illustrated. In Section 3, we study clustered events, especially those large observed values of the process. The meaning of signature patterns is defined and illustrated in an example graphically. Probabilistic properties of those clustered events are proved. Finally, concluding remarks will be given in Section 4.

## 2 The model and some basic distributional properties

The definition of an M4 process is

$$Y_{id} = \max_l \max_k a_{l,k,d} Z_{l,i-k}, \quad d = 1, \ldots, D, \ -\infty < i < \infty, \tag{2.1}$$

where $\{Z_{li}, l \geq 1, -\infty < i < \infty\}$ are an array of independent unit Fréchet random variables with distribution function $F(x) = e^{-1/x}, 0 < x < \infty$. The constants $\{a_{l,k,d}, l \geq 1, -\infty < k < \infty, 1 \leq d \leq D\}$ are nonnegative constants satisfying

$$\sum_{l=1}^{\infty} \sum_{k=-\infty}^{\infty} a_{l,k,d} = 1 \text{ for } d = 1, \ldots, D. \tag{2.2}$$

For this process, we have

$$\begin{aligned}
&\Pr\{Y_{id} \leq y_{id}, 1 \leq i \leq r, 1 \leq d \leq D\} \\
&= \Pr\{Z_{l,i-k} \leq \tfrac{y_{id}}{a_{l,k,d}} \text{ for } l \geq 1, -\infty < k < \infty, 1 \leq i \leq r, 1 \leq d \leq D\} \\
&= \Pr\{Z_{l,m} \leq \min_{1-m \leq k \leq r-m} \min_{1 \leq d \leq D} \tfrac{y_{m+k,d}}{a_{l,k,d}}, l \geq 1, -\infty < m < \infty\} \\
&= \exp\left[ -\sum_{l=1}^{\infty} \sum_{m=-\infty}^{\infty} \max_{1-m \leq k \leq r-m} \max_{1 \leq d \leq D} \tfrac{a_{l,k,d}}{y_{m+k,d}} \right].
\end{aligned} \tag{2.3}$$

Hence

$$\Pr^n\{Y_{id} \leq n y_{id}, 1 \leq i \leq r, 1 \leq d \leq D\} = \Pr\{Y_{id} \leq y_{id}, 1 \leq i \leq r, 1 \leq d \leq D\}.$$

A process satisfying these conditions for all $r \geq 1$ is called max-stable (de Haan 1984). The main result of Smith and Weissman (1996) was to characterize the conditions under which the multivariate extremal index – introduced by Nandagopalan (1990) – for a general stationary process could be approximated by one from an M4 process. First, they proved that under suitable mixing conditions, the multivariate extremal index from a stationary time series could be calculated from that of a max-stable process with the same limiting distributions for any finite-dimensional multivariate extremes. Second, they argued that any max-stable process in $D$ dimensions could be approximated with arbitrary accuracy by one of M4 form. The latter result is a direct generalization to multivariate processes of a result due to Deheuvels (1983), whose own representation is the $D = 1$ case of (2.1). If we further simplify the process by omitting the 'max$_l$' part of the definition, we get the even simpler representation

$$Y_i = \max_k a_k Z_{i-k},$$

with $\{Z_i\}$ independent unit Fréchet, a process known as a 'moving maxima' (MM) process. Thus the M4 process may be considered a generalization of the MM process to a fully multivariate setting.

Since in practice estimating an infinite number of parameters will generally be too ambitious a task, usually we content ourselves with concentrating attention on $l = 1, \ldots, L$ and $-K_1 \leq k \leq K_2$ for some finite numbers $L$, $K_1$ and $K_2$. Here $L$ corresponds to the maximum number of different observed moving patterns when extreme events happen. Those moving patterns are known as signature patterns. We give an example to show what they are after we define our finite M4 processes. The constants $K_1$ and $K_2$ characterize the range of the sequential dependence. A finite dimensional M4 process can be written as follows:

$$Y_{id} = \max_{1 \leq l \leq L} \max_{-K_1 \leq k \leq K_2} a_{l,k,d} Z_{l,i-k}, \quad d = 1, \ldots, D, \tag{2.4}$$

where $\sum_{l=1}^{L} \sum_{k=-K_1}^{K_2} a_{l,k,d} = 1$ for $d = 1, \ldots, D$.

Under model (2.4), it is possible that a big value of $Z_{lk}$ dominates all other $Z$ values within a certain period of length $K_2 + K_1 + 1$ and creates a moving pattern, i.e. $Y_{id} = a_{l,i-k,d} Z_{lk}$ for $i$ close to $k$. A moving pattern is known as a signature pattern. Figure 1 shows two different signature patterns in a simulated M4 process.
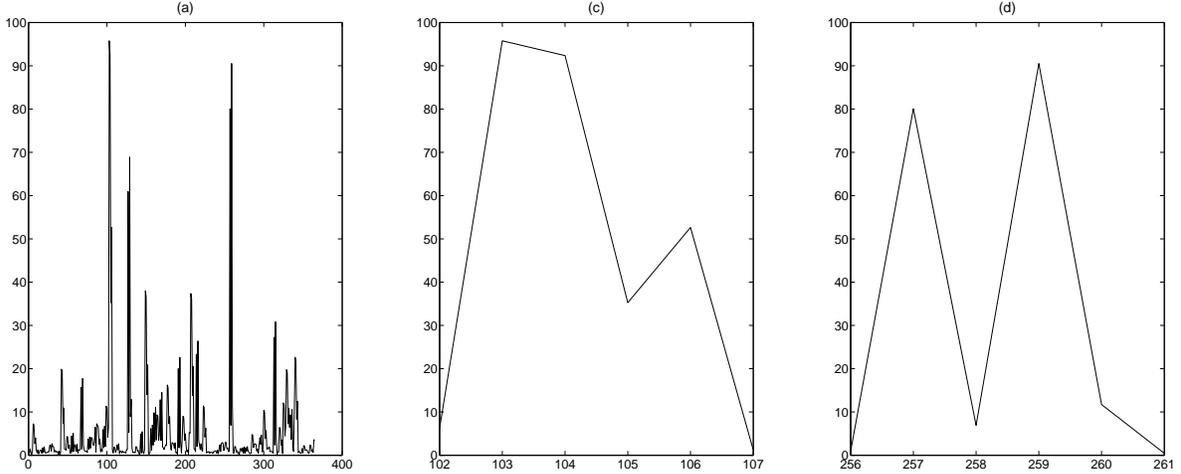


Figure 1: *A demonstration of an M4 process. Plot (a) is a simulated 365 days data set. Plots (b) and (c) are partial pictures drawn from the whole simulated data showing two different moving patterns, called signature patterns, in certain time periods when extremal events occur.*

A single value of $l$ corresponds just to a single signature pattern, but it seems unrealistic in a real data generating process to assume that a single signature pattern would be sufficient to describe the shape of the process every time it exceeds some high threshold. For example, stock market variation resulting from an internal or external big market price movement will last a certain period and form certain moving patterns, such as increasing pattern, decreasing pattern, or up and down pattern, etc. By allowing many values of $l$, we accommodate the feature that a real process may have different types of signature behavior. On the other hand, there are cross-sectional dependencies among different variables under study. For example, a portfolio contains many financial assets, and they are correlated. Motivated by these and other examples in insurance and finance as well as environmental engineering, it would be fair enough to say M4 process is very useful to model extremal cross-sectional dependencies and serial temporal dependencies.

The above analysis suggests that there are strong dependencies and moving patterns among those large observed values during corresponding time periods. Actually, this kind of phenomenon occurs an infinite number of times if our observed sequence is from $i = -\infty$ to $i = \infty$. These phenomena will be shown theoretically and illustrated using examples in Section 3.

Under model (2.4), it is easy to obtain the finite distribution of $\{Y_{id}, \ 1 \leq i \leq r, \ 1 \leq d \leq D\}$. It

3

follows immediately from (2.3) that

$$\Pr(Y_{id} \le y) \;=\; e^{-1/y}, \tag{2.5}$$

$$\Pr(Y_{id} \le y_{id}, Y_{i+1,d} \le y_{i+1,d}) \;=\; \exp\left[-\sum_{l=1}^{L}\sum_{m=-K_1-1}^{K_2} \max\left\{\frac{a_{l,m,d}}{y_{id}}, \frac{a_{l,m+1,d}}{y_{i+1,d}}\right\}\right], \tag{2.6}$$

$$\Pr(Y_{1d} \le y_{1d}, Y_{1d'} \le y_{1d'}) \;=\; \exp\left[-\sum_{l=1}^{L}\sum_{m=-K_1}^{K_2} \max\left\{\frac{a_{l,m,d}}{y_{1d}}, \frac{a_{l,m,d'}}{y_{1d'}}\right\}\right], \tag{2.7}$$

where $a_{l,K_2+1,d}=0,\; a_{l,-K_1-1,d}=0,\; a_{l,K_2+1,d'}=0,\; a_{l,-K_1-1,d'}=0.$

The results show that the marginal distribution function contains no information about the time series dependence of the process, but the bivariate distributions of $\{Y_{id}, Y_{i+1,d},\; 1 \le d \le D\}$ depend on all the parameters of the process and therefore it is reasonable to try to estimate the parameters from these joint distributions.

# 3 Events of clustered observations and their probabilistic properties

## 3.1 The case of $L = 1$ (single pattern)

Consider now a simplified model,

$$Y_{id} = \max_{-K_1 \le k \le K_2} a_{kd} Z_{i-k},\; d = 1, \ldots, D,\; -\infty < i < \infty, \tag{3.1}$$

which is corresponding to the case of $L = 1$ (single pattern).

For any fixed $d$, define the event $A_{d,t}$ at time $t$ by

$$A_{d,t} = \left\{ a_{sd} Z_{t+K_1} = \max_{-K_1 \le k \le K_2} a_{kd} Z_{t+K_1+s-k} \text{ for } s = -K_1, \ldots, K_2 \right\} \tag{3.2}$$

The event of $A_{d,t}$ is equivalent to $Y_{t+K-1+s,d} = a_{sd} Z_{t+K_1}$ for $s = -K_1, \ldots, K_2$. These equations can be put into one equation, i.e.

$$(Y_{td},\; Y_{t+1,d},\; \ldots,\; Y_{t+K_1+K_2,d}) = (a_{-K_1 d},\; a_{-K_1+1,d},\; \ldots,\; a_{K_2 d}) \times Z_{t+K_1}.$$

We can immediately see that the values of $(a_{-K_1 d},\; a_{-K_1+1,d},\; \ldots,\; a_{K_2 d})$ completely determine the values of $(Y_{td},\; Y_{t+1,d},\; \ldots,\; Y_{t+K_1+K_2,d})$ when an extreme event $A_{d,t}$ occurs. The values of $(a_{-K_1 d},\; a_{-K_1+1,d},\; \ldots,\; a_{K_2 d})$ then determine the moving pattern (up or down) of the process. The moving pattern formed by $(a_{-K_1 d},\; a_{-K_1+1,d},\; \ldots,\; a_{K_2 d})$ is called a signature pattern.

It is clear that $\Pr(A_{d,t}) > 0, t \ge 1$. The following probability can be calculated directly though the calculation is not short – see detail in Zhang (2002).

$$\Pr(A_{d,t}) = \frac{1}{[1 + \sum\limits_{j=1}^{K_1+K_2}\left(\min\limits_{1 \le i \le j}\frac{a_{-K_1+i-1,d}}{a_{K_2-j+i,d}} + \min\limits_{1 \le i \le j}\frac{a_{K_2-i+1,d}}{a_{-K_1+j-i,d}}\right)]^2}. \tag{3.3}$$

For $\Pr\{A_{d,t} A_{d,t+m}\}$, it is clear that $\Pr\{A_{d,t} A_{d,t+m}\} = \left[P(A_{d,t})\right]^2$ if $m > K_1 + K_2$.

Suppose $1 \leq m \leq K_1 + K_2$, then from the definition of $A_{d,t}$, we get

$$a_{-K_1+m,d}Z_{t+K_1} \geq a_{-K_1 d}Z_{t+m+K_1}, \tag{3.4}$$

but from the event $A_{d,t+m}$, we get

$$a_{-K_1 d}Z_{t+m+K_1} \geq a_{-K_1+m,d}Z_{t+K_1}. \tag{3.5}$$

The equations (3.4) and (3.5) imply $a_{-K_1+m,d}Z_{t+K_1} = a_{-K_1 d}Z_{t+m+K_1}$ which is true with probability 0, thus

$$\Pr\{A_{d,t}A_{d,t+m}\} = \begin{cases} [P(A_{d,t})]^2 & \text{if } m > K_1 + K_2, \\ 0 & \text{if } 1 \leq m \leq K_1 + K_2. \end{cases} \tag{3.6}$$

Recall that a sequence of events $\{B_n\}$ is said to occur infinitely often (i.o.) or infinitely many times if $\Pr(\limsup_n B_n) = 1$. The mnemonic notation is $\{B_n, i.o.\}$. The following lemma tells that $A_{d,t}$ occurs infinitely often.

**Lemma 3.1** *Under model (3.1), for each $d$ and its entire process we have $\Pr\{A_{d,t}, \quad i.o. \} = 1$, or equivalently*

$$\Pr\left\{Y_{td} = a_{-K_1 d}Z_{t+K_1}, Y_{t+1,d} = a_{-K_1+1,d}Z_{t+K_1}, \ldots, Y_{t+K_2+K_1,d} = a_{K_2 d}Z_{t+K_1}, \ i.o.\right\} = 1,$$

*which implies*

$$\Pr\left\{\frac{Y_{t+m,d}}{Y_{td} + Y_{t+1,d} + \cdots + Y_{t+K_2+K_1,d}} = a_{-K_1+m,d}, \ m = 0, \ldots, K_1 + K_2, \quad i.o. \right\} = 1. \tag{3.7}$$

*Proof.* Let $\{t_1, t_2, \ldots, \}$ be a subsequence of $\{t \geq 1\}$ such that $t_{i+1} - t_i > K_1 + K_2, i \geq 1$, then $\{A_{d,t_i}\}$ is an independent sequence of events. Since $\Pr(A_{d,t_i}) > 0$, by the Borel-Cantelli lemma for independent events, we get

$$\Pr\left\{A_{d,t}, \ i.o.\right\} = \Pr\left\{\bigcap_{t=1}^{\infty}\bigcup_{n=t}^{\infty} A_{d,n}\right\} \geq \Pr\left\{\bigcap_{i=1}^{\infty}\bigcup_{n=i}^{\infty} A_{d,t_n}\right\} = 1,$$

and this completes the proof. □

This lemma shows that there are an infinite number of time periods within which the process is driven by a single extreme jump. For example, a real-world interpretation might be that a flood in a certain region and a certain time period is caused by a specific hurricane. The strengths of different hurricanes are different and the costs are different. Or, we say they follow different patterns.

Lemma 3.1 also shows that the values of all coefficients can be obtained by observing certain proportions among the observed processes.

The following two theorems show how those proportions give the exact value of all coefficients.

**Theorem 3.2** *Under the model in (3.1), for each $d$ and its entire process, if*

$$\Pr\left\{\frac{Y_{t+m,d}}{Y_{td} + Y_{t+1,d} + \cdots + Y_{t+K_2+K_1,d}} = c_{md}, \quad i.o.\right\} = 1$$

*for $m = 0, \ldots, K_1 + K_2$, then $c_{md} = a_{-K_1+m,d}$, and those $\{Y_{td}, \ldots, Y_{t+K_1+K_2,d}\}$ define the events $A_{d,t}$ if $Y_{t+m,d}/(Y_{td} + Y_{t+1,d} + \cdots + Y_{t+K_2+K_1,d}) = c_{md}$ for all $m = 0, \ldots, K_1 + K_2$.*

**Remark 1** *The theorem says, for example when $m = 0$, there is only one constant $c_{0d} = a_{-K_1d}$ such that (3.7) is true. On the other hand, if $Y_{t+m,d}/(Y_{td} + Y_{t+1,d} + \cdots + Y_{t+K_2+K_1,d}) = c_{md}$ is true for one $m$, it is true for all $m$.*

*Proof.* We only prove the case of $m = 0$. The proofs for other values of $m$ are similar.

Define random variables
$$T_{td} = sI_{(Y_{td} = a_{t-s,d}Z_s)},$$
where $I_{(.)}$ is an indicator function. Notice that $t - K_2 \leq s \leq t + K_1$ and $T_{td}$ is uniquely defined for each $t$ and hence for all $t$ because $Z_s$ has an absolutely continuous distribution. The event $A_{d,t}$ gives
$$T_{td} = T_{t+1,d} = \cdots = T_{t+K_1+K_2,d} = t + K_1.$$

Suppose that
$$\frac{Y_{td}}{Y_{td} + Y_{t+1,d} + \cdots + Y_{t+K_2+K_1,d}} = p_d \tag{3.8}$$

occurs infinitely many times for $p_d \neq a_{-K_1d}$, and $p_d \in (0,1)$. Then $T_{td}, T_{t+1,d}, \ldots, T_{t+K_1+K_2,d}$ must follow one of the following two cases.

**(1)** $T_{td} = T_{t+1,d} = \cdots = T_{t+K_1+K_2,d} = t + K$, for some $K \in \{-K_2, -K_2 + 1, \ldots, K_1 - 1\}$,

**(2)** $T_{td}, T_{t+1,d}, \ldots, T_{t+K_1+K_2,d}$ contain at least two different values.

For case (1), we have $Y_{td} = a_{-Kd}Z_{t+K}, Y_{t+1,d} = a_{1-K,d}Z_{t+K}, \ldots, Y_{t+K_1+K_2,d} = a_{K_1+K_2-K,d}Z_{t+K} = 0$ since $K_1 + K_2 - K > K_2$, but $a_{K_2+1,d} = a_{K_2+2,d} = \cdots = 0$. This is a contradiction to $Y_{td} > 0$ for all $t$.

For case (2), define
$$S_1 = \{j : T_{td} = T_{t+j,d}, \ j = 1, \ldots, K_1 + K_2\}, \quad S_2 = \{j : T_{td} \neq T_{t+j,d}, \ j = 1, \ldots, K_1 + K_2\}.$$

Then (3.8) can be written as
$$(1 - p_d)Y_{td} - \sum_{j \in S_1} p_d Y_{t+j,d} = \sum_{j \in S_2} p_d Y_{t+j,d}. \tag{3.9}$$

Then the random variables on both side are independent and continuous. Consequently, the probability of observing (3.9) is zero.

Both cases have shown contradictions when $p_d \neq a_{K_1}$. So $c_{0d} = a_{-K_1d}$ and those $\{Y_{td}, \ldots, Y_{t+K_1+K_2,d}\}$ define events $A_{d,t}$, therefore the proof is completed. $\qquad\square$

**Remark 2** *The equation (3.8) cannot be true even for two different $t$ values – say $t_1$ and $t_2$. Suppose it can be true for $t_1$ and $t_2$. Then we will have the following equation:*

$$\Pr\{\frac{Y_{t_1d}}{Y_{t_1d} + Y_{t_1+1,d} + \cdots + Y_{t_1+K_2+K_1,d}} = \frac{Y_{t_2d}}{Y_{t_2d} + Y_{t_2+1,d} + \cdots + Y_{t_2+K_2+K_1,d}} = p_d\} > 0 \tag{3.10}$$

*which in turn implies (3.8) will occur infinitely many times – a contradiction to the results in Theorem 3.2.*

The following theorem tells that the range cannot be over $K_2 + K_1 + 1$ numbers in order to get infinitely many ratios which are equal to a constant.

**Theorem 3.3** *Under the model (3.1), for each $d$ and its entire process*

$$\Pr\left\{\frac{Y_{td}}{Y_{td} + Y_{t+1,d} + \cdots + Y_{t+K_2+K_1+1,d}} = c_d, \quad i.o.\right\} = 0$$

*for any constant $c_d$.*

*Proof.* Because $Y_{td}$ and $Y_{t+K_1+K_2+1,d}$ cannot be written as functions of just one $Z_t$, the ratio $Y_{td}/(Y_{td} + Y_{t+1,d} + \cdots + Y_{t+K_2+K_1+1,d})$ is a function of at least two different $Z_t$'s. The proof then follows by the same arguments as in Theorem 3.2. □

**Remark 3** *Suppose a data set is generated from an M4 process without knowing the moving range parameters $K_1$ and $K_2$. If too few summands are considered, we will observe more than $L$ dotted lines in the plot of the ratios. The idea is to start from denominator including two summands, three summands, until there are no horizontal dotted lines are observed. The maximal number of summands which give dotted lines are the number of signature patterns.*

Theorems 3.2 and 3.3 and remarks tell that in order to estimate the parameter $a_{-K_1+m}$, we only need to observe two equal values of

$$\frac{Y_{t_1+m,d}}{Y_{t_1 d} + \cdots + Y_{t_1+K_1+K_2,d}}, \ \frac{Y_{t_2+m,d}}{Y_{t_2 d} + \cdots + Y_{t_2+K_1+K_2,d}}$$

at time $t_1$ and $t_2$, where $t_2 - t_1 > K_1 + K_2$, because these two points must be from those dotted lines which are related to parameter values. We now turn to some examples to demonstrate applications of the theorems.

**Example 3.1** . *For the following process*

$$Y_{td} = \max\left(a_{1d}Z_{t-1}, a_{0d}Z_t, a_{-1d}Z_{t+1}\right).$$

*We have $\Pr(A_{d,t}) > 0$ for each $d$ and its entire process. Figure 2 demonstrates the behavior of the process. Notice that the right figure contains an almost dark and solid horizontal line which is formed by those points defined by $A_{d,t}$. The value of $a_{-1,d}$ can be read from the right figure. The interception of the dark dotted line to the vertical axis is the value of $a_{-1,d}$.*

Example 3.1 has shown how to get the moving coefficients in each individual process, but we are mainly interested in multivariate processes. In other words, we need to know how to distinguish different processes. For example, we have two bivariate processes

$$\begin{cases} Y_{i1} = \frac{1}{2}\max(Z_{1,i-1}, \ Z_{1,i}), \\ Y_{i2} = \frac{1}{3}\max(Z_{1,i-1}, \ Z_{1,i}, \ Z_{1,i+1}), \end{cases} \tag{3.11}$$

and

$$\begin{cases} Y'_{i1} = \frac{1}{2}\max(Z_{1,i-1}, \ Z_{1,i}), \\ Y'_{i2} = \frac{1}{3}\max(Z_{1,i}, \ Z_{1,i+1}, \ Z_{1,i+2}). \end{cases} \tag{3.12}$$
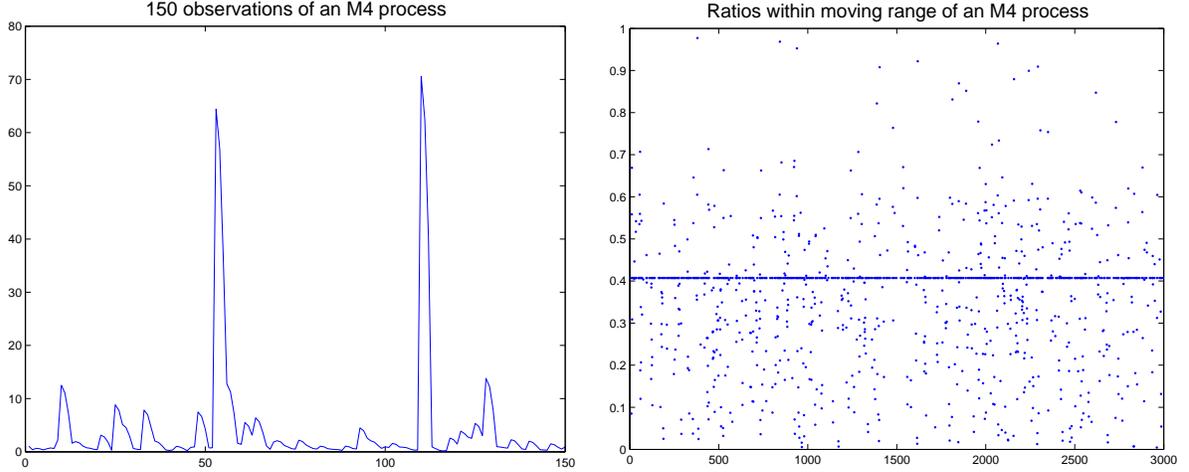
7

Figure 2: *Left figure is a time series plot of 150 observations of process $Y_{td} = \max(a_{1d}Z_{t-1}, a_{0d}Z_t, a_{-1d}Z_{t+1})$ for some $d$. Right figure is a time series plot of ratios $Y_{td}/(Y_{td} + Y_{t+1,d} + Y_{t+2,d})$ of 3000 observations. The value of $a_{-1,d}$ can be read from the right figure. The interception of the dark dotted line to the vertical axis is the value of $a_{-1,d}$.*

Obviously, the joint distribution of $\{Y_{i2}, i = i_1, i_2, \ldots, i_k\}$ is the same as the joint distributions of $\{Y'_{i2}, i = t + i_1, t + i_2, \ldots, t + i_k, \ -\infty < t < \infty\}$. But the joint distributions of the bivariate sequence defined in (3.11) are different from the joint distributions of the bivariate sequence defined in (3.12). For any given data set generated from either (3.11) or (3.12), when we plot ratios for each sequence and observe the dotted lines, we can obtain values $1/2$ and $1/3$. But the plots do not tell that data was generated from which of (3.11) and (3.12). This problem can be solved by comparing the appearing locations of those dots between two horizontal dotted lines since there will be a location shift if the data was not generated from (3.11).

However, if we consider ratio $Y_{i-1,1}/Y_{i,2}$, under model (3.11), this takes the value $3/2$ with positive probability. Under model (3.12), there is no such degeneracy. Similarly, if we consider ratio $Y_{i,1}/Y_{i+1,2}$, under model (3.12), this takes the value $3/2$ with positive probability. Under model (3.11), there is no such degeneracy. So the model can be identified. Some other comparisons also can be done in order to distinguish the models.

## 3.2 The case of $L > 1$ (multiple patterns)

Now consider the case of $L > 1$ and the model which is defined by (2.4). Define

$$A_{ld,t} = \left\{ a_{l,s,d}Z_{l,t+K_1} = \max_{-K_1 \leq k \leq K_2} a_{l,k,d}Z_{l,t+K_1+s-k} \text{ for } s = -K_1, \ldots, K_2 \right\} \tag{3.13}$$

for each $l$ of the model in (2.4).

**Remark 4** *We can define such events for all $l$ simultaneously, but we don't need them here.*

Notice that $\Pr(A_{ld,t}) > 0$ for each $d$ and its entire process, so by Theorem 3.2, for each $m = 0, 1, \ldots, K_1 + K_2$, we have

$$\Pr\left\{ \frac{Y_{t+m,d}}{Y_{td} + Y_{t+1,d} + \cdots + Y_{t+K_2+K_1,d}} = \frac{a_{l,-K_1+m,d}}{a_{l,-K_1,d} + a_{l,-K_1+1,d} + \cdots + a_{l,K_2,d}}, \quad i.o. \right\} = 1. \tag{3.14}$$

8

We expect to have $L$ horizontal dotted lines when plotting $Y_{t+m,d}/(Y_{td} + Y_{t+1,d} + \cdots + Y_{t+K_2+K_1}, d)$. Those lines give estimates of $a_{l,-K_1+m,d}/(a_{l,-K_1,d} + a_{l,-K_1+1,d} + \cdots + a_{l,K_2,d})$, $1 \leq l \leq L$. Figure 3 shows three different signature patterns (points fall onto 3 horizontal lines) which correspond to the case of $L = 3$. As we have already seen, the plots give accurate estimates of the ratios. When $L = 1$,
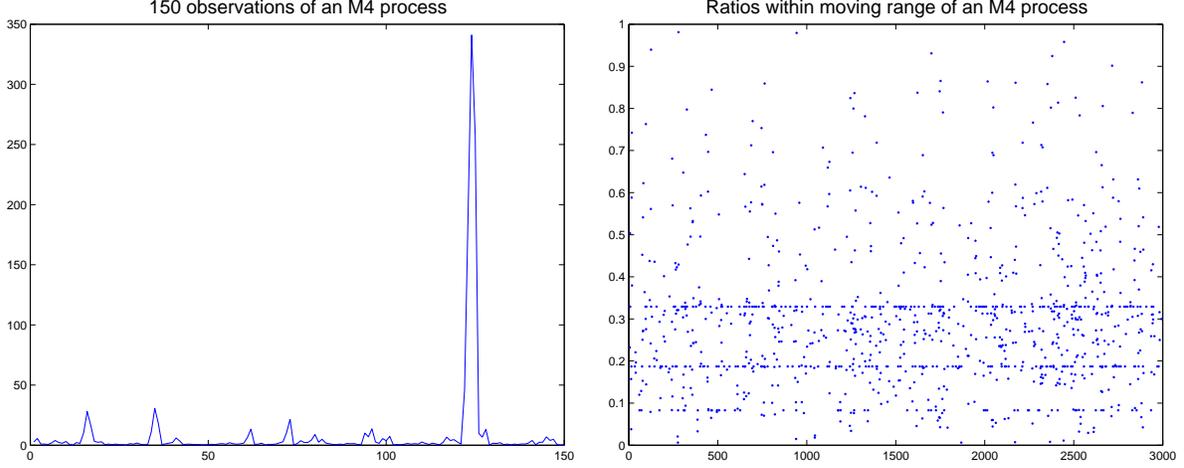


Figure 3: *A demo of multiple signature patterns. Left figure is a time series plot of 150 observations of process $Y_{td} = \max_{1 \leq l \leq 3} \max(a_{l1d}Z_{l,t-1}, a_{l0d}Z_{lt}, a_{l,-1,d}Z_{l,t+1})$ for some $d$. Right figure is a time series plot of ratios $Y_{td}/(Y_{td} + Y_{t+1,d} + Y_{t+2,d})$ of 3000 observations.*

we can exactly get all the values of $a_{kd}$. But for $L > 1$, we cannot. What we get are the ratios $a_{l,-K_1+m,d}/(a_{l,-K_1,d} + a_{l,-K_1+1,d} + \cdots + a_{l,K_2,d})$, not the parameters themselves. Rewrite the model as

$$
\begin{aligned}
Y_{id} &= \max_{1 \leq l \leq L} \max_{-K_1 \leq k \leq K_2} a_{l,k,d} Z_{l,i-k} \\
&= \max_{1 \leq l \leq L} b_{ld} \max_{-K_1 \leq k \leq K_2} c_{l,k,d} Z_{l,i-k}
\end{aligned}
\tag{3.15}
$$

where $b_{ld}$ is the weight of $l$'s signature pattern such that $\sum_l b_{ld} = 1$ for each $d$, and $\sum_k c_{l,k,d} = 1$ for each $l$ and $d$.

Under the model (3.15), all $c_{l,k,d}$ can be estimated by looking into

$$
\Pr\left\{ \frac{Y_{t+m,d}}{Y_{td} + Y_{t+1,d} + \cdots + Y_{t+K_2+K_1,d}} = c_{l,-K_1+m,d}, \quad \text{i.o.} \right\} = 1
\tag{3.16}
$$

for each $d$ and the observed process, therefore we only need to estimate $b_{ld}$.

It is obvious that

$$
\max(a_{l,0,d} Z_{l,i}, a_{l',0,d} Z_{l',i}) \overset{\mathcal{L}}{=} (a_{l,0,d} + a_{l',0,d}) Z_{l,i}, \ l \neq l',
$$

where $\overset{\mathcal{L}}{=}$ means equal in distribution, and

$$
\Pr(a_{l,0,d} Z_{l,i} > a_{l',0,d} Z_{l',i}) = a_{l,0,d}/(a_{l,0,d} + a_{l',0,d}),
$$

and

$$
\Pr(a_{l,0,d} Z_{l,i} > a_{l',0,d} Z_{l',i} | \max(a_{l,0,d} Z_{l,i}, a_{l',0,d} Z_{l',i}) > u) = a_{l,0,d}/(a_{l,0,d} + a_{l',0,d}),
$$

9

where $u$ is any large threshold value. Using these properties to the model (3.15), we can count the frequency of different signature patterns from the observed sequence and get the estimates of $b_{ld}$ and eventually $a_{l,k,d}$.

# 4 Concluding remarks

In this paper, we have studied the behavior of the M4 processes. All the coefficients can be obtained by looking at the ratios of observed values when the data are from an M4 process. Those ratios also tell the numbers of signature patterns and the order of moving range. These results provide guidance to real data modeling. For example, when data nearly follow an M4 process we may be able to apply an M4 process modeling. The coefficients are obtained by looking into the ratios of observed values.

The results of this paper are to some extent pathological, because in practice we would not expect to see repeated identities of the form of (3.14). We view the results as clarifying the role of signature patterns in M4 processes, while serving as a warning against the over-literal interpretation of these processes. For example, if an M4 process was observed with a small amount of added noise, the relationships of the form of (3.14) would no longer hold, but it nevertheless still makes sense to think of an M4 process as a possible approximation to the true multivariate extremal behavior of the process. In a forthcoming paper (Zhang and Smith 2003), we develop practical estimation techniques in more detail.

The use of M4 processes is a new approach to modeling the extremal behavior of multivariate time series. There are clearly many other issues of such M4 processes that need to be studied. For example, it can be used to study the cross-sectional tail dependencies of multiple time series. Zhang (2003) uses (2.4) to construct a characterization and Gamma test of the asymptotic tail dependencies between two random variables. The Gamma test has effectively detected the asymptotic independencies and dependencies of all models using simulated examples.

# References

[1] Davis, R. A. and Resnick, S. I. (1989) Basic Properties and Prediction of Max-ARMA Processes. *Adv. Appl. Prob.* , **21**, 781-803

[2] Davis, R. A. and Resnick, S. I. (1993) Prediction of Stationary Max-Stable Processes. *Ann. Appl. Prob.* , **3**, no 2, 497-525

[3] Deheuvels, P. (1983) Point Processes and Multivariate Extreme Values. *Jour. of Multi. Anal.* **13**, pp. 257-272

[4] de Haan, L. (1984) A Spectral Representation for Max-Stable Processes. *Ann. of Prob.* Vol 12, No. 4, 1194-1204

[5] Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1983) *Extremes and related properties of random sequences and processes.* Springer, Berlin.

[6] Nandagopalan, S. (1990) *Multivariate extremes and the estimation of the extremal index.* Ph.D. dissertation, Dept. Of Statistics, University of North Carolina, Chapel Hill.

[7] Pickands, J. III (1981) Multivariate Extreme Value distributions. *Proc. 43rd Session I.S.I.* (Buenos Aires), 859-878.

[8] Resnick, S. I. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

[9] Smith, R. L., and Weissman, I. (1996) Characterization and Estimation of the Multivariate Extremal Index. Technical Report, Department of Statistics, University of North Carolina.

[10] Zhang, Z., (2002) *Multivariate Extremes, Max-Stable Process Estimation and Dynamic Financial Modeling*, PhD Dissertation, Department of Statistics, University of North Carolina.

[11] Zhang, Z., (2003) Testing Hypothesis of Tail Independence and New Tail Dependence Measures. Manuscript. Submitted.

[12] Zhang, Z. and Smith, R. L. (2003) On the Estimation and Application of Max-Stable Processes. Submitted.