Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Weighted multifractal cross-correlation analysis based on Shannon entropy

Hui Xiong⁎, Pengjian Shang
Department of Mathematics, School of Science, Beijing Jiaotong University, Beijing 100044, PR China

Article info
Article history:
Received 5 February 2015
Revised 16 May 2015
Accepted 27 June 2015
Available online 3 July 2015

Keywords:
Multifractality
Statistical moments
Shannon entropy
Weight
Scaling exponent ratio
Delay

Abstract
In this paper, we propose a modification of multifractal cross-correlation analysis based on statistical moments (MFSMXA) method, called weighted MFSMXA method based on Shannon entropy (W-MFSMXA), to investigate cross-correlations and cross-multifractality between time series. Robustness of this method is verified by numerical experiments with both artificial and stock returns series. Results show that the proposed W-MFSMXA method not only keep the multifractal structure unchanged, but contains more significant information of series compared to the previous MFSMXA method. Furthermore, analytic formulas of the binomial multifractal model are generated for W-MFSMXA. Theoretical analysis and finite-size effect test demonstrate that W-MFSMXA slightly outperforms MFSMXA for relatively shorter series. We further generate the scaling exponent ratio to describe the relation of two methods, whose profile is found approximating a centrosymmetric hyperbola. Cross-multifractality is found in returns series but then destroyed after being shuffled as a consequence of the removed long memory in separate series.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Sequences in complex systems generally exhibit fluctuations on a wide range of time scales of the data, where the distinct fluctuations follow a scaling relation over several orders of magnitude. Such scaling laws provide a description of the data and the complex system by fractal scaling exponents. Nevertheless, many records do not present a simple monofractal scaling behavior and different parts of the series require different scaling exponents. In even more complicated cases, such different scaling behavior can be observed for many interwoven fractal subsets of the time series, where a multitude of scaling exponents is required for a full characterization of the scaling behavior in the same range of time scales and a multifractal analysis must be applied. In order to observe the fractals and multifractals in time series, algorithms have been developed one by another [1,2]. Since detrended fluctuation analysis (DFA) has been proposed by Peng et al. [3] to detect the long-range power-law correlations in DNA sequences, it has been successfully applied to diverse fields [4–11]. Then Podobnik and Stanley [12] generalized DFA and introduced the detrended cross-correlation analysis (DCCA) for two non-stationary time series, which has aroused increasing interest in analysis of long-range cross-correlation and multifractality [13–25]. Specifically, the analysis is based on the bivariate Hurst exponent hxy estimation, which is related to an asymptotic power-law decay of the cross-correlation function. A power-law cross-correlated process has the cross-correlation function \( C_{xy}(k) \propto k^{2h_{xy}-2} \) for \( k \to +\infty \) and the cross-power spectrum \( |f_{xy}(s)| \propto s^{1-2h_{xy}} \) for \( s \to 0^+ \). \( h_{xy} = 0.5 \) is characteristic for the absence of power-law cross-correlation, while processes with

⁎ Corresponding author. Tel.: +8613810435421.
E-mail address: 11271099@bjtu.edu.cn (H. Xiong).
http://dx.doi.org/10.1016/j.cnsns.2015.06.029
1007-5704/© 2015 Elsevier B.V. All rights reserved.
h_{xy} > 0.5 are cross-persistent and h_{yx} < 0.5 indicates the anti-persistent cross-correlation of the data [2,12,25]. Basis of these, the cross-correlation coefficient ($\sigma_{DCCA}$) was introduced with the objective of quantifying the level of cross-correlation between non-stationary time series [26–28]. In the DFA–DCCA–$\sigma_{DCCA}$ line, Kristoufek [29] has recently introduced the DFA framework to estimate regression parameters at different scales and under potential non-stationarity and power-law correlations. Alternatively, the detrending moving-average (DMA) method can also be used for fractal or multifractal analysis, found that the backward multifractal detrending moving average (MFDMA) algorithm outperforms the multifractal detrended fluctuation analysis (MF DFA) [30–33]. As an alternative to $\sigma_{DCCA}$, based on the detrending moving-average cross-correlation analysis (DMCA), the DMCA coefficient $\rho_{DMCA}$ was proposed as well [17,34]. To investigate the multifractal features of two cross-correlated non-stationary series, several methods have been proposed for this purpose. Zhou [14] further proposed multifractal detrended cross-correlation analysis (MFDCCA) as a generation of DCCA. Jiang and Zhou [35] developed a class of MFDCCA algorithms based on DMA, called MFXDMA whose performances are comparative to MFDCCA’s. Moreover, a new method for the detection of long-range cross-correlations and multifractality based on scaling of qth order covariances, multifractal height cross-correlation analysis (MF-HXA), was proposed by Kristoufek [18]. Meanwhile, statistical moments function or partition function has been widely applied to analyze the multifractal features of financial time series [36–40]. Based on this approach, Jiang and Zhou [38] investigated the multifractal characteristics in intraday minutely data for four stock market indexes within individual trading days and found the so-called multifractality is merely an illusion. Dissimilarly, when they performed the partition function approach on the 1-min volatility of two indexes and 1139 stocks in the Chinese stock markets, multifractal nature is significant [39]. Extending that from one dimension to two dimensions, Wang et al. [40] introduced the multifractal cross-correlation analysis based on statistical moments (MFSMXA) as an application of the joint multifractal measures [41], finding that MFSMXA has comparative performance with MFXDMA and sometimes perform slightly better than MFDCCA. By far, multifractal analysis has been employed successfully in various provinces, such as human nature [42–44], financial time series [45–47], river flow [48] and traffic signals [49], etc. However, there are limitations in these algorithms and the common one is that they ignored the difference of the measure in each segment that contributes to the final fluctuation functions unequally theoretically. While in other academic fields, weighting function is selected for improvement, such as Shannon entropy [50–53], variance or energy [54], and so on [55]. Inspired by their work, basis of MFSMXA, we propose in this paper a weighted method based on Shannon entropy ($W$-MFSMXA) to analyze the multifractal cross-correlation between two series. Results show that W-MFSMXA is comparable to and in some extent better than MFSMXA.

The remainder of this paper is organized as follows. In Section 2, the methodologies of MFSMXA and W-MFSMXA are introduced. In Section 3, we test the effectiveness of the W-MFSMXA algorithms with three types of artificial time series: two-component ARFIMA stochastic processes, binomial multifractal model and the NBVP time series. Performances are also compared with these in MFSMXA. Application to financial time series is presented in Section 4. At last, Section 5 gives the conclusions and appendix exhibits charts of the finite-size effect for the binomial multifractal model.

2. Methodologies

2.1. MFSMXA method

The MFSMXA method [40] consists of four steps. Consider two time series $\{X(i)\}$ and $\{Y(i)\}$ of the same length $N$, where $i = 1, 2, \ldots, N$.

Step 1: Divide each series into $N_s = \text{int}(N/s)$ non-overlapping segments of equal length $s$. For a given segment size $s$, the elements in the $v$th segment can be described as:

$$x(i, s) = x((v - 1)s + i) \quad \text{and} \quad y(i, s) = y((v - 1)s + i),$$

where $v = 1, 2, \ldots, N_s$ and $i = 1, 2, \ldots, s$.

Step 2: For each segment, we define a quantity $u$ as follows,

$$u_X = u_X[(v - 1)s + 1, vs] = \sum_i x[(v - 1)s + i]$$

$$u_Y = u_Y[(v - 1)s + 1, vs] = \sum_i y[(v - 1)s + i],$$

where $[(v - 1)s + 1, vs]$ is the $v$th segment. The measure $\mu$ in each segment is constructed as follows,

$$\mu_X(v, s) = u_X(v, s) / \left(\sum_{i=1}^{N_s} u_X(v, s)\right),$$

$$\mu_Y(v, s) = u_Y(v, s) / \left(\sum_{i=1}^{N_s} u_Y(v, s)\right).$$
Step 3: Calculate the statistical moments function $F_X(q, s)$

$$
F_X(q, s) = \sum_{i=1}^{N_i} |\mu_X(v, s) \cdot \mu_Y(v, s)|^{q/2} \equiv N_i \cdot (|\mu_X(v, s) \cdot \mu_Y(v, s)|^{q/2}).
$$

Step 4: Determine scaling behavior of cross statistical moment functions by analyzing log-log plots of $F_X(q, s)$ versus $s$ for each value of $q$. We expect it to scale as

$$
F_X(q, s) \sim s^{\tau_{XY}(q)},
$$

where the exponent $\tau_{XY}(q)$ is a scaling exponent function, regarded as a characteristic function of the fractal behavior. Note that $\tau_{XY}(0) = -1$. For $q = 2$, the method is used to detect the long-range cross-correlations between series, which we call it cross-correlation analysis based on statistical moments (SMXA). When $\{X(i)\} = \{Y(i)\}$, the above method reduces to the multifractal analysis for separate series (MFSMA) [36–39].

In light of the relationship of two sets of multifractal scaling exponents, there is an analytic form between $\tau_{XY}(q)$ and the generalized bivariate Hurst exponent $h_{XY}(q)$:

$$
\tau_{XY}(q) = qh_{XY}(q) - D_f,
$$

where $D_f$ is the fractal dimension of the geometric support of the multifractal measure and $D_f = 1$ for time series analysis [1]. If $q$ dependence of the scaling exponent $\tau(q)$ is nonlinear, there is multifractality in the signal. The singularity strength $\lambda$ and its spectrum $f(\lambda)$ related to $\tau(q)$ through a Legendre transform [56] can be defined as,

$$
\lambda = \tau'(q) \quad \text{and} \quad f(\lambda) = q\lambda - \tau(q).
$$

If the singularity width $\Delta \lambda = \lambda_{max} - \lambda_{min}$ gets close to zero, the series has practically no multifractality. Otherwise, there is a multifractal nature in the series.

According to Ref. [18], for time series in a standard way,

$$
\langle |\mu_X(v, s)|^q \rangle \sim s^{\tau_{XX}(q)+1}, \quad \langle |\mu_Y(v, s)|^q \rangle \sim s^{\tau_{YY}(q)+1},
$$

and

$$
\langle |\mu_X(v, s) \cdot \mu_Y(v, s)|^{q/2} \rangle \sim s^{\tau_{XY}(q)+1}.
$$

Further, the left part of Eq. (9) can be rewritten through the definition of covariance:

$$
\langle |\mu_X(v, s) \cdot \mu_Y(v, s)|^{q/2} \rangle = \langle |\mu_X(v, s)|^{q/2} \rangle \langle |\mu_Y(v, s)|^{q/2} \rangle + \text{cov}(|\mu_X(v, s)|^{q/2}, |\mu_Y(v, s)|^{q/2}),
$$

and combining Eqs. (6) and (8), the first part of the right side of Eq. (10) implies

$$
\langle |\mu_X(v, s)|^{q/2} \rangle \langle |\mu_Y(v, s)|^{q/2} \rangle \sim s^{\tau_{XX}(q)+\tau_{YY}(q)/2},
$$

i.e,

$$
\langle |\mu_X(v, s)|^{q/2} \rangle \langle |\mu_Y(v, s)|^{q/2} \rangle \sim s^{\tau_{XX}(q)+\tau_{YY}(q)/2+1}.
$$

Now consider a new scaling exponent $\eta(q)$ and relationship

$$
\text{cov}(|\mu_X(v, s)|^{q/2}, |\mu_Y(v, s)|^{q/2}) \sim s^{\eta(q)}.
$$

If $\eta(q) = \frac{\tau_{XX}(q) + \tau_{YY}(q)}{2}$ + 1, then

$$
\tau_{XY}(q) = \frac{\tau_{XX}(q) + \tau_{YY}(q)}{2}
$$

holds, which is identical with $\alpha(q) = \frac{h_{XX}(q) + h_{YY}(q)}{2}$ in Ref. [18]. Otherwise, it does not establish unless covariances do not scale with $s$. Thus, the importance of long-range cross-correlations and multifractality resides in the scaling of covariances between $|\mu_X(v, s)|^{q/2}$ and $|\mu_Y(v, s)|^{q/2}$ regarding varying $s$ and $q$. As a consequence, there are two types of cross-multifractality: (i) cross-multifractality caused by multifractality of the separate series, and (ii) cross-multifractality by scaling of covariances between $|\mu_X(v, s)|^{q/2}$ and $|\mu_Y(v, s)|^{q/2}$.
2.2. Weighted MFSMXA method

It can be noted that in the MFSMXA, the product $\mu_X\mu_Y$ of each segment in both sequences contributes to the final function $F_{XY}(q,s)$ equally, but it is rational to make $F_{XY}(q,s)$ proportional to $\mu_X\mu_Y$ in a simplest view. And to a certain extent, the measure $\mu$ is kind of frequency or probability. To improve that, we propose a modification of current MFSMXA, called weighted MFSMXA method based on Shannon entropy (W-MFSMXA), to take into account that not all measurements contribute equally when constructing the statistical moments function and to save as much information as possible carried by the series. Specifically, the motivation behind this is to weight proportionally segments with distinct measurements $\mu$ but the same final $F_{XY}(q,s)$. For example, consider a simple case and let $N_s = 5$ and $q = 2$. where $\mu_1^2 = (0.2, 0.2, 0.2, 0.2, 0.2), \mu_2^2 = (0.1, 0.2, 0.2, 0.2, 0.3)$ and $\mu_3^2 = (0.2, 0.2, 0.2, 0.2, 0.2), \mu_4^2 = (0.1, 0.1, 0.2, 0.3, 0.3).$ Generally we will get $F_{XY}^1 = F_{XY}^2 = 0.2$, but if we consider the effect of weight on the contrary, $F_{XY}^1 = 0.0408$ and $F_{XY}^2 = 0.0417$ are not equal. Weights here are $\omega^1 = (0.1744, 0.2033, 0.2033, 0.2033, 0.2157)$ and $\omega^2 = (0.1773, 0.1773, 0.2067, 0.2193, 0.2193)$ according to the procedure described detailedly in the following. The selection of Shannon entropy [57] to be the weight factor makes sure that $\mu$ with higher value is assigned more weights. Furthermore, Shannon entropy as a complexity (or regularity or orderliness) measure quantifies the information richness of $\mu$ in series and this novel information-based approach has been successfully employed in literary authorship disputes [50], SARS coronavirus [51], biologic signals [52] and financial time series analysis [53].

Therefore, we replace the statistical moments function in step 3 in MFSMXA method with a weighted statistical moments function, which is calculated as follows:

$$ F_{XY}^{(\omega)}(q,s) = \sum_{i=1}^{N_s} \omega(v,s) \cdot |u_X(v,s) \cdot u_Y(v,s)|^{q/2}, $$

(15)

where the weighting function $\omega(v,s)$ is selected to be the sum of Shannon entropy for the $v$th segment of series $\{x(i)\}$ and $\{y(i)\)$ [50–53]:

$$ \omega(v,s) = [-\mu_X(v,s) \cdot \log (\mu_X(v,s)) - \mu_Y(v,s) \cdot \log (\mu_Y(v,s))]/Z, $$

(16)

and $Z$ is a normalization factor such that $\sum_{v=1}^{N_s} \omega(v,s) = 1$. The normalization factor $Z$ in Eq. (16) is given by

$$ Z = \sum_{v=1}^{N_s} [-\mu_X(v,s) \cdot \log (\mu_X(v,s)) - \mu_Y(v,s) \cdot \log (\mu_Y(v,s))]. $$

(17)

Then, same as MFSMXA in the step 4, we expect $F_{XY}^{(\omega)}(q,s)$ to scale as

$$ F_{XY}^{(\omega)}(q,s) \sim s_{XY}^{\omega(q)}. $$

(18)

It reduces to the long-range cross-correlation analysis in a weighted way for $q = 2$ (W-SMXA) and the multifractal analysis for individual series when $\{X(i)\} = \{Y(i)\}$ (W-MFSMXA). Note that $s_{XY}^{\omega(0)}(0) = 0$.

3. Numerical results for artificial time series

In this section, we use three types of artificial time series, the two-exponent ARFIMA processes [58], binomial multifractal model [56] and the NBVP time series [59], to test the effectiveness of this new procedure and make comparison between two methods. The latter two have been proved the existence of cross-multifractality.

3.1. Two-component ARFIMA stochastic processes

The power-law autocorrelations in stochastic variables can be modeled by the fractionally autoregressive integrated moving-average process (ARFIMA) process [60]:

$$ x_t = \sum_{n=1}^{\infty} a_n(d)x_{t-n} + \epsilon_t, $$

(19)

where $\epsilon_t \sim N(0,1), d \in (0,0.5)$ is a memory parameter related to Hurst exponent as $h_X = 0.5 + d$ [12], and $a_n(d)$ are the weights defined by $a_n(d) = d^n(\Gamma(n-d)/\Gamma(1-d)\Gamma(n+1))$, where $\Gamma$ denotes the Gamma function and $n$ is the time scale.

The two-component ARFIMA stochastic processes can be defined by two variables $x_t$ and $y_t$ [58]:

$$ x_t = [WX_t + (1-W)Y_t] + \epsilon_t, $$

$$ y_t = [(1-W)X_t + WY_t] + \nu_t, $$

$$ X_t = \sum_{n=1}^{\infty} a_n(d_1)x_{t-n}, $$

$$ Y_t = \sum_{n=1}^{\infty} a_n(d_2)y_{t-n}. $$

(20)
(a) The power-law dependence of the statistical moments function $F_{xy}^{(o)}(q, s)$ versus the scale $s$ in log-log plots for $q = -4 : 2 : 4$. The dashed lines are the best power-law fits for $F_{xy}^{(o)}(q, s)$. (b) The corresponding scaling exponents $t_{xy}^{(o)}(q)$ versus the moment orders $q$. The straight line is the linear fit for $t_{xy}^{(o)}(q)$. (c) The singularity spectrum $f_{xy}^{(o)}(\lambda)$ versus the singularity strength $\lambda$. (d–f). Processes with $d_1 = 0.1$ and $d_2 = 0.4$. Similar to (a–c), $F_{xy}^{(o)}(q, s)$ (d), $t_{xy}^{(o)}(q)$ (e) and $f_{xy}^{(o)}(\lambda)$ (f) are plotted as well.

Fig. 1. Long-range cross-correlation analysis of the two-component ARFIMA processes in Eq. (20) by W-MFSMXA. $W = 0.8$. (a–c). Processes with $d_1 = d_2 = 0.4$. where $\epsilon_{xt}, \nu_{yt} \sim N(0, 1)$, $a_n(d_1)$ and $a_n(d_2)$ are the weights defined in Eq. (19) through the scaling parameters $d_1$ and $d_2$ ($0 < d_{1,2} < 0.5$), and $W$ is a free parameter controlling the coupling strength between $x_t$ and $y_t$ ($0.5 \leq W \leq 1$). When $W = 1$, the process in Eq. (20) reduces to two decoupled ARFIMA processes defined in Eq. (19), and consequently the long-range cross-correlations between $x_t$ and $y_t$ vanish while both $x_t$ and $y_t$ remain long-range power-law auto-correlated. Additionally, the mixed-correlated ARFIMA process (MC-ARFIMA) has been proposed as a good alternative to the two-component ARFIMA process with analytically solved asymptotic properties [61].

In our simulations, $W = 0.8$ and parameters $(d_1, d_2)$ are set as $d_1 = d_2 = 0.4$ and $d_1 = 0.1$, $d_2 = 0.4$ separately. Series pre-processed with $x_t = \frac{x_{t \max} - \min(x_t)}{\max(x_t) - \min(x_t)}$ are then investigated by the proposed W-MFSMXA. Figs. 1a–c show the numerical results of the process in Eq. (20) with $d_1 = d_2 = 0.4$. Fig. 1a illustrates the dependence of the statistical moments function $F_{xy}^{(o)}(q, s)$ regarding the scale $s$ in log-log plots for $q = -4:2:4$. Excellent power-law cross-correlations are observed. Figs. 1b and c show the corresponding scaling exponents $t_{xy}^{(o)}(q)$ and singularity spectrum $f_{xy}^{(o)}(\lambda)$, respectively. The linear dependence of $t_{xy}^{(o)}(q)$ with respect to $q$ as well as the singularity widths $\Delta \lambda^{(o)} = 0.0328$ closer to 0 unveils the fractal property in two-component ARFIMA stochastic processes. Similar results can also be obtained for the process in Eq. (20) with $d_1 = 0.1$ and $d_2 = 0.4$ in Figs. 1d–f. The singularity width $\Delta \lambda^{(o)}$ under this case is 0.00765.

3.2. Binomial multifractal series

Using the binomial multifractal model (BMF) [56], we construct several binomial multifractal series $\{x(i)\}$ with the length of $N = 2^{n_{\text{max}}}$. The kth number in series is defined as

$$x_k = p_x^{n(k-1)}(1 - p_x)^{n_{\text{max}} - n(k-1)}, \tag{21}$$

where $0 < p_x < 0.5$ is a parameter, $n(k)$ is the number of digits equal to 1 in the binary representation of index $k$. For example, $n(14) = 3$, since 14 corresponds to binary 1110. Series generated by this model has known analytic multifractal properties. The scaling exponent $\tau(q)$ has an analytic form [62]:

$$\tau_{x}(q) = -\frac{\ln[p_x^q + (1 - p_x)^q]}{\ln 2}, \tag{22}$$

satisfying Eq. (14) [35,40].

In this paper, we generated series with length of $2^{16}$ with $p_x = 0.2$ for $\{x(i)\}$, $p_y = 0.3$ for $\{y(i)\}$ and $p_z = 0.4$ for $\{z(i)\}$, just same with those in Ref. [40]. The two pairs of time series, $\{x(i)\} \& \{y(i)\}$ and $\{y(i)\} \& \{z(i)\}$, are highly correlated with a coefficient of
with respect to weighted did not change the structure of the statistical moments function, which is particularly evident in Fig. 2c. It is noteworthy that $\tau(\omega)$ more, the slopes of $\tau(\omega)$, than zero and differences of $\Delta \lambda$ correspondingsingularitywidths $\Delta \lambda$ closer to each other. When shifting the figure of $\tau(q)$ against the moment order $q$ in Fig. 3a. Obviously, $\tau(q)$ are nonlinear and upwards convex, which indicates the multifractal nature between two time series. What's more, $\tau(q)$ and $\tau(q)$ are almost parallel to each other. When shifting the figure of $\tau(q)$ towards zero in W-MFSMXA, it is noteworthy that $\tau(q)$ in W-MFSMXA is more synchronous with $\tau(q)$ in MFSMXA apparently when $q$ is non-negative and especially when $q$ equals zero in W-MFSMXA while one in MFSMXA.

Analogous phenomena turn up when plotting the scaling exponent function $\tau(q)$ versus the moment order $q$ in Figs. 2a and 3a. Obviously, $\tau(q)$ are nonlinear and upwards convex, which indicates the multifractal nature between two time series. What's more, $\tau(q)$ and $\tau(q)$ are almost parallel to each other. When shifting the figure of $\tau(q)$ towards unit 1, $\tau(q)$ and $\tau(q)$ coincide pretty well (not shown here), which is identical with $\tau(q)$ closer to $F(x)$ or $F(y)$ with the step-length of 1. Similar results can be obtained for series $\{x(i)\}$ and $\{z(i)\}$ for series $\{y(i)\}$ and $\{z(i)\}$, which are obvious in Figs. 2 and 3a. Table 1 shows the corresponding singularity widths $\Delta \lambda$ captured before and after the weighting procedure. For two algorithms, $\Delta \lambda$ are much larger than zero and differences of $\Delta \lambda$ deviate insignificantly from zero. Thus, we assume that except for the phase, the multifractal structure of the series will not be changed after being weighted, and it's reasonable to believe there is a delay in MFSMXA compared to W-MFSMXA.

The power-law dependence of $F_{xy}(q)$ and $F_{yy}(q)$ versus $q$ for separate series $\{x(i)\}$ and $\{y(i)\}$ based on weighted statistical moments are also displayed in Fig. 3b, which demonstrates that the slope of weighted cross statistical moments function $F_{xy}(q)$ is the average slopes of $F_{xy}(q)$ and $F_{yy}(q)$, satisfying Eq. (14). Same results can also be obtained when computing the slopes of $F_{xy}(q)$ and $F_{yy}(q)$. As the scaling exponent function for $q$ in W-MFSMXA is synchronous with that in MFSMXA for $q + 1$, then we assume $\tau(q)$ against $\tau(q)$ and theoretically when $\{x(i)\} = \{y(i)\}$

\[
\begin{align*}
\tau(x) = & \tau(x) + 1 = -\frac{\ln[p_x^{q+1} + (1 - p_x)^{q+1}]}{\ln 2}, \\
\tau(y) = & \tau(y) + 1 = -\frac{\ln[p_y^{q+1} + (1 - p_y)^{q+1}]}{\ln 2}, \\
\lambda_{xy}(q) = & -p_{xy}^{q+1}\ln p_x + (1 - p_x)^{q+1}\ln(1 - p_x) \\
& + \frac{\ln[p_x^{q+1} + (1 - p_x)^{q+1}]}{\ln 2}, \\
F_{xy}(q) = & -p_{xy}^{q+1}\ln p_x + (1 - p_x)^{q+1}\ln(1 - p_x) \\
& + \frac{\ln[p_x^{q+1} + (1 - p_x)^{q+1}]}{\ln 2}.
\end{align*}
\]
Fig. 3. Multifractal analysis of BMF series when $q$ is from $-5$ to 5 with the step-length of 0.2. (a) Scaling exponents $\tau(q)$ in MFSMXA (filled circles) and W-MFSMXA (open squares). Different colors stand for different types of scaling exponents: $\tau_{xy}(q)$ (blue), $\tau_{xz}(q)$ (black) and $\tau_{yz}(q)$ (red). Theoretical values are calculated by Eq. (22). (b) Scaling exponents $\tau(q)$ obtained for separate series in W-MFSMA. To further verify the Eq. (14), $\tau_{xy}(q)$ and the average of $\tau_{xx}(q)$ and $\tau_{yy}(q)$ are also obtained. (c) The singularity spectra $f(\lambda)$ for the separate series in W-MFSMA. The solid lines are the corresponding theoretical values in Eq. (23). (d) The singularity spectra $f(\lambda)$ between two series in W-MFSMXA. The solid lines are their average values calculated according to Eq. (23). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Table 1

| $\Delta \lambda$ | MFSMXA | W-MFSMXA | Diff. |
|-----------------|--------|----------|-------|
| BMF $\Delta \lambda_{xy}$ | 1.5708 | 1.5613 | 0.0095 |
| $\Delta \lambda_{xz}$ | 1.2401 | 1.2405 | 0.0004 |
| $\Delta \lambda_{yz}$ | 0.8146 | 0.8134 | 0.0012 |
Note the difference of the q index is 1 rather than the step-length. The singularity spectrum \( f(\lambda) \) of the theoretical analysis in Eq. (23) is displayed in Fig. 3c. Numerical results are significantly in line with their theoretical values. Fig. 3d shows the singularity spectra between two series and they basically accord with the averages of theoretical values on the basis of Eq. (23).

It is worthwhile to note that short length \( N \) of the investigated series or small scale \( s \) often results in a spurious detection of multifractal behavior from a monofractal model [19]. To evaluate the finite-size effect [25] of the BMF model, performances of the scaling exponent \( \tau(q) \) estimators in MFSM(X)A and the proposed W-MFSM(X)A are studied. In all cases, time series with length \( N = 2^{n_{\text{max}}} \) are analyzed, where \( n_{\text{max}} = \{10, 11, 12, 13, 14, 15, 16\} \). The scales \( s \) are given by \( s = \text{round}([\log_{10}(T_{\text{min}}), \log_{10}(T_{\text{max}}), 30]) \), where \( T_{\text{min}} = 100, T_{\text{max}} = N/4 \) and \( \text{round}(X) \) means rounding the elements of \( X \) to the nearest integers. For series with different parameters \( p_x = \{0.1, 0.2, 0.25, 0.3, 0.4\} \) in Eq. (21), results of the evaluation of (cross-) multifractality are presented in appendix. Significantly, for analyses of separate series in MFSMA and W-MFSMA (see Figs. 9 and 10), two \( \tau(q) \) estimators, namely \( \tau_{xx}(q) \) and \( \tau_{xx}^{(0)}(q) \), have the same performances: (1) For sequences with different length sizes, numerical results of \( \tau_{xx}(q) \) and \( \tau_{xx}^{(0)}(q) \) coincide well with the corresponding theoretical values respectively except for \( n_{\text{max}} = 10 \); (2) For \( p_x = 0.4 \) in Figs. 9e and 10e, all series perform well especially in W-MFSMA; (3) Generally, \( n_{\text{max}} = 11 - 13 \) are acceptable while \( n_{\text{max}} = 14 - 16 \) are more reliable, which is accordant with the results of MFSMXA procedure between two time series in Fig. 11. Results of \( \tau_{xy}(q) \) between two time series in W-MFSMXA are plotted in Fig. 12. Yet again, nine pairs of data used here for cross-multifractal analysis are significantly correlated (see Table 2). It is clear to see that \( \tau_{xy}(q) \) deviate significantly from the theoretical values when \( n_{\text{max}} = 10 \), same as in MFSM(X)A and W-MFSMA. In most of cases, performances of \( \tau_{xy}(q) \) when \( n_{\text{max}} = 11 - 16 \) are comparative. While when \( p_x = 0.1, p_y = 0.25 \) and \( p_x = 0.1, p_y = 0.3 \) (see Figs. 12b and c), \( n_{\text{max}} = 11 - 13 \) are the best. In general, on the basis of investigations in this section, performances of W-MFSMA are as good as MFSMA’s and \( n_{\text{max}} = 11 - 16 \) are suggested for use in MFSM(X)A or W-MFSM(X)A. Nevertheless, W-MFSMXA outperforms MFSMXA for series with relatively shorter length, i.e., \( n_{\text{max}} = 11 - 13 \). At this point, W-MFSM(X)A is comparable to MFSM(X)A. In order not to seem rigmarole, we only illustrate the results of one pair of experimental data in following experiments.

In fact, from the log-log plot, the function \( \tau_{xy}^{(0)}(q) \) in Eq. (18) can be defined as follows [44,63,64]:

\[
\tau_{xy}^{(0)}(q) = \lim_{s \to 0} \frac{\log F_{xy}^{(0)}(q, s)}{\log s},
\]

so can be the Eq. (5). Combining Eqs. (5), (18) and (24), we can get a new equation:

\[
\frac{\tau_{xy}^{(0)}(q)}{\tau_{xy}(q)} = \lim_{s \to 0} \frac{\log F_{xy}^{(0)}(q, s)}{\log F_{xy}(q, s)} = a(q),
\]

where \( a(q) \) is the exponent of the best power law fit for \( F_{xy}^{(0)}(q, s) \) versus \( F_{xy}(q, s) \) in logarithms, which can be regarded as a scaling exponent ratio related to the moment order \( q \) and used to explain why there is a delay. As for positive time series \( \{x(i)\} \) and \( \{y(i)\} \), the values of \( F_{xx}(1, s) \) and \( F_{yy}(1, s) \) are equivalent to the sum of the entire series individually, which is uncorrelated with the scale \( s \). Therefore, according to Eq. (14), \( \tau_{xx}(1) = 0 \) and \( a(1) \) diverges. Notice \( a(0) = 0 \).

The \( a(q) \) with respect to \( q \) for series \( \{x(i)\} \) and \( \{y(i)\} \) is depicted in Fig. 4a. In Fig. 3a, \( \tau_{xy}^{(0)}(q) \) and \( \tau_{xy}(q) \) increase with the increasing orders \( q \) and \( \tau_{xy}^{(0)}(q) > \tau_{xy}(q) \) for all values of \( q \). Thus, for \(-5 < q < 0 \), \( \tau_{xy}(q) < \tau_{xy}^{(0)}(q) < 0 \) then \( 0 < a(q) < 1 \); for \( 1 < q < 5 \), \( 0 < \tau_{xy}(q) < \tau_{xy}^{(0)}(q) \) then \( a(q) > 1 \); otherwise, \( \tau_{xy}^{(0)}(q) > 0 \) and \( -1 < \tau_{xy}(q) < 0 \), then \( a(q) < 0 \). Since the \( a(q) \) functions diverges at \( q = 1 \), \( a(1) \) is removed for better illustrations as well in the following experiments. Then, the graph of \( a(q) \) with
Fig. 5. Multifractal cross-correlation analysis of the NBVP time series. (a) Statistical moments functions $F_{xy}(q, s)$ in MFSMXA and W-MFSMXA with respect to $s$ are shown in log-log plots for $q$ from $-3$ to $3$ (from top to bottom) with the step-length of $1$. (b) Scaling exponents $\tau_{xy}(q)$ are calculated by the proposed W-MFSMXA and MFSMXA. The theoretical values according to Eq. (25) are also performed. (c) Scaling exponent ratio $a(q)$ as a function of $q$. (d) Scaling of covariances between $|\mu_X(v, s)|^{q/2}$ and $|\mu_Y(v, s)|^{q/2}$ by MFSMXA according to Eq. (13). Scaling of covariances here is plotted as $\eta(q) - 1$ for better comparison. $\tau_{xx}(q), \tau_{yy}(q)$ and their average, and $\tau_{xy}(q)$ are also shown.

respect to $q$ can be approximately viewed as a centrosymmetric hyperbola. For $q \rightarrow \pm 5$, $a(q)$ approximates to 1, demonstrating two methods are practically consistent, while differences are apparent for relatively smaller orders $q$. Actually, for higher orders $q > 5$ or $q < -5$, $a(q) \rightarrow 1$ establishes as well. However, $a(q) \neq 1$ holds for all $q$, consequently delay between W-MFSMXA and MFSMXA is substantial. As we can see in Fig. 4b, the profiles of $\tau_{xy}^{(\omega)}(q)$ and $\tau_{xy}(q)$ after multiplying corresponding $a(q)$ are coincident, confirming the Eq. (25).

3.3. NBVP time series

Deriving from the classical continuous BVP oscillator, a two-dimensional map by using the forward Euler discrete scheme can be written as:
where \(0 < \rho \ll 1, 0 < a < 1, 0 < b < 1\) is a stimulus intensity and \(0 < \delta < 1\) is the step size. Given a set of variables and initial values \(x_0\) and \(y_0\), we can get time series \(\{x(i)\}\) and \(\{y(i)\}\).

First, we generate two more complex BVP series with the length of \(N = 10^4\), which are called NBVP time series, whose initial value \(\mu\) is a random variable, satisfying \(\mu \sim \text{Norm}(0.2, 1)\) and other initial values are constant values, with \(\delta = 0.1, \rho = 0.001, b = 0.5, a = 0.2, x_0 = 0\) and \(y_0 = 0.01\), where the cross-correlation relationship is found to be multifractal by Xue et al. [59]. To our knowledge, no analytic expression has been generated for this model. Series analyzed here are with a given preparation as in Section 3.1, then we apply MFSMXA and W-MFSMXA algorithms to the NBVP series to examine the validity of the assumption about delay in Section 3.2. Results are shown in Fig. 5. Singularity widths in this case are \(\tau(\omega)\) and \(\tau(\lambda)\) measure in both methods. The fractal structure of the series remains unchanged. \(\tau(q)\) measurements in both methods are apparently different for relatively smaller \(q\) but more accordant for higher \(q\) while never equal. Proposition about the delay is further verified. It can be seen in Fig. 5d, the scaling of covariances \(\eta(q) \sim \delta \rho\) is higher than the average of scaling of separate processes for negative moment orders \(q\), the deviations are nevertheless not significant for higher positive \(q\). In contrast, \(\eta(q) \sim \delta \rho\) and \(\tau(q)\) both basically coincide with \(\tau(q)\), implying that the cross-multifractality of series is mainly caused by the scaling of covariances between \(|\mu_X(v,s)|^{1/2}\) and \(|\mu_Y(v,s)|^{1/2}\) as the construction of the NBVP mixes the long memory of the separate processes.

4. Numerical results for financial time series

In order to prove the applicability of the proposed W-MFSMXA method in real-world data, we then apply this method to financial time series to analyze their cross-multifractality. Here, we perform two US stock indices: the Dow Jones Industrial Average (DJI) and the National Association of Securities Dealers Automated Quotations (NAS). The original data were obtained from yahoo finance (http://finance.yahoo.com) from Jan. 2, 1997 to Dec. 29, 2014.

To make sure of the synchronization of the time series, we exclude the asynchronous data and then reconnect the remaining parts of the original series. In so doing, the resulting length is 4526. The data we analyze in this paper are defined as the absolute logarithmic returns, this is, \(r_t = |\log(x_t) - \log(x_{t-1})|\), where \(x_t\) denote the closing price for a stock index on day \(t\). The daily absolute returns for the two stock indices are shown in Fig. 6. Profiles of two series are very similar to each other.

Let \(x\) denotes DJI series and \(y\) denotes NAS throughout this section. Fig. 7a depicts the statistical moments function \(F_{xy}(q, s)\) as a function of scale \(s\) in log-log plots for \(q\) from \(-3\) to \(3\) with the step-length of \(1\) by W-MFSMXA and MFSMXA methods. Excellent power laws are also observed for both algorithms. Same as before, \(F_{xy}(q, s)\) in W-MFSMXA is more similar to \(F_{xy}(q + 1, s)\) in MFSMXA, which is conform for both algorithms. The scaling exponent ratio \(a(q)\) is plotted in Fig. 7c and clearly approximates a hyperbola. The nonlinearity of \(r(q)\) demonstrates the
Fig. 7. Multifractal cross-correlation analysis for the absolute logarithmic daily returns series. (a) Statistical moments functions in MFSMXA and W-MFSMXA with respect to \( s \). (b) The scaling exponents \( \tau_{xy}(q) \) are calculated by the proposed W-MFSMXA and MFSMXA. Theoretical values in light of Eq. (25) are also obtained. (c) Scaling exponent ratio \( a(q) \) as a function of \( q \). \( a(1) = 332.4224 \) is removed for its divergence. (d) Scaling of covariances between \( |\mu_X(v,s)|^{q/2} \) and \( |\mu_Y(v,s)|^{q/2} \) by MFSMXA in Eq. (13). \( \tau_{xx}(q) \), \( \tau_{yy}(q) \) and the average, and \( \tau_{xy}(q) \) are also depicted.

Fig. 8. Multifractal cross-correlation analysis for the original and shuffled series in W-MFSXMA. \( \tau(q) \) and \( f(\lambda) \) for the shuffled series are accordingly calculated by the average of results of 1000 surrogate data. (a) Scaling exponents \( \tau_{xy}(q) \). (b) Singularity spectra \( f(\lambda) \). \( \Delta \lambda = 0.3764 \). (c) Singularity widths \( \Delta \lambda_{\text{rnd}} \) of the shuffled data. \( \Delta \lambda_{\text{rnd}} = 0.0188 \pm 0.0038 \).
existence of but weak multifractality in the absolute logarithmic daily returns series. Besides, in Fig. 7d, \( \tau_{xy}(q) \) are practically consistent with the averages of \( \tau_{xx}(q) \) and \( \tau_{yy}(q) \) while the scaling of covariances \( \eta(q) - 1 \) is obviously the lowest, indicating that the cross-multifractality in series is mainly dominated by the multifractality of separate series.

To give a more systemic statistical test on these returns series, we adopt the following bootstrapping approach to get the possibility that the empirical multifractality could be spurious [38,39]. We reshuffle each series 1000 times to remove any potential temporal correlation, then carry out the W-MFSMXA procedure for the surrogate data as for the original data. A very strict null hypothesis is imposed to investigate whether the singularity spectrum \( f(\lambda) \) is wider than those produced by chance. The null hypothesis is the following:

\[
H_0 : \Delta \lambda \leq \Delta \lambda_{\text{rnd}}.
\]

The associated probability of false alarm for multifractality is defined by

\[
p = \frac{\# \{ \Delta \lambda \leq \Delta \lambda_{\text{rnd}} \}}{n},
\]

where \( n \) is the number of shufflings and \( \# \{ \Delta \lambda \leq \Delta \lambda_{\text{rnd}} \} \) counts the number of \( \Delta \lambda \) whose value is not greater than \( \Delta \lambda_{\text{rnd}} \). As \( n \to \infty \), it is clear that the estimated bootstrap \( p \)-value will tend to the ideal bootstrap \( p \)-value. Under the conventional significance level of 0.01, the multifractal phenomenon is statistically significant if and only if \( p \leq 0.01 \). When \( p > 0.01 \), the null hypothesis cannot be rejected.

Results of the statistical test are presented in Fig. 8. A remarkable deviation of the singularity spectrum \( f(\lambda) \) of the original data from \( f_{\text{rnd}}(\lambda_{\text{rnd}}) \) of the surrogate data is clear to see in Fig. 8b. Hence, we can infer that cross-multifractality in series is destroyed by the shuffling procedure as a consequence of the eliminated long memory in the separate returns series. Further, we find \( p = 0 \) using the significance level of 0.01, which is a significant evidence for the presence of cross-multifractality in the investigated DJI and NAS data. Comparing the scaling exponent ratio \( a(q) \) in Figs. 4a, 5c and 7c, it can be found the similar hyperbolic feature of \( a(q) \). Moreover, the statistical moments functions as well as the scaling exponents in W-MFSMXA are almost exactly the same as those in MFSMXA if taking no account of their phases or positions.

5. Conclusions

In this paper, we have proposed a modified MFSMXA method called weighted MFSMXA method based on Shannon entropy (W-MFSMXA). To test the robustness of this new procedure and compare to the heuristic MFSMXA, we first employed three types of artificial time series: the two-exponent ARFIMA processes with monofractal feature, the binomial multifractal model (BMF) and the NBVP time series, which have been proved the existence of multifractal properties. However, unlike BMF, cross-multifractality in NBVP is mainly caused by the scaling of covariances rather than the correlated separate processes.

Long-range cross-correlation analysis of the two-component ARFIMA stochastic processes by W-MFSMXA uncovered the fractal nature in these power-law cross-correlated series. Results of multifractal analyses show that the statistical moments function versus scale \( s \) for \( q \) in W-MFSMXA is more similar to that in MFSMXA for \( q + 1 \). The scaling exponents obtained by two methods are almost parallel to each other. What’s more, differences of the corresponding singularity widths captured by two algorithms do not deviate significantly from zero. Thus, it is rational to believe that the multifractal structure of series has not been changed by the weighting procedure without considering their positions. Basis of these, we generated analytic formulas of the BMF model for W-MFSMXA. Theoretical analysis and finite-size effect test verified the good performances of W-MFSMXA. Specifically, numerical results are well in line with the theoretical values and W-MFSMXA outperforms MFSMXA for analysis of series with relatively shorter length. Meanwhile, we generated the scaling exponent ratio \( a(q) \), whose profile is approximately a centrosymmetric hyperbola, to describe the relation between W-MFSMXA and MFSMXA. Scaling exponent estimators in both methods are more consistent for relatively higher orders \( q \) while never equal, indicating the delay in MFSMXA compared to W-MFSMXA.

At last, we applied W-MFSMXA to the absolute logarithmic returns of two US stock market indices: the Dow Jones Industrial Average (DJI) and the National Association of Securities Dealers Automated Quotations (NAS). The results obtained are same with those obtained for artificial time series. Besides, weak cross-multifractal nature is found in the returns series and their cross-multifractality is mainly due to the multifractality of separate series. Further, we compared outcomes of the original and shuffled series. Statistical test proved that cross-multifractality in the investigated DJI and NAS data is destroyed by the shuffling procedure as a consequence of the removed long memory in separate returns series.

In general, all these illustrations make we believe that the proposed W-MFSMXA method not only will not change the multifractal structure of series but contains more significant information of series compared to the previous MFSMXA method for the introduced weighting function. In some extent, it’s comparable to and better than MFSMXA. Though there is no criterion for Shannon entropy as the weight factor as the best choice, it did exhibit elegant performances as we can see in the text. This goes nevertheless beyond the scope of current work.

Acknowledgment

The financial support by China National Science (61371130, 61304145) and Beijing National Science (4122059) are gratefully acknowledged.
Appendix A. Finite-size effect

To evaluate the finite-size effect of BMF series, plots of the scaling exponent $\tau(q)$ in MFSM(X)A and the proposed W-MFSM(X)A are exhibited in this appendix. Series investigated here are generated from the process in

Fig. 9. The scaling exponents $\tau(q)$ for separate series in MFSMA. Theoretical values are calculated by Eq. (22) for different parameters $p_x$.

Fig. 10. The scaling exponents $\tau(q)$ for separate series in W-MFSMA. Corresponding theoretical values are calculated according to Eq. 23 for different parameters $p_x$. 
Fig. 11. The scaling exponents $\tau(q)$ between two time series in MFSMXA. Theoretical values are calculated in light of Eqs. (14) and (22).

Table 2  
Correlation coefficients of two series generated by the process in Eq. (21) with varied parameters.

| Corrcoef | $p_x$ | $p_y$ | $p_x$ | $p_y$ | $p_x$ | $p_y$ | $p_x$ | $p_y$ |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|
|          | 0.1   | 0.2   | 0.25  | 0.3   | 0.4   |       |       |       |
| $p_x$    |       |       |       |       |       |       |       |       |
| 0.1      | 1     | 0.8930| 0.7606| 0.5980| –     |       |       |       |
| 0.2      | 1     | 0.9662| 0.8676| 0.5667|       |       |       |       |
| 0.25     | 1     | 0.9642| 0.7257|       |       |       |       |       |
| 0.3      | 1     | 0.8689|       |       |       |       |       |       |
| 0.4      | 1     |       |       |       |       |       |       |       |

Eq. (21) with different parameters $p_x = \{0.1, 0.2, 0.25, 0.3, 0.4\}$. Time series with $N = 2^{n_{max}}$ are analyzed here, where $n_{max} = \{10, 11, 12, 13, 14, 15, 16\}$. In each figure, straight lines in different colors denote $\tau(q)$ of series with different length sizes and hidden legends in other graphs are same as the first graph’s. Notice the different intervals for the vertical axes.
Fig. 12. The scaling exponents $\tau(q)$ between two series in W-MFSMX. Theoretical values are calculated based on Eqs. (14) and (23).

References

[1] Kantelhardt JW, Zschiegner SA, Koscielny-Bunde E, Havlin S, Bunde A, Stanley HE. Multifractal detrended fluctuation analysis of nonstationary time series. Physica A 2002;316:87–114.
[2] Kantelhardt J. Fractal and multifractal time series. In: Meyers RA, editor. Encyclopedia of Complexity and Systems Science. New York: Springer; 2009. p. 3754–79.
[3] Peng CK, Buldyrev S, Simons M, Stanley H, Goldberger A. Mosaic organization of DNA nucleotides. Phys Rev E 1994;49:1685–9.
[4] Ossadnik SM, Buldyrev SV, Goldberger AL, Havlin S, Mantegna RN, Peng CK. Correlation approach to identify coding regions in DNA sequences. Biophys J 1994;67:64–70.
[5] Buldyrev SV, Goldberger AL, Havlin S, Mantegna RN, Matsa ME, Peng CK. Long-range correlation properties of coding and noncoding DNA sequences: GenBank analysis. Phys Rev E 1995;51:5084–91.
[6] Peng CK, Havlin S, Stanley HE, Goldberger AL. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos 1995;5:82–7.
[7] Buldyrev SV, Dokholyan NV, Goldberger AL, Havlin S, Peng CK, Stanley HE. Analysis of DNA sequences using methods of statistical physics. Physica A 1998;249:430–8.
[8] Liu Y, Gopikrishnan P, Cizeau P, Meyer M, Peng CK, Stanley HE. Statistical properties of the volatility of price fluctuations. Phys Rev E 1999;60:1390–400.
[9] Grau-Carles P. Long-range power-law correlations in stock returns. Physica A 2001;299:521–7.
[10] Alvarez-Ramirez J, Alvarez J, Rodriguez E. Short-term predictability of crude oil markets: a detrended fluctuation analysis approach. Energy Econ 2008;30:2645–56.
[11] Xia J, Shang P, Wang J. Estimation of local scale exponents for heartbeat time series based on DFA. Nonlinear Dyn 2013;74:1183–90.
[12] Podobnik B, Stanley HE. Detrended cross-correlation analysis: a new method for analyzing two nonstationary time series. Phys Rev Lett 2008;100:084102.
