Metrika
DOI 10.1007/s00184-009-0285-9

Estimating the codifference function of linear time series models with infinite variance

Dedi Rosadi · Manfred Deistler

Received: 12 February 2009

Abstract We consider the *codifference* and the *normalized codifference* function as dependence measures for stationary processes. Based on the empirical characteristic function, we propose estimators of the *codifference* and the *normalized codifference* function. We show consistency of the proposed estimators, where the underlying model is the ARMA with symmetric α-stable innovations, $0 < \alpha \leq 2$. In addition, we derive their limiting distribution. We present a simulation study showing the dependence of the estimator on certain design parameters. Finally, we provide an empirical example using some stocks from Indonesia Stock Exchange.

Keywords ARMA · Infinite variance · Codifference · Empirical characteristic function

1 Introduction

In many cases, the assumption of normality for the observations seems to be reasonable. On the other hand, in a number of applications, such as in signal processing, telecommunications, finance, physics and chemistry, the leptokurtic distribution, i.e., the distribution which is heavy-tailed and peaked around the center, seems to be more

D. Rosadi (✉)
Department of Mathematics, Gadjah Mada University,
Sekip Utara, Yogyakarta, Indonesia
e-mail: dedirosadi@ugm.ac.id

M. Deistler
Institute for Mathematical Methods in Economics,
Research Unit Econometrics and System Theory, Vienna University of Technology,
Argentinierstrasse 8, A1040 Vienna, Austria
e-mail: manfred.deistler@tuwien.ac.at

Published online: 30 September 2009

appropriate (e.g., Rachev and Mittnik 2000; Nikias and Shao 1995). An important class of distributions in this context is the stable distribution, which is a flexible class for data modeling and contains normal distributions as its special case. The importance of this class of distributions is strongly supported by generalized central limit theorems, which indicate that the stable distributions are the only possible limiting distributions for normed sums of independent and identically distributed random variables. For information on stable distribution, the reader is referred to, e.g., Samorodnitsky and Taqqu (1994).

In this paper, we consider univariate strictly stationary linear processes $\{X_t, t \in \mathbb{Z}\}$, where $\mathbb{Z}$ denotes the integers, given by

$$X_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j} \tag{1}$$

where the following holds

(C1) The coefficients c_j are real-valued and satisfy $|c_j| < cQ^{-j}$ for some $c > 0$, $Q > 1$

(C2) ϵ_t is i.i.d. symmetric α stable ($S\alpha S$) distributed, i.e., ϵ_t has a characteristic function of the form

$$E \exp(is\epsilon_t) = \exp(-\sigma^\alpha |s|^\alpha) \tag{2}$$

where α denotes *the index of stability* ($0 < \alpha \leq 2$), and $\sigma \geq 0$ denotes *the scale parameter*.

Under conditions C1 and C2, $\sum_{j=0}^{\infty} |c_j|^\alpha < \infty$ holds, and the infinite sum (1) is well defined in the sense of a.s. convergence. Moreover, the process $\{X_t\}$ will be a strictly stationary $S\alpha S$ process with the same index of stability α and with scale parameter $\sigma_X = \sigma(\sum_{j=0}^{\infty} |c_j|^\alpha)^{1/\alpha}$ (Samorodnitsky and Taqqu 1994, Theorem 7.12.2). Consider the $ARMA(p, q)$ system

$$\Phi(z)X_t = \Theta(z)\epsilon_t \tag{3}$$

for $t \in \mathbb{Z}$, where the polynomials Φ and Θ are given as

$$\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p,$$
$$\Theta(z) = 1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q,$$

with the real coefficients $\phi_1, \phi_2, \ldots, \phi_p$ and $\theta_1, \ldots, \theta_q$, when z denotes backward-shift operator ($z(X_t) = (X_{t-1})$) as well as the complex variable and when ϵ_t fulfils the condition C2. We assume that the polynomial $\Phi(z)$ has no roots in the closed unit disk $\{z : |z| \leq 1\}$ and then the system (3) has a unique stationary solution of the form (1) which fulfils condition C1. In addition, it is assumed that the polynomials $\Theta(z)$ and $\Phi(z)$ have no common zeros.

Here notice that if $\alpha = 2$, then ϵ_t is i.i.d. Gaussian with $\text{var}(\epsilon_t) = 2\sigma^2$. When $\alpha < 2$, $E|\epsilon_t|^p = \infty$ for $p \geq \alpha$ and $E|\epsilon_t|^p < \infty$ for $0 < p < \alpha$. The second moments of ϵ_t (and X_t) exist only for $\alpha = 2$, and for $\alpha < 2$, one can not use the covariance function to describe the dependence structure of the process $\{X_t\}$. Some generalizations of the autocovariance function as dependence measures of stationary process with infinite variance have been proposed in the literature, e.g., the *autocovariation* (see, e.g., Samorodnitsky and Taqqu 1994), the *codifference* function (e.g., Kokoszka and Taqqu 1994; Samorodnitsky and Taqqu 1994) and the *dynamical function* Janicki and Weron (1994). In this paper, we consider the codifference function and analyze the properties of its estimator, both by an analytical investigation and a simulation study.

The rest of this paper is organized as follows. In Sect. 2, we present the main results of this paper. In this section, we give the definition of the codifference and the normalized codifference function, and also propose their estimators. Furthermore, for a class of linear processes, we show consistency of the sample codifference function, and further establish the limiting distribution of the proposed estimator. In Sect. 3, we present several simulation studies for the estimation of the normalized codifference function of pure moving average processes. Finally, in Sect. 4 we provide an empirical example using some stocks from Indonesia Stock Exchange (IDX).

2 Dependence structure of linear time series model with infinite variance

2.1 Definition of the codifference function and its estimator

We consider the codifference function as proposed in Kokoszka and Taqqu (1994) and Yang et al. (2001)

$$\begin{aligned}\tau(k) = \tau(s, -s; k) = &- \ln \mathrm{E} \exp(is(X_{t+k} - X_t)) + \ln \mathrm{E} \exp(isX_{t+k}) \\ &+ \ln \mathrm{E} \exp(-isX_t)\end{aligned} \tag{4}$$

where $s \in \mathbb{R}$ and $k \in \mathbb{Z}$. Because the characteristic function always exists, the codifference function requires no moment conditions for the original process $\{X_t\}$. In the Gaussian case, the codifference function is proportional to the covariance function, i.e., $\tau(s, -s; k) = -s^2\gamma(k)$, where $\gamma(\cdot)$ denotes the covariance function of the stationary process $\{X_t\}$. Moreover, by defining the normalized codifference function $I(k)$ as

$$I(k) = \frac{\tau(s, -s; k)}{\tau(s, -s; 0)} \tag{5}$$

one directly obtains $I(k) = \rho(k)$ in the Gaussian case and it does not depend on s, where $\rho(k)$ denotes the correlation function.

Note that in general $\tau(-k) = \tau(k)^*$, where $\tau(k)^*$ denotes the conjugate of $\tau(k)$. In particular, under the assumptions C1 and C2, we obtain that the codifference function

$\tau(k)$ of the linear process (1) is of the form (see Kokoszka and Taqqu 1994)

$$\tau(k) = \sigma^{\alpha}|s|^{\alpha}\left[\sum_{j=0}^{\infty}(\left|(c_{j+k}-c_j)\right|^{\alpha} - \left|c_{j+k}\right|^{\alpha} - \left|-c_j\right|^{\alpha})\right], \quad k \geq 0 \tag{6}$$

Note that under conditions C1 and C2, the normalized codifference function (5) is independent of the choice of s. In what follows, we consider $S\alpha S$ processes $\{X_t, t \in \mathbb{Z}\}$ (see, e.g., Samorodnitsky and Taqqu 1994). For these processes, $\tau(1,-1;k)$ is identical to the codifference function $u(k)$ as given in Samorodnitsky and Taqqu (1994), eq. (4.7.1), where for a given (strictly) stationary $S\alpha S$ process $\{X_t, t \in \mathbb{Z}\}$, $u(k)$ is defined as

$$u(k) = 2(\sigma_{X_t})^{\alpha} - (\sigma_{X_{t+k}-X_t})^{\alpha} \tag{7}$$

where σ_Z denotes a scale parameters of Z.

Notice that if X_t's are independent, then for $k \neq 0$, $u(k) = 0$, and clearly, $\tau(k) = 0$ for all s. Conversely, if $u(k) = 0, k \neq 0$ and $0 < \alpha < 1$, then X_t are independent. When $1 \leq \alpha < 2$, $u(k) = 0$ does not imply that X_{t+k} and X_t are independent Samorodnitsky and Taqqu (1994).

Using Property 2.10.5 in Samorodnitsky and Taqqu (1994), we obtain that the normalized codifference $I(k)$ has the following property

$$0 \leq I(k) \leq 1 \quad \text{if } 0 < \alpha \leq 1 \tag{8}$$

$$1 - 2^{\alpha-1} \leq I(k) \leq 1 \quad \text{if } 1 \leq \alpha \leq 2 \tag{9}$$

when $\alpha = 2$, (9) becomes $-1 \leq \rho(k) \leq 1$, where $\rho(\cdot)$ denotes the autocorrelation function. Further theoretical properties of the codifference function were studied in Kokoszka and Taqqu (1994), Samorodnitsky and Taqqu (1994), Nowicka (1997) and Nowicka and Weron (1997), Rosadi (2005).

As the codifference function is defined via characteristic functions (*cf*), it can be estimated by empirical characteristic functions (*ecf*) (see, e.g., Yu 2004 for a review on *ecf*). Given a sample $X_1, X_2, \ldots, X_N$, an estimator for the codifference function at lag $k \in \mathbb{Z}$ can be defined as ($s \in \mathbb{R}$)

$$\hat{\tau}(s,-s;k) = \sqrt{(N-k)/N} \times [-\ln\phi(s,-s;k) + \ln\phi(s,0;k) + \ln\phi(0,-s;k)] \tag{10}$$

where for $u, v \in \mathbb{R}$

$$\phi(u,v;k) = \begin{cases} (N-k)^{-1}\sum_{t=1}^{N-k}\exp(i(uX_{t+k}+vX_t)) & \text{when } k \geq 0 \\ (N+k)^{-1}\sum_{t=1}^{N+k}\exp(i(uX_{t-k}+vX_t)) & \text{when } k < 0 \end{cases} \tag{11}$$

Accordingly, $\hat{I}(s,-s;k) = \frac{\hat{\tau}(s,-s;k)}{\hat{\tau}(s,-s;0)}$ can be used as the estimator of the normalized codifference $I(k)$. Here we consider a discrete estimation procedure, i.e., we evaluate the codifference function at r points $s_1 < \cdots < s_r$, for $s_i \in \mathbb{R}$, $s_i \neq 0$, $i = 1, \ldots, r$. We define the vectors $\mathbf{s} = [s_1, \ldots, s_r]^T$,

$$\hat{\tau}(\mathbf{s},k) = [\hat{\tau}(s_1,-s_1;k), \hat{\tau}(s_2,-s_2;k), \ldots, \hat{\tau}(s_r,-s_r;k)]^T$$

and

$$\hat{I}(\mathbf{s},k) = [\hat{I}(s_1,-s_1;k), \hat{I}(s_2,-s_2;k), \ldots, \hat{I}(s_r,-s_r;k)]^T$$

We similarly denote $\tau(\mathbf{s},k)$ and $I(\mathbf{s},k)$ as the vectors of the codifference and normalized codifference function, respectively. Note that one can replace the factor $\sqrt{(N-k)/N}$ in (10) by unity, and also the divisor $(N-k)$ in (11) by N without changing the asymptotic properties of the estimator, however, the choices in (10) and (11) will give a better finite sample performance than the alternative. Note that $\hat{\tau}(-k) = \hat{\tau}(k)$ such one can restrict the analysis to the case of $k \geq 0$. Two similar estimators for the codifference function have been proposed in Yang et al. (2001) and in Hong (1999).

2.2 The asymptotic properties of the estimator

The asymptotic properties of the sample codifference are summarized in the following theorems.

Theorem 1 *Let $\{X_t, t \in \mathbb{Z}\}$ be a stationary linear process (1) satisfying conditions C1 and C2. For $\mathbf{s} \in \mathbb{R}^r$, $s_i \neq 0$, $i = 1, \ldots, r$, $k \in \{0, 1, 2, \ldots\}$ its sample codifference $\hat{\tau}(\mathbf{s},k)$ and the sample normalized codifference $\hat{I}(\mathbf{s},k)$ are (weakly) consistent estimators for $\tau(\mathbf{s},k)$ and $I(\mathbf{s},k)$, respectively.*

The proof is given in Appendix A.

The asymptotic distribution of the sample codifference function (and the sample normalized codifference function) of the linear process (1) can be derived using a central limit theorem for empirical characteristic functions (Hesse 1990, Theorem 1 and Remark 2.6.). For convenience, we split $\hat{\tau}$ into its real and imaginary parts. We write

$$\mathrm{Re}\,\hat{\tau}(\mathbf{s},k) = \left[\mathrm{Re}\,\hat{\tau}(s_1,-s_1;k), \mathrm{Re}\,\hat{\tau}(s_2,-s_2;k), \ldots, \mathrm{Re}\,\hat{\tau}(s_r,-s_r;k)\right]^T$$

and

$$\mathrm{Im}\,\hat{\tau}(\mathbf{s},k) = \left[\mathrm{Im}\,\hat{\tau}(s_1,-s_1;k), \mathrm{Im}\,\hat{\tau}(s_2,-s_2;k), \ldots, \mathrm{Im}\,\hat{\tau}(s_r,-s_r;k)\right]^T$$

Here, $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$, $z \in \mathbb{C}$ denote the real and imaginary parts of z. As $\hat{\tau}(s,-s,0)$ by definition is a real function, we therefore obtain

$$\operatorname{Re}\hat{I}(\mathbf{s},k)=\begin{bmatrix}\operatorname{Re}\hat{\tau}(s_1,-s_1;k)/\hat{\tau}(s_1,-s_1;0)\\ \operatorname{Re}\hat{\tau}(s_2,-s_2;k)/\hat{\tau}(s_2,-s_2;0)\\ \vdots\\ \operatorname{Re}\hat{\tau}(s_r,-s_r;k)/\hat{\tau}(s_r,-s_r;0)\end{bmatrix},$$

$$\operatorname{Im}\hat{I}(\mathbf{s},k)=\begin{bmatrix}\operatorname{Im}\hat{\tau}(s_1,-s_1;k)/\hat{\tau}(s_1,-s_1;0)\\ \operatorname{Im}\hat{\tau}(s_2,-s_2;k)/\hat{\tau}(s_2,-s_2;0)\\ \vdots\\ \operatorname{Im}\hat{\tau}(s_r,-s_r;k)/\hat{\tau}(s_r,-s_r;0)\end{bmatrix} \tag{12}$$

In the following theorem, a result regarding the asymptotic distribution of the sample normalized codifference is given. The proof is given in Appendix B.

Theorem 2 *Let $\{X_t, t\in\mathbb{Z}\}$ be a stationary linear process (1), satisfying conditions C1 and C2. Then for $h\in\{1,2,\ldots\}$,*

$$\left[\begin{pmatrix}\operatorname{Re}\hat{I}(\mathbf{s},1)\\ \operatorname{Im}\hat{I}(\mathbf{s},1)\end{pmatrix},\begin{pmatrix}\operatorname{Re}\hat{I}(\mathbf{s},2)\\ \operatorname{Im}\hat{I}(\mathbf{s},2)\end{pmatrix},\ldots,\begin{pmatrix}\operatorname{Re}\hat{I}(\mathbf{s},h)\\ \operatorname{Im}\hat{I}(\mathbf{s},h)\end{pmatrix}\right]^T$$

is

$$AN\left(\left[\begin{pmatrix}l_rI(1)\\ l_r0\end{pmatrix},\begin{pmatrix}l_rI(2)\\ l_r0\end{pmatrix}\ldots,\begin{pmatrix}l_rI(h)\\ l_r0\end{pmatrix}\right]^T,N^{-1}\mathbf{W}\right) \tag{13}$$

The matrix variance-covariance $\mathbf{W}$ is given in (43) and $l_r=[1,1,\ldots,1]^T\in\mathbb{R}^r$.

Applying this theorem to the special case of i.i.d. observations, we obtain:

Corollary 1 *Let $\{X_t, t\in\mathbb{Z}\}$ be an i.i.d. sequence satisfying the condition C2. Then for $k\in\{1,2,\ldots\}$,*

$$\operatorname{Re}\hat{I}(\mathbf{s},k)\ \textit{is}\ AN(0,N^{-1}\mathbf{W}_1) \tag{14}$$

and

$$\operatorname{Im}\hat{I}(\mathbf{s},k)\ \textit{is}\ AN(0,N^{-1}\mathbf{W}_2) \tag{15}$$

where the (i,j)th elements of matrix $\mathbf{W}_1$ and $\mathbf{W}_2$ are,

$$W_1(i,j)=\frac{f_{ij}}{g_{ij}}\ \textit{and}\ \ W_2(i,j)=\frac{h_{ij}}{g_{ij}}\ \ ,i,j=1,\ldots,r \tag{16}$$

with

$$
\begin{aligned}
f_{ij} &= e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i-s_j|^\alpha)}\left\{\frac{1}{2}e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i-s_j|^\alpha)}-1\right\}\\
&\quad+e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i+s_j|^\alpha)}\left\{\frac{1}{2}e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i+s_j|^\alpha)}-1\right\}+1\\
h_{ij} &= e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i-s_j|^\alpha)}\left\{\frac{1}{2}e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i-s_j|^\alpha)}-1\right\}\\
&\quad+e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i+s_j|^\alpha)}\left\{1-\frac{1}{2}e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i+s_j|^\alpha)}\right\}
\end{aligned}
$$

and

$$
g_{ij} = 4\sigma^{2\alpha}\,|s_i|^\alpha\,\left|s_j\right|^\alpha.
$$

The proof is given in Appendix C.

3 Simulation evidence

In this section, we present a simulation study for investigating the small sample properties of the sample normalized codifference function.

3.1 Practical considerations

Before we proceed, we make a remark about the sample and the population codifference function. From (6) and the fact that all c_j's are real, we see that the population codifference function is a real valued function. On the other hand, its estimator (10) is complex valued. Because the estimator we propose is consistent, one possibility is to use only the real part of the estimator.

In what follows, we are only considering with the sample normalized codifference $\hat{I}(k)$. We note that unlike the population, the sample normalized codifference function depends on $\mathbf{s} = \{s_1, \ldots, s_r\}$. Apparently $\hat{I}(\cdot)$ is defined for all $\mathbf{s} > \mathbf{0}$. However, in a finite sample, the accuracy of the estimator of the population values depends on the choice of $\mathbf{s}$. Therefore, for estimation, $\mathbf{s}$ is a design parameter, which has to be chosen appropriately. In other words, $\hat{I}(\cdot)$ should be calculated from those values of $\mathbf{s}$ which give the most accurate estimates of the true function $I(\cdot)$.

To be more precise, in practice the number of grid points r and more importantly, the location of $s_1, \ldots, s_r$, have to be chosen. As the normalized codifference function is defined based on the *ecf*, we can apply here the known results for the *ecf*. For a fixed r, Koutrouvelis (1980) and Kogon and Williams (1998) showed that for calculating the *ecf*, the location of the grid points $s_1, \ldots, s_r$, should be chosen close, but not equal to zero, since then the *ecf* will give accurate estimates of the characteristic function. Such a choice of grid points seems to be reasonable for calculating $\hat{I}(\cdot)$ too, see Fig. 1. To determine the location of s_i's, we suggest to plot $\mathrm{Re}\,\hat{I}(s, k)$ for $0.01 \leq s \leq 2$,

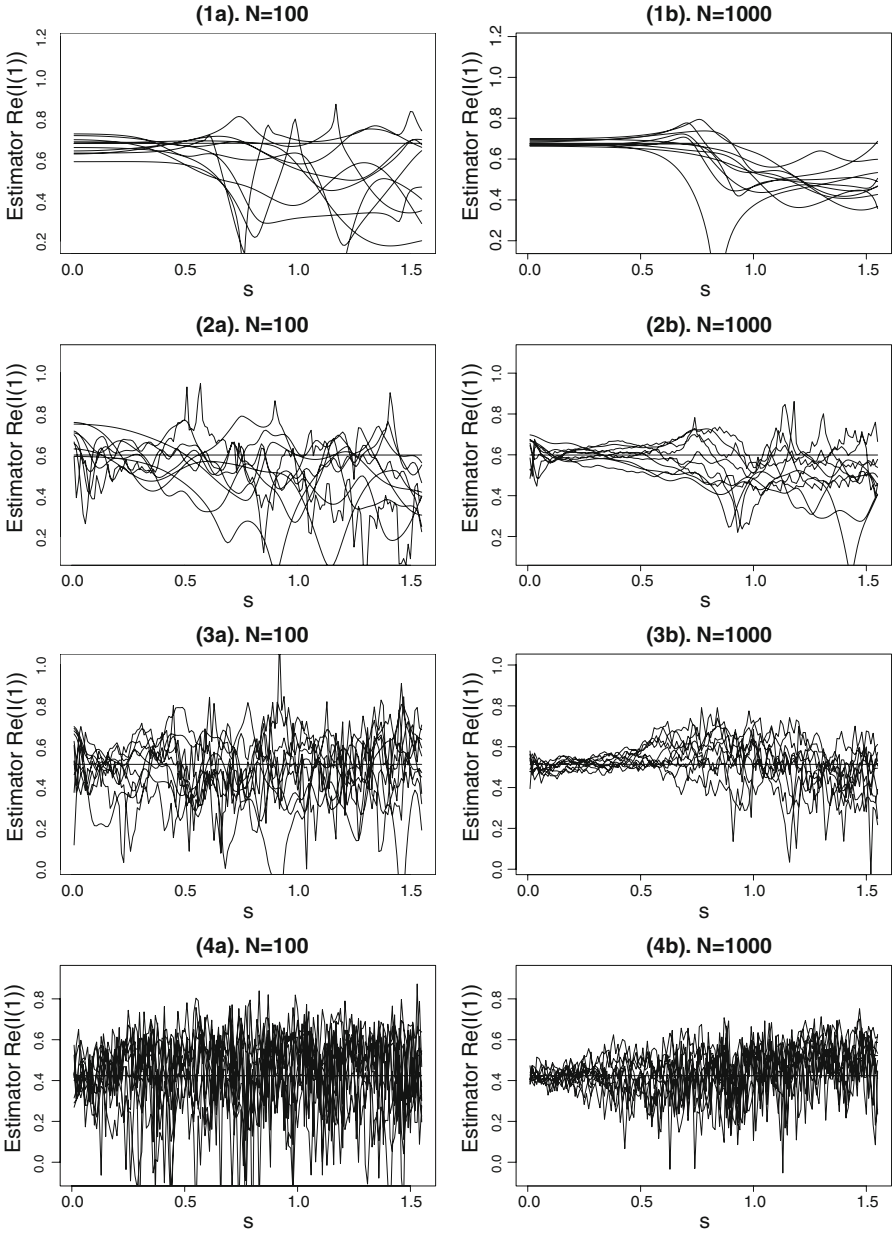

Fig. 1 Plots of Re $\hat{I}(1)$ for several simulation runs. Here $\alpha = 2$ (**1a**–**1b**) , $\alpha = 1.5$ (**2a**–**2b**), $\alpha = 1$ (**3a**–**3b**) and $\alpha = 0.5$ (**4a**–**4b**) and $\sigma = 1$, $s \in [0.01, 1.55]$. Data are generated from experiment I that is MA(2) process with $c_0 = 1$, $c_1 = 2$ and $c_2 = 1.111$. The *straight lines* denote the true values of $I(1)$

for some values of lag $k > 0$. These graphs will show ranges for s where the bias is small. The best choice for range of s clearly depends on the data itself and in general also on the lag k. However, we suggest to use $s_a = 0.01$ as the left bound of range,

where the best choice for the right bound should be determined from the graphs, i.e., as the threshold of $s = s_b$ where the graphs of Re $\hat{I}(k)$, for some lag k, are still relatively flat. The individual choices for s_i's can be chosen in one of the following two ways:

1. If we wish to use equal spacing of s_i's, we can set $\mathbf{s} = \{s_1 = 0.01, 0.01 + i\frac{s_b - s_a}{r-1}, s_b\}, i = 1, \ldots, r-2$. Here we choose s_i's by plotting the sample codifference for given values of s_i as given in Fig. 1. Figure 1 indicates these sample codifferences apparently depend on the sample size N and additionally on α. From our simulation studies, we observe that the sample codifference has a strong dependency on α, less dependence on lag k. For $\alpha = 2$, $\hat{I}(k)$ is relatively smooth, where for $\alpha < 2$, $\hat{I}(k)$ has an erratic behavior, and this behavior is stronger for smaller α. This result suggests that when $\alpha = 2$, there is no benefit by choosing the s_i's very close to each other. When α is getting smaller, then we need to choose s_i's sufficiently close to each other. Here, we propose to use s_i's with distance $d = 0.01$ for $\alpha \leq 1$, $0.01 < d \leq 0.05$ for $1 < \alpha \leq 1.5$, $0.05 < d \leq 0.1$ for $1.5 < \alpha < 2$ and $d = 0.1$ or larger for $\alpha = 2$. Especially in the i.i.d. case we can show that these choices are sufficient, in the sense that for given α, choosing a smaller distance between grid points will not significantly decrease the determinant of the covariance matrix (14). However, as the erratic behavior of the sample codifference is typical for given α, in practice it is not necessary to know α in advance.
2. In case of unequal spacing, when r has been fixed, the choice of s_i's can be chosen using a similar consideration as above, i.e., we choose the s_i's sufficiently close, depending on the erratic behavior of the estimates Re $\hat{I}(\cdot)$.

Finally we have to choose the number of points r. For $r \geq 1$, the final estimate $\hat{I}(k)$ can be defined as the weighted average of the estimates at the grid points $s_1, \ldots, s_r$, i.e., we define $\hat{I}(k) = \sum_{i=1}^{r} w_i \hat{I}(s_i, -s_i; k)$ with $\sum_{i=1}^{r} w_i = 1$. For instance, we can use a simple average with $w_i = 1/r$ or a negative exponentially weighted average with $w_i = \exp(-s_i^2)/\sum_{j=1}^{r} \exp(-s_j^2)$. In i.i.d. case, we obtain that by simply averaging the estimates at different points, the asymptotic variance of the estimator will be smaller or equal to the variance of the estimator obtained at a single point, which can be seen directly from Fig. 2. Figure 2 shows that for $\alpha = 2$, there is no difference in terms of the asymptotic variance between estimating $\hat{I}(\cdot)$ either at a single point or at more points, whereas for $\alpha < 2$, the difference is significant, especially when α is small. Furthermore, Fig. 2 also shows that the smaller α is, the smaller is the covariance between the grid points. For the finite sample case, this fact is consistent with the typical erratic behavior of the plot Re $\hat{I}(\cdot)$ of some non i.i.d. samples, as shown in Fig. 1. Based on these results and from our simulations, we suggest the choice of the number of grid points r as follows. For $\alpha = 2$ (i.e., for smooth graphs of Re $\hat{I}(k)$, $k > 0$), we observe that $r = 1$ is sufficient, whereas for $\alpha < 2$ (i.e., for erratic graphs of Re $\hat{I}(k)$, $k > 0$), at least two points should be chosen, and the number of points should be larger for smaller α (i.e., for more erratic graphs of Re $\hat{I}(k)$, $k > 0$). It is important to note that from our simulation experience, the accuracy of the estimator is more sensitive to the location of the grid points than to their number.

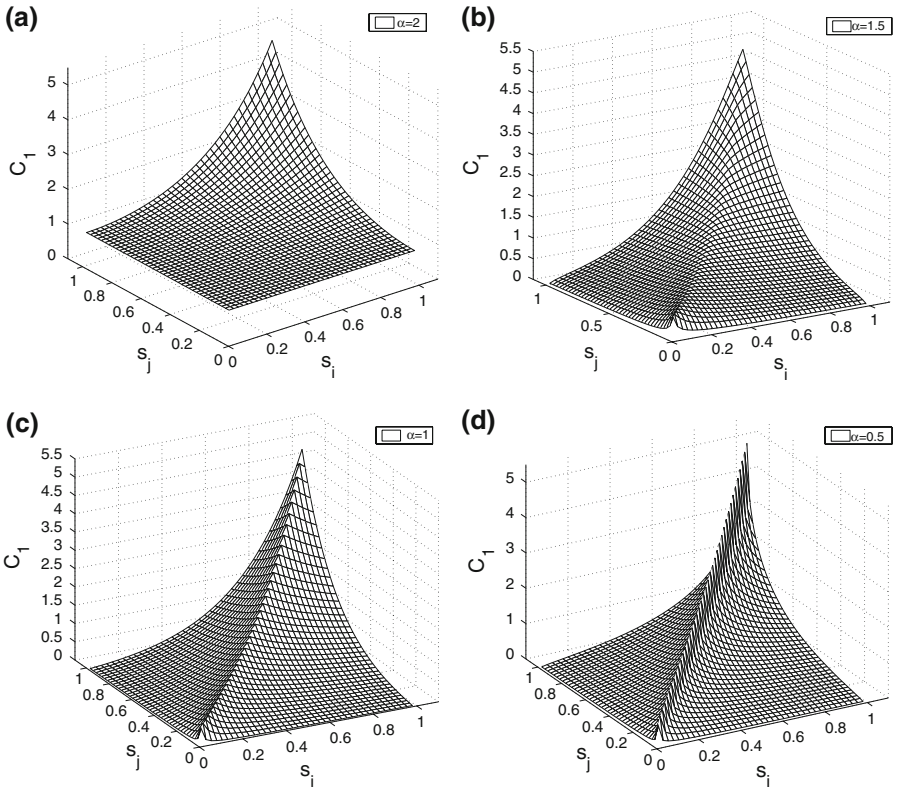

Fig. 2 Plots of $W_1(i, j)$ [see eq. (16)], for $s_i, s_j \in [0.01, 1]$, and some α's

In the following subsection, we will investigate the choice of grids points through monte carlo simulations.

3.2 Simulation results

To investigate the finite sample behavior of the sample codifference, in particular its dependence on the choice of **s**, we run several monte-carlo simulations using *R/GNU-S* version 2.7.0 R Development Core Team (2008) (available on the Web at http://www.r-project.org), where we use function *rstable* in the extension package `stable` (available on the Web at http://alpha.luc.ac.be/~jlindsey/rcode.html), to generate the unit symmetric α stable innovations (based on the method presented in Chambers et al. 1976) and function *arima.sim* in the package `stats` to generate $X_t = \epsilon_t + c_1\epsilon_{t-1} + c_2\epsilon_{t-2}$ where (c_1, c_2) are

$$I.\,(2, 1.111) \quad II.\,(-1, 0.5) \quad III.\,(0.55, 0.05) \quad IV.\,(-0.4, 0.7) \tag{17}$$

and from now on we refer to these as experiment I–experiment IV, respectively.

In the Gaussian framework, the models correspond to the experiments I, II and III are examined in Bhansali (1983). The roots of the polynomial $1 + c_1 z + c_2 z^2 = 0$ are as follows. In experiment I, the roots are $-0.9 \pm 0.3i$, close to the invertibility region. In experiment II and IV, the roots are $0.5 \pm 0.5i$, and $-0.2857 \pm 0.247i$, so the absolute values of the roots are 0.71, and 0.378, respectively. In these experiments, the models have similar roots properties which are neither too close to 1 nor to 0. In experiment III, the roots are real-valued, equal to -0.435 and -0.115, one close to 0.5 and the other close to 0.

For $\alpha = 2$, the true values of the normalized population codifference (equal to the correlation function) at lag $k = 1, 2$, $(I(1), I(2)) = (\rho(1), \rho(2))$ in experiments I–IV are:

$$(I).\,(0.677, 0.178), \quad (II).\,(-0.667, 0.222), \quad (III).\,(0.443, 0.038),$$
$$(IV).\,(-0.412, 0.424)$$

In experiments I and II, the values of I(k) are closer to 1 at lag 1 and not too close to 0 at lag 2 while for experiment III, at lag 1 close to 0.5 but almost 0 at lag 2. For the last experiment, at lag 1 it is negative but it is positive at lag 2, with absolute values near 0.5.

In order to see the performance of the estimator, we simulate the time series in experiments I–IV for several values of α with $\sigma = 1$ and two sample sizes, the "small" one is $N = 100$ and the "large" one is $N = 1{,}000$. All experiments are replicated $T = 1{,}000$ times. The sample normalized codifferences are calculated for lags 1 to 10. Figure 1 suggests that in the interval $0.01 \leq s \leq 0.5$, $\mathrm{Re}\,\hat{I}(\cdot)$ is relatively less biased, although the best range for $\mathbf{s}$ depends on the index α. For checking the impact of the choice of grid points $\mathbf{s}$, we choose several different sets of $\mathbf{s}_i = \{s_1, \ldots, s_r\}$, $i = 1, 2, \ldots, 28$. Here we consider equidistant as well as non equidistant grid points. The complete list of the choices is as follows: $\mathbf{s}_1 = \{0.01\}$, $\mathbf{s}_2 = \{0.1\}$, $\mathbf{s}_3 = \{0.2\}$, $\mathbf{s}_4 = \{0.3\}$, $\mathbf{s}_5 = \{0.5\}$, $\mathbf{s}_6 = \{1\}$, $\mathbf{s}_7 = \{0.01, 0.1\}$, $\mathbf{s}_8 = \{0.01, 0.2\}$, $\mathbf{s}_9 = \{0.01, 0.5\}$, $\mathbf{s}_{10} = \{0.01, 1\}$, $\mathbf{s}_{11} = \{0.1, 0.2\}$, $\mathbf{s}_{12} = \{0.1, 0.5\}$, $\mathbf{s}_{13} = \{0.1, 1\}$, $\mathbf{s}_{14} = \{0.5, 1\}$, $\mathbf{s}_{15} = \{0.01, 0.1, 0.2\}$, $\mathbf{s}_{16} = \{0.01, 0.1, 0.5\}$, $\mathbf{s}_{17} = \{0.01, 0.1, 1\}$, $\mathbf{s}_{18} = \{0.1, 0.2, 0.3\}$, $\mathbf{s}_{19} = \{0.1, 0.3, 0.5\}$, $\mathbf{s}_{20} = \{0.01, 0.5, 1\}$, $\mathbf{s}_{21} = \{0.1, 0.5, 1\}$, $\mathbf{s}_{22} = \{0.1, 0.2, 0.3, 0.4, 0.5\}$, $\mathbf{s}_{23} = \{0.1, 0.2, \ldots, 1\}$, $\mathbf{s}_{24} = \{0.01, 0.06, 0.11, 0.16, 0.21\}$, $\mathbf{s}_{25} = \{0.01, 0.02, \ldots, 0.2\}$, $\mathbf{s}_{26} = \{0.01, 0.02, \ldots, 0.1\}$, $\mathbf{s}_{27} = \{0.11, 0.12, \ldots, 0.2\}$ and $\mathbf{s}_{28} = \{0.5, 0.55, \ldots, 1\}$. For each choice of $\mathbf{s}_i$ in run h, $h = 1, \ldots, T$, the final estimates are calculated as the weighted value of estimates corresponding to the choices of grid points s_{ij}, $j = 1, \ldots, r_i$. They are denoted by $\mathrm{Re}\,\hat{I}(\cdot)_{ih} = \sum_{j=1}^{r_i} w_{ij}\,\mathrm{Re}\,\hat{I}(s_{ij}, -s_{ij}, \cdot)_h$, where $\mathrm{Re}\,\hat{I}(s_{ij}, -s_{ij}, \cdot)_h$ denotes the real part of the sample normalized codifference in run h at a certain lag, calculated at s_{ij}, $j = 1, \ldots, r_i$. Here we consider two methods for weighting the estimates, first we use a simple average of the estimates and second, we use a negative exponential weighted average. To save space, we only present the result of experiment I, which is summarized in Table 1, but the results in the other experiments are similar. In the table, we also record the best choices of $\mathbf{s}$, which are defined as the values of grid points which minimize the sum of mean absolute deviation (MAD) of estimates at lag 1 and lag 2, among all considered choices of grid points above. Here, MAD at lag k and

Table 1 The true values $I(\cdot)$ and the estimates $\hat{I}(\cdot)$

N	α	Method	s	$I(1)$	Avg. $\hat{I}(1)$	MAD_1	$I(2)$	Avg. $\hat{I}(2)$	MAD_2
100	2	Avg.	{0.01}		0.66180	0.04600		0.14647	0.10186
		Exp.	{0.01, 0.1, 1}	0.67722	0.63195	0.05732	0.17821	0.19387	0.09025
		ACF	–		0.65848	0.04644		0.14500	0.10115
	1.8	Avg.	{0.01, 0.2}	0.64700	0.64938	0.04415	0.19237	0.15860	0.09297
		Exp.	{0.01, 0.1, 1}		0.62509	0.04760		0.20409	0.08011
	1.5	Avg.	{0.01, 0.1, 0.2}	0.59903	0.60826	0.04839	0.21343	0.17158	0.07927
		Exp.	{0.01, 0.1, 1}		0.59728	0.04888		0.21062	0.06729
	1.3	Avg.	{0.01, 0.06, . . . , 0.21}	0.56554	0.57235	0.05032	0.22665	0.19017	0.06974
		Exp.	{0.01, 0.1, 1}		0.57364	0.05548		0.22380	0.06105
	1	Avg.	{0.01, 0.02, . . . , 0.2}	0.51350	0.50708	0.04745	0.24325	0.21576	0.06049
		Exp	{0.01, 0.02, . . . , 0.2}		0.50717	0.04739		0.21577	0.06036
	0.8	Avg.	{0.01, 0.02, . . . , 0.1}	0.47792	0.47274	0.04357	0.25014	0.22644	0.05599
		Exp.	{0.01, 0.02, . . . , 0.1}		0.47276	0.04355		0.22643	0.05597
	0.5	Avg.	{0.01, 0.02, . . . , 0.1}	0.42379	0.40713	0.04463	0.24809	0.22776	0.06033
		Exp.	{0.01, 0.02, . . . , 0.1}		0.40715	0.04459		0.22776	0.06028

Table 1 continued

N	α	Method	s	$I(1)$	Avg. $\hat{I}(1)$	MAD_1	$I(2)$	Avg. $\hat{I}(2)$	MAD_2
1,000	2	Avg.	$\{0.01\}$		0.67577	0.01335		0.17495	0.03096
		Exp.	$\{0.01\}$	0.67722	0.67577	0.01335	0.17821	0.17495	0.03096
		ACF	–		0.67544	0.01336		0.17477	0.03094
	1.8	Avg.	$\{0.01, 0.06, \ldots, 0.21\}$	0.64700	0.65059	0.01645	0.19237	0.18859	0.02708
		Exp.	$\{0.01, 0.06, \ldots, 0.21\}$		0.65067	0.01649		0.18854	0.02708
	1.5	Avg.	$\{0.01, 0.02, \ldots, 0.2\}$	0.59903	0.60204	0.01815	0.21343	0.20879	0.02155
		Exp.	$\{0.01, 0.02, \ldots, 0.2\}$		0.60209	0.01822		0.20878	0.02154
	1.3	Avg.	$\{0.01, 0.02, \ldots, 0.2\}$	0.56554	0.56751	0.01640	0.22665	0.22312	0.01996
		Exp	$\{0.01, 0.02, \ldots, 0.2\}$		0.56754	0.01644		0.22310	0.01992
	1	Avg.	$\{0.01, 0.02, \ldots, 0.1\}$	0.51350	0.51396	0.01521	0.24325	0.24105	0.01577
		Exp.	$\{0.01, 0.02, \ldots, 0.1\}$		0.51396	0.01521		0.24105	0.01576
	0.8	Avg.	$\{0.01, 0.02, \ldots, 0.1\}$	0.47792	0.47742	0.01383	0.25014	0.24811	0.01511
		Exp.	$\{0.01, 0.02, \ldots, 0.1\}$		0.47742	0.01382		0.24811	0.01510
	0.5	Avg.	$\{0.01, 0.02, \ldots, 0.1\}$	0.42379	0.42239	0.01296	0.24809	0.24548	0.01852
		Exp.	$\{0.01, 0.02, \ldots, 0.1\}$		0.42239	0.01295		0.24548	0.01851

The true values $I(\cdot)$ and the estimates $\hat{I}(\cdot)$ from the experiment I, that is MA(2) process with $c_0 = 1$, $c_1 = 2$ and $c_2 = 1.111$ for $T = 1,000$ replication, and for some sample size N. The ϵ_t is $S\alpha S$ process with some α and $\sigma = 1$. Here, $Avg.\hat{I}(i) = \frac{1}{T}\sum_{j=1}^{T} \text{Re}\,\hat{I}(i)_j$, and $MAD_i = \frac{1}{T}\sum_{j=1}^{T} |\,\text{Re}\,\hat{I}(i)_j - I(i)|$, $i = 1, 2$, where $\text{Re}\,\hat{I}(i)_j$ denotes the estimates at lag i in run j. The weighting methods here denote by the simple average (method Avg.) and the negative exponential weighted average (method Exp.). Further explanation about the table is given in Sect. 3.2

for grid $\mathbf{s}_i$ is defined as $MAD_{ik} = \frac{1}{T}\sum_{h=1}^{T} |\operatorname{Re}\hat{I}(k)_{ih} - I(k)|$, $k = 1, 2$. We also calculated mean square error (MSE) of estimates, but the results are very similar to MAD. For the sake of comparison, when $\alpha = 2$ we also record the estimates of the sample (central) ACF.

As expected, we observe that the estimation accuracy will be improved when the sample size is increased. Furthermore, throughout the simulation studies, the results indicate that the accuracy of the estimates of the normalized codifference function depends on the choice of grid points $\mathbf{s}$, where the optimal choices of grids depend on the index α and the sample size N. When $\alpha = 2$, surprisingly under suitable choices of grid points, we find in some cases that the sample normalized codifference can provide a better estimate (in terms of total MAD for the first two lags) for the true values (of the normalized codifference, which is equal to ACF) than the estimates given by the sample ACF. When $\alpha < 2$, it seems that there is a great benefit by evaluating $\operatorname{Re}\hat{I}(\cdot)$ at several points of s, that is $r > 1$, where under the appropriate choice of grid points $\mathbf{s}$, the performance of the weighting methods (the simple average and the exponential weight) are approximately the same. For MA(2) models, in all cases we consider here, we also find that the estimation accuracies are significantly better if we choose the grids point in $0.01 \leq s < 0.5$. In general, there is a benefit in terms of the estimation accuracy if we include a point close to zero. We further observe that when $\alpha < 1.5$, the choice of equidistant grids with a distance between 0.01 and 0.05 seems to be adequate. For $\alpha \geq 1.5$, a distance 0.1 seems to be adequate, a smaller grid distance does not really improve the accuracy of the estimates. As a general conclusion, from this simulation studies, we may conclude that the optimal choice of grid points $\mathbf{s}$ will follow the lines of our proposed choice of grid points $\mathbf{s}$ described in Sect. 3.1.

4 Application

For the practical application, we collected some daily closing price data of stocks INDF and the composite index IHSG in Indonesia Stock Exchange (IDX) during the period January 1, 2003 until December 31, 2007. The data are transformed into daily returns using $Y_t = \ln(X_{t+1}/X_t)$, $t = 1, \ldots, N-1$.

For checking stability of these return series, we estimate the index α of the data over different time horizons (e.g., over daily, weekly and monthly aggregate levels. See for instance, Paolella 2001). For each time horizon, we informally compare whether the confidence intervals of estimates of α (i.e., the average of estimates plus or minus two standard errors of the estimates) are below the threshold 2.0 and reject the possibility of using stable distribution if they are above 2.0. For this purpose, we estimate the parameters of the stable distribution using Nolan's numerical maximum likelihood method (Nolan 1999a), implemented in package fBasics in Rmetrics. The time horizons used in the study are $j = 1, 2, \ldots, 10$ days, where for $j > 1$, we calculate the aggregates as the average of the data over non-overlapping j-length segments of the data. Our choice of maximum time horizon $j = 10$ days here is only restricted by the length of the data available in the study. The fitting results using non-normal stable distribution for the stocks returns are given in Table 2.

Table 2 Estimated of α

Frequency (days)	1	2	3	4	5	6	7	8	9	10	$\bar{\hat{\alpha}} \pm 2 \times \mathrm{SE}(\hat{\alpha})$
Sample size	1,211	605	403	302	242	201	173	151	134	121	
Stocks											
IHSG composite index	1.62	1.56	1.57	1.49	1.61	1.72	1.80	1.87	1.75	1.80	1.68 ± 0.08
Indofood sukses (INDF)	1.74	1.78	1.70	1.68	1.80	1.64	1.73	1.77	1.68	1.70	1.72 ± 0.031

Estimated of α-values for samples defined with respect to the time intervals of different sizes. The last column denotes the mean plus or minus two standard error of $\hat{\alpha}$ over all time frequencies considered here

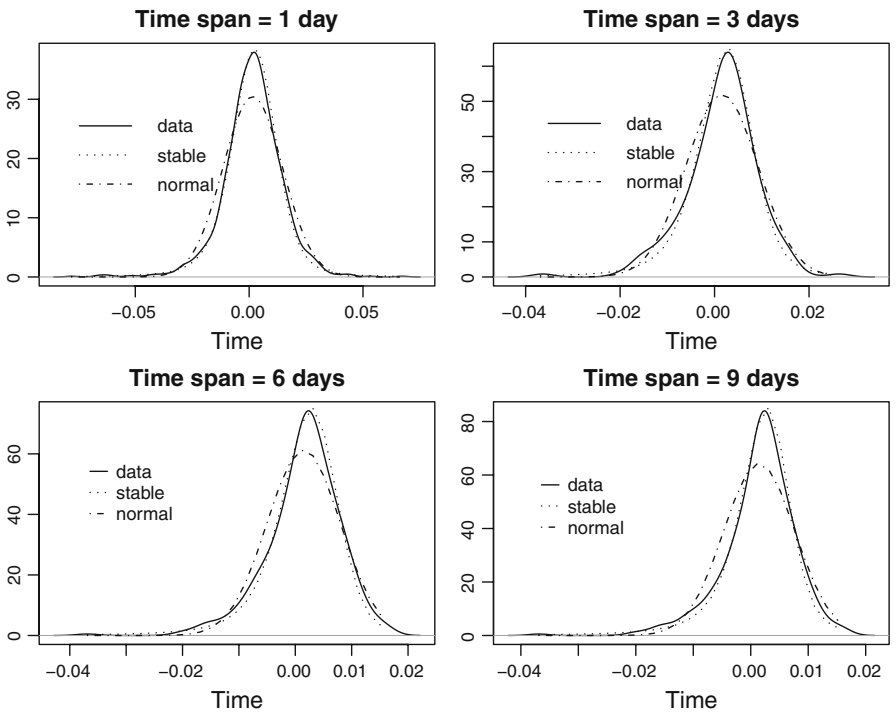

Fig. 3 Density plots for the composite index IHSG returns over different time horizons

We further obtain that based on the density plot (see also Nolan 1999b), the non-normal stable distributions show a better shape to the data than the fitting result obtained using the normal distribution (where the sample mean and the sample variance are used as the estimated parameters). An example for the density plot over different time horizons is given in Fig. 3. The similar results can also be obtained for INDFR stocks.

To obtain the estimates of the normalized codifference with a good accuracy, we plot $\mathrm{Re}\,\hat{I}(k)$ within the interval $s \in [0.01, 2]$, for lag $k = 1$ (see Fig. 4). These graphs are smooth, which indicate that $\hat{\alpha}$ is closer to 2, as confirmed from Table 2. Therefore, we may choose the grid points in the interval $0 < s < 0.5$ with a relatively large distance

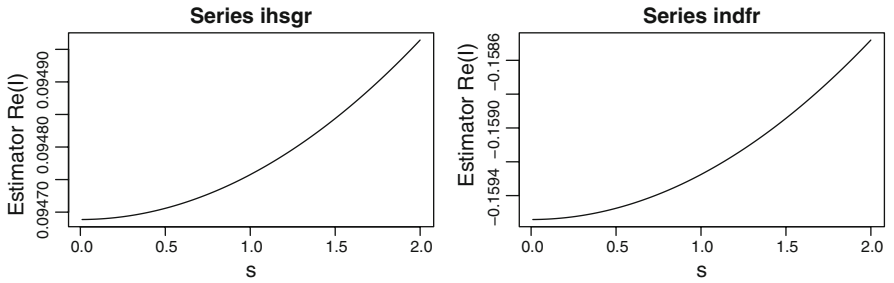

Fig. 4 Plots of Re $\hat{I}(1)$ for INDFR and IHSGR

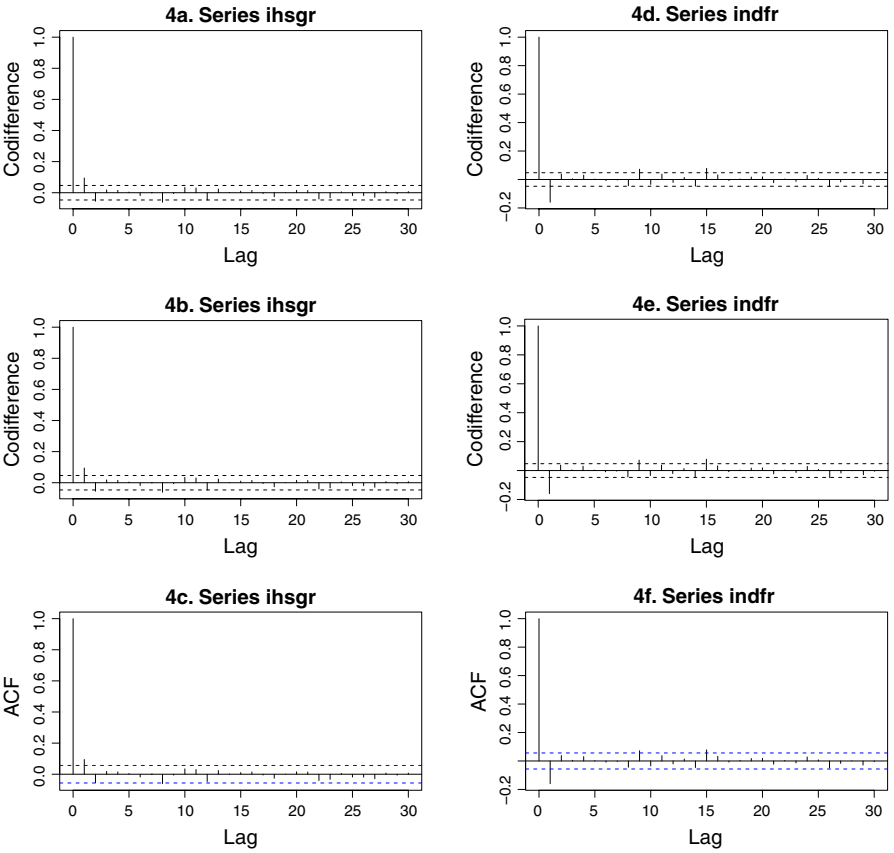

Fig. 5 Plots of the sample acf and the sample codifference function. In **4a** and **4d**, the sample codifference is evaluated at grids $s = 0.01$, where in **4b** and **4e**, we consider grids $s = \{0.01, 0.1, 0.5\}$

between the grid points. In Fig. 5, we plot Re $\hat{I}(k)$ at $\mathbf{s} = 0.01$ and $\mathbf{s} = (0.01, 0.1, 0.5)$ using a simple average as the weighting method. The confidence interval of estimates is calculated using Corollary 1 [see Eqs. (14) and (16)]. For comparison purpose, we also plot the sample ACF of the data. These results indicate that the data will follow

a moving average process with order 1. A further discussion on the identification of moving average process with infinite variance process is given in Rosadi (2007).

5 Conclusion

In this paper, we propose estimators of the codifference and the normalized codifference function, where for the linear processes with geometrically bounded coefficients and $S\alpha S$ innovations, we established the asymptotic properties of the proposed estimators. Notice that unlike the ACF estimator, we obtain that there is no discontinuity in either the normalization or the limiting distribution of the proposed estimators when $\alpha \to 2$. Moreover, we note that unlike the sample ACF which has an unfamiliar limiting distribution when $\alpha < 2$ and relatively difficult to obtain the quantiles of the limit distribution, estimators of the codifference and the normalized codifference will be asymptotically normally distributed at the same rate as the sample ACF in the classical case although the asymptotic variance is different. We also present a simulation study to observe the small sample properties of the normalized codifference estimator.

In the practical situation, to obtain the estimates of the normalized codifference with a good accuracy for the real data, first we suggest to plot Re $\hat{I}(k)$ within the interval $s \in [0.01, 2]$, for some values of lag $k > 0$. These graphs will show two important things. First, they suggest the interval of s near zero which has a small bias (i.e., the interval $0.01 < s < s_b$, s_b denotes the threshold point of s where the graphs Re $\hat{I}(k)$, for some k, are still relatively flat), as the best location for evaluating $\hat{I}(\cdot)$. Secondly, it reveals the erratic behavior of the estimates. When the graphs are smooth, the choice of one point $s = 0.01$ is sufficient for estimating Re $\hat{I}(k)$. If the graphs are erratic, at least two points are required and more points are better for more erratic graphs. If the equidistant points $s_1, \ldots, s_r$ are considered, when the graphs are highly erratic, we can use a small distance between points, e.g., 0.01, where for less erratic graphs, we can use a bigger distance, such as 0.05 or 0.1. If the non equidistant points are used, we should include one point close to zero in the choice of s_i's and the chosen points are sufficiently close to each other. Finally, we define the final estimate as the weighted average of the estimates at $s_1, \ldots, s_r$.

Notice that in this paper we have considered a method for calculating the codifference and the normalized codifference "directly" from the data. As an obvious alternative, once one knows the estimated parameters and the orders of the estimated models, one may directly estimate the codifference and the normalized codifference using equation (6). The methods for estimating the parameters of stable ARMA models have been reviewed in, e.g., Embrechts et al. (1997), Chapter 7. Notice that for small order MA and AR processes, the tail index α can be well estimated using a quantile based estimator (i.e McCulloch's method), see, e.g., Adler et al. (1998) for simulation evidences. In our opinion, for identification purpose (see, e.g., Rosadi 2006, 2007) or other purposes, the "direct" estimation method is more preferable than estimating the codifference function via estimated parameters.

Acknowledgments We wish to thank D. Bauer for various discussions and helpful suggestions. The financial support for D. Rosadi from the Ministry of Science and Technology of Austria via ÖAD TSA Project and WCRU Program 2009 in UGM, Indonesia is gratefully acknowledged.

Appendix A: Proof of Theorem 1

Before we give the lemmas which are necessary for the consistency proof of the codifference estimator, a related result from Kokoszka and Taqqu (1994) will be presented, which shows the codifference function in ARMA case is bounded by an exponentially decaying function just like the covariance function in the classical case. Kokoszka and Taqqu (1994) consider more general definition of the codifference function (for $\theta_1, \theta_2 \in \mathbb{R}$)

$$\tau_G(\theta_1, \theta_2; k) = -\ln \mathrm{E} \exp(i(\theta_1 X_{t+k} + \theta_2 X_t)) + \ln \mathrm{E} \exp(i\theta_1 X_{t+k}) + \ln \mathrm{E} \exp(i\theta_2 X_t) \tag{18}$$

but contain (4) as the special case ($\theta_1 = s, \theta_2 = -s$).

Theorem 3 (Kokoszka and Taqqu (1994), Theorem 2.1.) *If the coefficients c_j's of the linear process (1), satisfying conditions C1 and $\{\epsilon_t\}$ satisfying C2 then ($\theta_1, \theta_2 \in \mathbb{R}$)*

$$\limsup_{k \to \infty} Q^{\alpha k} |\tau_G(k)| \leq 2(1 - Q^\alpha)^{-1/\alpha} |\theta_1|^\alpha \quad \text{for } 0 < \alpha \leq 1 \tag{19}$$

and

$$\limsup_{k \to \infty} Q^k |\tau_G(k)| \leq \alpha \left(\sum_{j=0}^{\infty} |c_j|^\alpha \right)^{\frac{\alpha-1}{\alpha}} (1 - Q^\alpha)^{-1/\alpha} |\theta_1| |\theta_2|^{\alpha-1} \quad \text{for } 1 < \alpha \leq 2 \tag{20}$$

To show consistency of the codifference estimator, the following two lemmas are necessary.

Lemma 1 *Let $X_t, t \in \mathbb{Z}$ be the stationary linear process (1), satisfying conditions C1 and C2, and let $\Phi(s) = \mathrm{E} \exp(isX_t)$ denote its first order characteristic function. For $k \in \{0, 1, 2 \ldots\}$ and $s \in \mathbb{R}, s \neq 0$*

$$\ln \phi(s, k) = \ln \left((N - k)^{-1} \sum_{t=1}^{N-k} \exp(isX_t) \right)$$

is a consistent estimator of $\ln \Phi(s)$.

Proof Let $y_t = \exp(isX_t)$. Apparently, the magnitude of y_t is equal to one, and therefore it is a second order stationary process. For the notation simplicity, instead of working with $\phi(s, k)$, we first show consistency of $\phi^*(s, k) = N^{-1} \sum_{t=1}^{N} \exp(isX_t)$.

Here, $\phi^*(s,k)$ is an unbiased estimator for $\Phi(s) = E(y_t)$. To show the weak consistency of this estimator, we show that y_t is a mean ergodic process. A sufficient condition for y_t to be mean ergodic, i.e., $\phi^*(s,k) \to E(y_t)$ in the mean square sense, is that its covariance function tends to zero as time lags tends to ∞ (e.g., Theorem 7.1.1 in Brockwell and Davis 1987). The covariance function of y_t at lag k can be expressed as

$$\begin{aligned} c(k) &= |\Phi(s)|^2 \left(\frac{\mathrm{E}(\exp(is(X_{t+k} - X_t)))}{\mathrm{E}(\exp(isX_{t+k}))\mathrm{E}(\exp(-isX_t))} - 1 \right) \\ &= |\Phi(s)|^2 \left(\exp(-\tau(k)) - 1\right) \end{aligned} \tag{21}$$

From Theorem 3, we see that $c(k) \to 0$ when $k \to \infty$ exponentially fast. As mean square convergence entails convergence in probability, $\phi^*(s,k) \xrightarrow{p} \Phi(s)$. Moreover, under assumptions C1 and C2, we have $\Phi(s) = \exp(-\sum_{j=0}^{\infty} \sigma^\alpha \left|sc_j\right|^\alpha)$, a real-valued function. Therefore we can conclude $\mathrm{Re}\,\phi^*(s,k) \xrightarrow{p} \mathrm{Re}\,\Phi(s) = \Phi(s)$ and $\mathrm{Im}\,\phi^*(s,k) \xrightarrow{p} \mathrm{Im}\,\Phi(s) = 0$.

By taking the principal value of $\ln(\cdot)$ function in the complex domain, we can see that $\ln(\cdot)$ is a continuous and well-defined function on $\mathbb{C}$ minus the negative real line. Because $|c_j| < cQ^{-j}$ for some $c > 0$, $Q > 1$, we conclude $\mathrm{Re}\,\Phi(s)$ always strictly greater than 0, which implies with the probability converging to 0, $\mathrm{Re}\,\phi^*(s,k)$ will be less than or equal to 0. Therefore, without loss of generality, we can restrict the definition of the real and imaginary parts of $\ln\phi^*(s,k)$ only on the right half plane where $\mathrm{Re}\,\phi^*(s,k) > 0$, and equal to 0 on the other case. From this consideration, we obtain $\mathrm{Re}\ln\phi^*(s,k) = \frac{1}{2}\ln((\mathrm{Re}\,\phi^*(s,k))^2 + (\mathrm{Im}\,\phi^*(s,k))^2)$ and $\mathrm{Im}\ln\phi^*(s,k) = \arctan(\frac{\mathrm{Im}\,\phi^*(s,k)}{\mathrm{Re}\,\phi^*(s,k)})$. From the continuity of the logarithm function in the considered domain, we can deduce that $\mathrm{Re}\ln\phi^*(s,k) \xrightarrow{p} \mathrm{Re}\ln\Phi(s) = \ln\Phi(s)$ and $\mathrm{Im}\ln\phi^*(s,k) = \arg\phi^*(s,k) \xrightarrow{p} 0$, when $N \to \infty$. In other words, we obtain $\ln\phi^*(s,k) \xrightarrow{p} \ln\Phi(s)$. To complete our proof, it is sufficient to show $\phi^*(s,k) - \phi(s,k) \xrightarrow{p} 0$. By assumption of the model, $\mathrm{Re}\,\Phi(s) > 0$, thus $E\,|\mathrm{Re}\,\phi^*(s,k) - \mathrm{Re}\,\phi(s,k)| < 2\frac{k}{N-k}$ and $E\,|\mathrm{Im}\,\phi^*(s,k) - \mathrm{Im}\,\phi(s,k)| < 2\frac{k}{N-k}$, and therefore we can conclude $\phi^*(s,k) - \phi(s,k) = o_p(1)$. □

Lemma 2 *Let $X_t, t \in \mathbb{Z}$ be the stationary linear process (1), satisfying conditions C1 and C2, and for $k \in \{0,1,2,\ldots\}$ and $s \in \mathbb{R}, s \neq 0$, let $\Phi(s,-s;k) = \mathrm{E}\exp(is(X_{t+k} - X_t))$ be its second-order characteristic function evaluated at $(s,-s)$. Then as $N \to \infty$*

$$\ln\phi(s,-s;k) \xrightarrow{p} \ln\Phi(s,-s;k)$$

where $\phi(s,-s;k)$ is as given in (11).

Proof For the proof, we can proceed in a similar way as the previous lemma. For simplicity, instead of working with $\phi(s,-s;k)$, we first show the consistency of $\phi^*(s,-s;k) = N^{-1}\sum_{t=1}^{N}\exp is(X_{t+k} - X_t)$. A sufficient condition for y_t to be autocovariance ergodic (Proakis and Manolakis 1996, p.A10), i.e., $\phi^*(s,-s,k) \to$

$\Phi(s,-s;k)$, in the mean square sense is that

$$\mathrm{E}\exp(is(X_t - X_{t+k} - X_{t+n} + X_{t+n+k})) \to |\Phi(s,-s;k)|^2$$

as $n \to \infty$ where the index n denotes the lag of covariance among the sample autocovariance function. Hence, we have

$$\begin{aligned}
&\mathrm{E}\exp(is(X_t - X_{t+k} - X_{t+n} + X_{t+n+k})) \\
&= |\Phi(s,-s;k)|^2 \frac{\mathrm{E}\exp(is((X_t - X_{t+k}) - (X_{t+n} - X_{t+n+k})))}{\mathrm{E}\exp(is(X_t - X_{t+k}))\mathrm{E}\exp(is(X_{t+n+k} - X_{t+n}))} \\
&= |\Phi(s,-s;k)|^2 \exp(-C_n)
\end{aligned}$$

where

$$C_n = -\ln \mathrm{E}\exp(is((X_t - X_{t+k}) - (X_{t+n} - X_{t+n+k}))) \tag{22}$$

$$+ \ln \mathrm{E}\exp(is(X_t - X_{t+k})) + \ln \mathrm{E}\exp(is(X_{t+n+k} - X_{t+n})) \tag{23}$$

Applying the similar technique as obtaining (6), one can write C_n as

$$\begin{aligned}
C_n &= \sigma^\alpha \left[\sum_{j=0}^{\infty} |s(c_j - c_{j+k} - c_{j+n} + c_{j+n+k})|^\alpha - |s(c_{j+n+k} - c_{j+n})|^\alpha \right. \\
&\qquad \left. - |s(c_j - c_{j+k})|^\alpha \right] \\
&= \sigma^\alpha \left[\sum_{j=0}^{\infty} |s(k_j - k_{j+n})|^\alpha - |-sk_{j+n}|^\alpha - |sk_j|^\alpha \right]
\end{aligned}$$

where $k_j = c_j - c_{j+k}$. This expression is the codifference function $\tau_G(n)$ for coefficients k_j's and parameters $\theta_1 = -s$, $\theta_2 = s$. Because $|c_j| < cQ^{-j}$ for some $c > 0$, $Q > 1$, then $|k_j| < c_1 Q^{-j}$ for some $c_1 = 2c > 0$, $Q > 1$. Therefore, by (20) and (19), we can conclude that $\exp(-C_n)$ will converge to 1 exponentially fast. In other words, $\mathrm{E}\exp(is(X_t - X_{t+k} - X_{t+n} + X_{t+n+k})) \to |\Phi(s,-s;k)|^2$ for $n \to \infty$, and we obtain the mean square convergence of $\phi^*(s,-s;k)$ to $\Phi(s,-s;k)$ and therefore $\phi^*(s,-s;k) \xrightarrow{p} \Phi(s,-s;k)$. For the rest of the proof, we can proceed similarly to the proof of previous lemma, as we have $\Phi(s,-s;k) = \exp(-\sigma^\alpha(\sum_{j=0}^{k-1} |sc_j|^\alpha - \sum_{j=0}^{\infty} |s(c_{j+k} - c_j)|^\alpha))$ also a real-valued function, strictly greater than 0. □

Proof (Proof of Theorem 1) As for finite k and $N \to \infty$ we obtain $\sqrt{1 - k/N} \to 1$, then using the results in lemma 1 and lemma 2, we have as $N \to \infty$, for $i = 1, \ldots, r$

$$\hat{\tau}(s_i,-s_i;k) \xrightarrow{p} -\ln \Phi(s_i,-s_i;k) + \ln \Phi(s_i) + \ln \Phi(-s_i) = \tau(s_i,-s_i;k) \tag{24}$$

□

Appendix B: The limit distribution of the sample codifference function

In this part, we will derive the asymptotic distribution of the sample codifference function of linear processes. The proof will be given as a series of propositions, where the main results are presented in Theorem 4 and also the proof of Theorem 2 at the end of this part. The proof will follow closely an approach for obtaining the limiting distribution of the sample ACF in the classical case, e.g., Theorem 7.2.1 in Brockwell and Davis (1987).

For notational simplicity, instead of working with $\hat{\tau}(s_i, -s_i; k)$, $i = 1, \ldots, r$, in the following first we will consider the similar estimator $\hat{\tau}^*(s_i, -s_i; k)$,

$$\hat{\tau}^*(s_i, -s_i; k) = -\ln \phi^*(s_i, -s_i; k) + \ln \phi^*(s_i, 0; k) + \ln \phi^*(0, -s_i; k) \tag{25}$$

where $\phi^*(u, v; k) = N^{-1} \sum_{t=1}^{N} \exp(i(uX_{t+k} + vX_t))$, $u, v \in \mathbb{R}$. The required result will be presented in Theorem 4.

Proposition 1 *Let $X_t, t \in \mathbb{Z}$ be the stationary linear process (1), satisfying conditions C1 and C2. Then if $p \geq 0$ and $q \geq 0$,*

$$\lim_{N\to\infty} N\mathrm{cov}\left(\begin{pmatrix} \mathrm{Re}\,\hat{\tau}^*(\mathbf{s}, p) \\ \mathrm{Im}\,\hat{\tau}^*(\mathbf{s}, p) \end{pmatrix}, \begin{pmatrix} \mathrm{Re}\,\hat{\tau}^*(\mathbf{s}, q) \\ \mathrm{Im}\,\hat{\tau}^*(\mathbf{s}, q) \end{pmatrix}\right) = \lambda \mathbf{L}_2^p \mathbf{V}_{pq} \mathbf{L}_2^q \lambda^T$$

where the matrices $\lambda, \mathbf{L}_2^k, k = p, q$ and $\mathbf{V}_{pq}$ are given in (27), (34) and (36) below. Here $\mathrm{cov}(X, Y)$ *denotes the covariance between X and Y.*

Proof To obtain a complete variance-covariance structure of the estimator, we consider the following representation of $\hat{\tau}^*(\mathbf{s}, k)$

$$\begin{pmatrix} \mathrm{Re}\,\hat{\tau}^*(\mathbf{s}, k) \\ \mathrm{Im}\,\hat{\tau}^*(\mathbf{s}, k) \end{pmatrix} = \begin{pmatrix} \mathrm{Re}\,\hat{\tau}^*(s_1, -s_1, k) \\ \mathrm{Re}\,\hat{\tau}^*(s_2, -s_2, k) \\ \vdots \\ \mathrm{Re}\,\hat{\tau}^*(s_r, -s_r, k) \\ \mathrm{Im}\,\hat{\tau}^*(s_1, -s_1, k) \\ \mathrm{Im}\,\hat{\tau}^*(s_2, -s_2, k) \\ \vdots \\ \mathrm{Im}\,\hat{\tau}^*(s_r, -s_r, k) \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{Y} \\ \mathbf{X} \end{pmatrix} \tag{26}$$

where

$$\lambda = \begin{pmatrix} \mathbf{I}_r \otimes \lambda_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_r \otimes \lambda_1 \end{pmatrix} \tag{27}$$
$$\lambda_1 = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$$

and

$$\mathbf{Y} = \begin{pmatrix} \operatorname{Re}\ln Y_1^k \\ \operatorname{Re}\ln Y_2^k \\ \vdots \\ \operatorname{Re}\ln Y_r^k \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} \operatorname{Im}\ln Y_1^k \\ \operatorname{Im}\ln Y_2^k \\ \vdots \\ \operatorname{Im}\ln Y_r^k \end{pmatrix}$$

Here $\mathbf{I}_r$ denotes the matrix identity of size r, where we denote

$$Y_i^k = \begin{pmatrix} \phi^*(0, -s_i; k) \\ \phi^*(s_i, 0; k) \\ \phi^*(s_i, -s_i; k) \end{pmatrix} = \begin{pmatrix} \phi_1(s_i, k) \\ \phi_2(s_i, k) \\ \phi_3(s_i, k) \end{pmatrix}$$

and the logarithm function is defined componentwise, i.e., we have

$$\operatorname{Re}\ln Y_i^k = \begin{pmatrix} \operatorname{Re}\ln\phi_1(s_i, k) \\ \operatorname{Re}\ln\phi_2(s_i, k) \\ \operatorname{Re}\ln\phi_3(s_i, k) \end{pmatrix}$$

and similarly for the imaginary part. Let us denote

$$\mathrm{E}Y_i^k = \begin{pmatrix} \mathrm{E}\phi_1(s_i, k) \\ \mathrm{E}\phi_2(s_i, k) \\ \mathrm{E}\phi_3(s_i, k) \end{pmatrix} = \begin{pmatrix} \Phi_1(s_i, k) \\ \Phi_2(s_i, k) \\ \Phi_3(s_i, k) \end{pmatrix}$$

Notice that $\Phi(u, v; k) = \mathrm{E}(\exp(i(uX_{t+k} + vX_t)))$, $u, v \in \mathbb{R}$. Using mean value theorem, we can expand the codifference function into

$$\begin{pmatrix} \operatorname{Re}\hat{\tau}^*(\mathbf{s}, k) \\ \operatorname{Im}\hat{\tau}^*(\mathbf{s}, k) \end{pmatrix} = \lambda \left\{ \mathbf{L}_1^k + \bar{\mathbf{L}}_2^k \mathbf{Z}_N^k \right\} \tag{28}$$

where

$$\mathbf{L}_1^k = \begin{pmatrix} \operatorname{Re}\mathbf{L}_1^k \\ \operatorname{Im}\mathbf{L}_1^k \end{pmatrix}, \quad \mathbf{Z}_N^k = \begin{pmatrix} \operatorname{Re}\mathbf{Z}_N^k \\ \operatorname{Im}\mathbf{Z}_N^k \end{pmatrix} = \begin{pmatrix} \operatorname{Re}\varphi_N^k - \operatorname{Re}\psi_N^k \\ \operatorname{Im}\varphi_N^k - \operatorname{Re}\psi_N^k \end{pmatrix}$$

with

$$\operatorname{Re}\mathbf{L}_1^k = \begin{pmatrix} \operatorname{Re}\ln \mathrm{E}Y_1^k \\ \operatorname{Re}\ln \mathrm{E}Y_2^k \\ \vdots \\ \operatorname{Re}\ln \mathrm{E}Y_r^k \end{pmatrix}, \quad \operatorname{Re}\varphi_N^k = \begin{pmatrix} \operatorname{Re}Y_1^k \\ \operatorname{Re}Y_2^k \\ \vdots \\ \operatorname{Re}Y_r^k \end{pmatrix}, \quad \operatorname{Re}\psi_N^k = \begin{pmatrix} \operatorname{Re}\mathrm{E}Y_1^k \\ \operatorname{Re}\mathrm{E}Y_2^k \\ \vdots \\ \operatorname{Re}\mathrm{E}Y_r^k \end{pmatrix}$$

and similarly for the imaginary parts, and where and $\bar{\mathbf{L}}_2^k = (\bar{d}_{ij}^k)_{i,j=1,\ldots,6}$ denotes Jacobian of (26), which is evaluated at $\mathbf{c}(\left\|\mathbf{c} - \psi_N^k\right\| < \left\|\varphi_N^k - \psi_N^k\right\|)$. From the assumption C2, we obtain

$$\Phi_3(s_i, k) = \Phi(s_i, -s_i; k) = \exp\left(-\sum_{j=0}^{k-1} \sigma^\alpha \left|s_i c_j\right|^\alpha - \sum_{j=0}^{\infty} \sigma^\alpha \left|s_i (c_{j+k} - c_j)\right|^\alpha\right) \tag{29}$$

and $\Phi_1(s_i, k) = \Phi_2(s_i, k)$, i.e.,

$$\Phi(s_i, 0; k) = \Phi(0, -s_i; k) = \exp\left(-\sum_{j=0}^{\infty} \sigma^\alpha \left|s_i c_j\right|^\alpha\right) \tag{30}$$

From identities (29)–(30) and further applying the assumption C1, we obtain that the elements of $\operatorname{Re} \psi_N^k$ are always strictly greater than 0. Therefore, with a probability convergent to 0, the elements of $\operatorname{Re} \varphi_N^k$ will be less than or equal to 0. Hence, without changing the limiting distribution of the estimator, we can restrict the definition of the real and the imaginary components of $\begin{pmatrix} \mathbf{Y} \\ \mathbf{X} \end{pmatrix}$ in (26) only in the right half plane where the elements of $\operatorname{Re}(\varphi_N^k) > 0$, and equal to 0 in the other case. Thus, we can conclude that the Jacobian matrix $\bar{\mathbf{L}}_2^k$ is well defined here. By Theorem 1, $\bar{\mathbf{L}}_2^k$ will converge in probability to $\mathbf{L}_2^k$, where

$$\mathbf{L}_2^k = \nabla \mathbf{L}_1^k$$

Here ∇g denotes the Jacobian of g. From (29), (30), we have the following identities

$$\operatorname{Re} \Phi(s_i, -s_i; k) = \mathrm{E} \cos(s_i (X_{t+k} - X_t)) = \Phi(s_i, -s_i, k) \tag{31}$$

$$\operatorname{Re} \Phi(s_i, 0; k) = \mathrm{E} \cos(s_i X_{t+k}) = \Phi(s_i, 0, k) \tag{32}$$

$$\operatorname{Re} \Phi(0, -s_i; k) = \mathrm{E} \cos(-s_i X_t) = \Phi(0, -s_i, k) \tag{33}$$

and $\operatorname{Im} \Phi(s_i, -s_i; k) = \mathrm{E} \sin(s_i (X_{t+k} - X_t)) = 0$, $\operatorname{Im} \Phi(s_i, 0; k) = \mathrm{E} \sin(s_i X_{t+k}) = 0$ and $\operatorname{Im} \Phi(0, -s_i; k) = \mathrm{E} \sin(-s_i X_t) = 0$. Using these identities, after some algebra we directly obtain

$$\mathbf{L}_2^k = \begin{pmatrix} \mathbf{I}_r \mathbf{d}^k & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_r \mathbf{d}^k \end{pmatrix} \tag{34}$$

where $(\mathbf{d}^k)^T = [\mathbf{d}_1^k, \mathbf{d}_2^k, \ldots, \mathbf{d}_r^k]$, and the elements of $\mathbf{d}_i^k, i = 1, \ldots r$ are

$$\begin{aligned} d_i^k(1,1) &= (\mathrm{Re}\,\Phi(0,-s_i;k))^{-1} \\ d_i^k(2,2) &= (\mathrm{Re}\,\Phi(s_i,0;k))^{-1} \\ d_i^k(3,3) &= (\mathrm{Re}\,\Phi(s_i,-s_i;k))^{-1}) \end{aligned}$$

and equal to 0, otherwise. The asymptotic variance-covariance matrix is obtained from (28) as

$$\lim_{N\to\infty} N\ \mathrm{cov}\left(\begin{pmatrix}\mathrm{Re}\,\hat{\tau}^*(s,-s;p)\\ \mathrm{Im}\,\hat{\tau}^*(s,-s;p)\end{pmatrix}, \begin{pmatrix}\mathrm{Re}\,\hat{\tau}^*(s,-s;q)\\ \mathrm{Im}\,\hat{\tau}^*(s,-s;q)\end{pmatrix}\right) = \lambda \mathbf{L}_2^p \mathbf{V}_{pq} \mathbf{L}_2^q \lambda^T \quad (35)$$

where

$$\mathbf{V}_{pq} = \begin{pmatrix}\mathbf{V}_{pq}^{RR} & \mathbf{V}_{pq}^{RI}\\ \mathbf{V}_{pq}^{IR} & \mathbf{V}_{pq}^{II}\end{pmatrix} = \lim_{N\to\infty} N \begin{pmatrix}\mathrm{cov}(\mathrm{Re}\,\mathbf{Z}_N^p, \mathrm{Re}\,\mathbf{Z}_N^q) & \mathrm{cov}(\mathrm{Re}\,\mathbf{Z}_N^p, \mathrm{Im}\,\mathbf{Z}_N^q)\\ \mathrm{cov}(\mathrm{Im}\,\mathbf{Z}_N^p, \mathrm{Re}\,\mathbf{Z}_N^q) & \mathrm{cov}(\mathrm{Im}\,\mathbf{Z}_N^p, \mathrm{Im}\,\mathbf{Z}_N^q)\end{pmatrix} \quad (36)$$

The matrix $\mathbf{V}_{pq}$ can be obtained by applying Theorem 1 and Remark 2.6 in Hesse (1990). Its elements can be derived in a similar way as obtaining variance-covariance matrix in Theorem 1 of Hesse (1990). This is possible, because it can be shown that all elements of $\mathbf{V}_{pq}$ (in the form of sum of the absolute components) are finite. Therefore, one can apply the property of the sample mean of ergodic processes (e.g., Theorem 7.1.1 in Brockwell and Davis 1987). Notice that here in particular, we obtain all elements of $\mathbf{V}_{pq}$ with respect to $\mathrm{cov}(\mathrm{Re}\,\mathbf{Z}_N^p, \mathrm{Im}\,\mathbf{Z}_N^q)$ and $\mathrm{cov}(\mathrm{Im}\,\mathbf{Z}_N^p, \mathrm{Re}\,\mathbf{Z}_N^q)$ are zeros. The elements of $\mathbf{V}_{pq}$ with respect to $\mathrm{cov}(\mathrm{Re}\,\mathbf{Z}_N^p, \mathrm{Re}\,\mathbf{Z}_N^q)$ and $\mathrm{cov}(\mathrm{Im}\,\mathbf{Z}_N^p, \mathrm{Im}\,\mathbf{Z}_N^q)$ can be shown to be finite using identities (29)–(30) and applying a similar approach as obtaining eq. (21) and (23), and further applying Theorem 3, or sometimes, eq.(2.7) in Kokoszka and Taqqu (1994) together with the similar steps as the proof of Theorem 3. However, we omit details. □

Proposition 2 *Let $X_t, t \in \mathbb{Z}$ be the moving average process of order m, $X_t = \sum_{j=0}^m c_j \epsilon_{t-j}$, satisfying conditions C1 and C2. Then for $h \in \{1,2,\ldots\}$, $s \in \mathbb{R}, s \neq 0$*

$$\left[\begin{pmatrix}\mathrm{Re}\,\hat{\tau}^*(\mathbf{s},0)\\ \mathrm{Im}\,\hat{\tau}^*(\mathbf{s},0)\end{pmatrix}, \ldots, \begin{pmatrix}\mathrm{Re}\,\hat{\tau}^*(\mathbf{s},h)\\ \mathrm{Im}\,\hat{\tau}^*(\mathbf{s},h)\end{pmatrix}\right]^T$$
$$\text{is } AN\left(\left[\begin{pmatrix}\tau(\mathbf{s},0)\\ \mathbf{0}\end{pmatrix}, \ldots, \begin{pmatrix}\tau(\mathbf{s},h)\\ \mathbf{0}\end{pmatrix}\right]^T, N^{-1}M\right)$$

where $\mathbf{M}$ *is the covariance matrix*

$$\mathbf{M} = \left[\lambda \mathbf{L}_2^p \mathbf{V}_{pq} \mathbf{L}_2^q \lambda^T\right]_{p,q=0,\ldots,h}$$

and the matrices $\lambda, \mathbf{L}_2^k, k = p, q$ *and* $\mathbf{V}_{pq}$ *are as given in Proposition 1 above.*

Proof To show this relation, define vectors $\{\mathbf{Y}_t\}$ by

$$\mathbf{Y}_t^T = (\mathbf{Z}_t, \mathbf{Z}_{t+1}, \ldots, \mathbf{Z}_{t+h})$$

where

$$\mathbf{Z}_{t+k} = \begin{pmatrix} \mathbf{X}_1^k \\ \mathbf{X}_2^k \\ \vdots \\ \mathbf{X}_r^k \end{pmatrix}$$

where for $j = 1, \ldots, r$

$$\mathbf{X}_j^k = \begin{pmatrix} \exp(-is_j X_t) \\ \exp(is_j X_{t+k}) \\ \exp(is_j (X_{t+k} - X_t)) \end{pmatrix}$$

By definition, $\{\mathbf{Z}_{t+k}\}$ is $m + k$-dependent sequence and therefore $\{\mathbf{Y}_t\}$ is $m + h$-dependent sequence. Next define

$$\zeta_t^T = (\xi_t, \xi_{t+1}, \ldots, \xi_{t+h})$$

where

$$\xi_{t+j} = \begin{pmatrix} \mathrm{Re}(\ln N^{-1} \sum_{t=1}^{N} Z_{t+j}) \\ \mathrm{Im}(\ln N^{-1} \sum_{t=1}^{N} Z_{t+j}) \end{pmatrix}$$

and

$$\mathrm{Re}(\ln N^{-1} \sum_{t=1}^{N} Z_{t+j}) = \begin{pmatrix} \mathrm{Re}(\ln N^{-1} \sum_{t=1}^{N} \mathbf{X}_1^j) \\ \mathrm{Re}(\ln N^{-1} \sum_{t=1}^{N} \mathbf{X}_2^j) \\ \vdots \\ \mathrm{Re}(\ln N^{-1} \sum_{t=1}^{N} \mathbf{X}_r^j)) \end{pmatrix}$$

where $l = 1, \ldots, r$

$$\mathrm{Re}(\ln N^{-1} \sum_{t=1}^{N} \mathbf{X}_l^j) = \begin{pmatrix} \mathrm{Re}(\ln N^{-1} \sum_{t=1}^{N} \exp(-i s_l X_t)) \\ \mathrm{Re}(\ln N^{-1} \sum_{t=1}^{N} \exp(i s_l X_{t+j}) \\ \mathrm{Re}(\ln N^{-1} \sum_{t=1}^{N} \exp(i s_l (X_{t+j} - X_t))) \end{pmatrix}$$

(similarly for the imaginary part. Note that the summation and the principal value of $\ln(\cdot)$ are defined component-wise), then we have

$$\begin{pmatrix} \mathrm{Re} \ln (N^{-1} \sum_{t=1}^{N} \mathbf{Y}_t) \\ \mathrm{Im} \ln (N^{-1} \sum_{t=1}^{N} \mathbf{Y}_t) \end{pmatrix} \lambda^T = \zeta_t \lambda^T = \left[\begin{pmatrix} \mathrm{Re}\, \hat{\tau}^*(\mathbf{s}, 0) \\ \mathrm{Im}\, \hat{\tau}^*(\mathbf{s}, 0) \end{pmatrix}, \ldots, \begin{pmatrix} \mathrm{Re}\, \hat{\tau}^*(\mathbf{s}, h) \\ \mathrm{Im}\, \hat{\tau}^*(\mathbf{s}, h) \end{pmatrix} \right]^T$$

where λ is as given in (27). We therefore need to show that when $N \to \infty$

$$\mathbf{a}^T (\zeta_t \lambda^T)^T \text{ is } AN \left(\mathbf{a}^T \left(\begin{pmatrix} \mathrm{Re}\, \tau(\mathbf{s}, 0) \\ \mathbf{0} \end{pmatrix}, \ldots, \begin{pmatrix} \mathrm{Re}\, \tau(\mathbf{s}, h) \\ \mathbf{0} \end{pmatrix} \right)^T, N^{-1} \mathbf{a}^T \mathbf{M} \mathbf{a} \right) \quad (37)$$

for all vectors $\mathbf{a} = (a_0, \ldots, a_h)^T \in \mathbb{R}^{h+1}$ such that $\mathbf{a}^T \mathbf{M} \mathbf{a} > 0$. For any such $\mathbf{a}$, the sequence $\{\mathbf{a}^T (\zeta_t \lambda^T)^T\}$ is $(m + h)$-dependent and since by Proposition 1

$$\lim_{N \to \infty} N \ \mathrm{var}(\mathbf{a}^T (\zeta_t \lambda^T)^T) = \mathbf{a}^T \mathbf{M} \mathbf{a} > 0$$

where $\mathbf{M}$ is the covariance matrix

$$\mathbf{M} = \left[\lambda \mathbf{L}_2^p \mathbf{V}_{pq} \mathbf{L}_2^q \lambda^T \right]_{p,q=0,\ldots,h}$$

and the vectors $\lambda, \mathbf{L}_2^p, \mathbf{L}_2^q$, matrix $\mathbf{V}_{pq}$ are as given in Proposition 1 above. We can conclude that $\{\mathbf{a}^T (\zeta_t \lambda^T)^T\}$ satisfies the conditions of central limit theorem for m-dependent processes (e.g., Brockwell and Davis 1987 Theorem 6.4.2), and therefore by this theorem, for $N \to \infty$, we obtain the required result (37). The relation $\mathrm{Im}\, \tau(\mathbf{s}, j) = \mathbf{0}$, $j = 0, 1, \ldots, h$ can be obtained directly from identities (29)–(30).

□

Proposition 3 *Proposition 2 remains true for $X_t, t \in \mathbb{Z}$ being a stationary linear process (1), satisfying conditions* C1 *and* C2.

Proof For the proof, we will apply the result of Proposition 2 to the truncated sequence $X_{tm} = \sum_{j=0}^{m} c_j \epsilon_{t-j}$ and then derive the result for X_t by letting $m \to \infty$. For

$0 \le p \le h$, we define

$$\hat{\tau}_m^*(s,-s;\,p) = -\ln \phi_m^*(s,-s;\,p) + \ln \phi_m^*(s,0;\,p) + \ln \phi_m^*(0,-s;\,p) \tag{38}$$

where $\phi_m^*(u,v;\,p) = N^{-1}\sum_{t=1}^{N} \exp(i(uX_{(t+p)m} + vX_{tm}))$. Then by Proposition 2

$$N^{1/2}\left[\begin{pmatrix} \operatorname{Re}\hat{\tau}_m^*(\mathbf{s},0) - \operatorname{Re}\tau_m(\mathbf{s},0) \\ \operatorname{Im}\hat{\tau}_m^*(\mathbf{s},0) - \operatorname{Im}\tau_m(\mathbf{s},0)\end{pmatrix}, \ldots, \begin{pmatrix} \operatorname{Re}\hat{\tau}_m^*(\mathbf{s},h) - \operatorname{Re}\tau_m(\mathbf{s},h) \\ \operatorname{Im}\hat{\tau}_m^*(\mathbf{s},h) - \operatorname{Im}\tau_m(\mathbf{s},h)\end{pmatrix}\right]^T \Rightarrow Y_m$$

where $Y_m \sim N(0, M_m)$. Here $\mathbf{M}_m$ is the covariance matrix

$$\begin{aligned}\mathbf{M}_m &= \begin{pmatrix} \operatorname{cov}(\operatorname{Re}\hat{\tau}_m^*(\mathbf{s},p), \operatorname{Re}\hat{\tau}_m^*(\mathbf{s},q)) & \operatorname{cov}(\operatorname{Re}\hat{\tau}_m^*(\mathbf{s},p), \operatorname{Im}\hat{\tau}_m^*(\mathbf{s},q)) \\ \operatorname{cov}(\operatorname{Im}\hat{\tau}_m^*(\mathbf{s},p), \operatorname{Re}\hat{\tau}_m^*(\mathbf{s},q)) & \operatorname{cov}(\operatorname{Im}\hat{\tau}_m^*(\mathbf{s},p), \operatorname{Im}\hat{\tau}_m^*(\mathbf{s},q))\end{pmatrix}_{p,q=0,\ldots,h} \\ &= \left[\lambda \mathbf{L}_2^p(m)\mathbf{V}_{pq}^m \mathbf{L}_2^q(m)\lambda^T\right]_{p,q=0,\ldots,h}\end{aligned}$$

where λ is defined as (27) and the Jacobian matrix $\mathbf{L}_2^k(m)$ and matrix $\mathbf{V}_{pq}^m$ are defined for X_{tm} as in (34) and (36), respectively. Now, as $m \to \infty$,

$$\mathbf{M}_m \to \mathbf{M}$$

where $\mathbf{M}$ is defined like $\mathbf{M}_m$ by replacing X_{tm} by X_t. Hence

$$Y_m \Rightarrow Y \text{ where } Y \sim N(0, \mathbf{M})$$

The proof now can be completed by applying Proposition 6.3.9 in Brockwell and Davis (1987) provided we can show that

$$\begin{aligned}\lim_{m\to\infty} \limsup_{N\to\infty} \mathrm{P}(N^{1/2}|\operatorname{Re}\hat{\tau}_m^*(\mathbf{s},p) - \operatorname{Re}\tau_m(\mathbf{s},p) - \operatorname{Re}\hat{\tau}^*(\mathbf{s},p) \\ + \operatorname{Re}\tau(\mathbf{s},p)| > \epsilon) = 0\end{aligned} \tag{39}$$

for $p = 0, 1, \ldots, h$ (and similarly for the imaginary part). The probability in (39) is bounded by

$$\begin{aligned}&\epsilon^{-2} N \operatorname{var}(\operatorname{Re}\hat{\tau}_m^*(\mathbf{s},p) - \operatorname{Re}\hat{\tau}^*(\mathbf{s},p)) \\ &\quad = \epsilon^{-2}\left[N \operatorname{var}(\operatorname{Re}\hat{\tau}_m^*(\mathbf{s},p)) + N \operatorname{var}(\operatorname{Re}\hat{\tau}^*(\mathbf{s},p))\right. \\ &\qquad \left. -2N\operatorname{cov}(\operatorname{Re}\hat{\tau}_m^*(\mathbf{s},p), \operatorname{Re}\hat{\tau}^*(\mathbf{s},p))\right]\end{aligned}$$

From the calculation of Proposition 1 and further noting that Theorem 1 and Remark 2.6 in Hesse (1990) can be applied for the finite moving average process by setting some of the coefficients c_j's to be zero, we obtain

$$\lim_{m\to\infty}\lim_{N\to\infty} N \operatorname{var}(\operatorname{Re}\hat{\tau}_m^*(\mathbf{s},p)) = \lim_{N\to\infty} N \operatorname{var}(\operatorname{Re}\hat{\tau}^*(\mathbf{s},p)) = m_{pp}^{RR}$$

where m_{pq}^{RR} denotes the covariance between the real elements in (p, q)- block of covariance matrix $\mathbf{M}$. Moreover, using the same steps to that given in the proof of Proposition 1, it can be shown that

$$\lim_{m\to\infty}\lim_{N\to\infty} N\text{cov}(\text{Re}\,\hat{\tau}_m^*(\mathbf{s}, p), \text{Re}\,\hat{\tau}^*(\mathbf{s}, p) = m_{pp}^{RR}$$

Thus

$$\lim_{m\to\infty}\limsup_{N\to\infty} \epsilon^{-2} N\ \text{var}(\text{Re}\,\hat{\tau}_m^*(\mathbf{s}, p) - \text{Re}\,\hat{\tau}^*(\mathbf{s}, p)) = 0$$

Similar results can be obtained for the imaginary part. This established (39). □

Theorem 4 *Let $X_t, t \in \mathbb{Z}$ be the stationary linear process (1), satisfying conditions* C1 *and* C2. *Then for $h \in \{1, 2, \ldots\}$, $s \in \mathbb{R}, s \neq 0$*

$$\left[\begin{pmatrix}\text{Re}\,\hat{\tau}(\mathbf{s}, 0)\\ \text{Im}\,\hat{\tau}(\mathbf{s}, 0)\end{pmatrix}, \ldots, \begin{pmatrix}\text{Re}\,\hat{\tau}(\mathbf{s}, h)\\ \text{Im}\,\hat{\tau}(\mathbf{s}, h)\end{pmatrix}\right]^T \text{ is } AN\left(\left[\begin{pmatrix}\tau(\mathbf{s}, 0)\\ \mathbf{0}\end{pmatrix}, \ldots, \begin{pmatrix}\tau(\mathbf{s}, h)\\ \mathbf{0}\end{pmatrix}\right]^T, N^{-1}M\right)$$

where $\mathbf{M}$ *is as given in Proposition 2 above.*

Proof To show the convergence of the estimator $\text{Re}\,\hat{\tau}(\mathbf{s}, j)$ and $\text{Im}\,\hat{\tau}(\mathbf{s}, j)$ to the same limit as $\text{Re}\,\hat{\tau}^*(\mathbf{s}, j)$ and $\text{Im}\,\hat{\tau}^*(\mathbf{s}, j)$, respectively, with $0 \leq j \leq h$, it suffices to show that as $N \to \infty$

$$N^{1/2}\left\{\lambda_2\begin{pmatrix}\text{Re}\,\phi^*(s_k, -s_k; j)\\ \text{Re}\,\phi^*(s_k, 0; j)\\ \text{Re}\,\phi^*(0, -s_k; j)\end{pmatrix} - \lambda_2\begin{pmatrix}\text{Re}\,\phi(s_k, -s_k; j)\\ \text{Re}\,\phi(s_k, 0; j)\\ \text{Re}\,\phi(0, -s_k; j)\end{pmatrix}\right\} = o_p(1)$$

(and similarly for the imaginary part), where $\phi^*(u, v; j) = N^{-1}\sum_{t=1}^{N}\exp(i(uX_{t+j} + vX_t))$, $\phi(u, v; j) = (N-j)^{-1}\sum_{t=1}^{N-j}\exp(i(uX_{t+j} + vX_t))$ and $\lambda_2 = \begin{bmatrix}-1 & 1 & 1\end{bmatrix}$. The required result then follows from Slutzky's theorem (e.g., Theorem 5.1.1 in Lehmann 1999).

Simple algebra gives, for $0 \leq j \leq h$,

$$N^{1/2}\text{E}\left|\lambda_2\begin{pmatrix}\text{Re}\,\phi^*(s_k, -s_k; j)\\ \text{Re}\,\phi^*(s_k, 0; j)\\ \text{Re}\,\phi^*(0, -s_k; j)\end{pmatrix} - \lambda_2\begin{pmatrix}\text{Re}\,\phi(s_k, -s_k; j)\\ \text{Re}\,\phi(s_k, 0; j)\\ \text{Re}\,\phi(0, -s_k; j)\end{pmatrix}\right|$$

$$= N^{1/2}\text{E}\left|\lambda_2\begin{pmatrix}\frac{j}{(N-j)}\frac{1}{N}\sum_{t=1}^{N}\cos(is_k(X_{t+j} - X_t)) - \frac{1}{N-j}\sum_{t=N-j+1}^{N}\cos(is(X_{t+j} - X_t))\\ \frac{j}{(N-j)}\frac{1}{N}\sum_{t=1}^{N}\cos(is_k X_{t+j}) - \frac{1}{N-j}\sum_{t=N-j+1}^{N}\cos(is_k X_{t+j})\\ \frac{j}{(N-j)}\frac{1}{N}\sum_{t=1}^{N}\cos(-is_k X_t)) - \frac{1}{N-j}\sum_{t=N-j+1}^{N}\cos(-is_k X_t)\end{pmatrix}\right|$$

$$\leq 6j(N-j)^{-1/2}\left(\frac{N}{N-j}\right)^{1/2}$$

The required result is obtained from $3j(N-j)^{-1/2} \to 0$ and $N/(N-j) \to 1$ as $N \to \infty$. Using the same arguments, similar results can be obtained for the imaginary part. The conclusion of the theorem then follows from Proposition 3 above. □

Proof (Proof of Theorem 2) Let $\mathbf{g}(\cdot)$ be the function from $\mathbb{R}^{(h+1)\times 2r}$ to $\mathbb{R}^{h\times 2r}$ defined by

$$\mathbf{g}\left(\left[\begin{pmatrix}\hat{\tau}(\mathbf{s},0)\\ \mathbf{0}\end{pmatrix}, \begin{pmatrix}\mathrm{Re}\,\hat{\tau}(\mathbf{s},1)\\ \mathrm{Im}\,\hat{\tau}(\mathbf{s},1)\end{pmatrix}, \ldots, \begin{pmatrix}\mathrm{Re}\,\hat{\tau}(\mathbf{s},h)\\ \mathrm{Im}\,\hat{\tau}(\mathbf{s},h)\end{pmatrix}\right]^T\right) = \left[\begin{pmatrix}\mathrm{Re}\,\hat{I}(\mathbf{s},1)\\ \mathrm{Im}\,\hat{I}(\mathbf{s},1)\end{pmatrix}, \begin{pmatrix}\mathrm{Re}\,\hat{I}(\mathbf{s},2)\\ \mathrm{Im}\,\hat{I}(\mathbf{s},2)\end{pmatrix}, \ldots, \begin{pmatrix}\mathrm{Re}\,\hat{I}(\mathbf{s},h)\\ \mathrm{Im}\,\hat{I}(\mathbf{s},h)\end{pmatrix}\right]^T$$

where for $0 < j \leq h$ and $\hat{\tau}(0) \neq 0$, we have $\mathrm{Re}\,\hat{I}(s_i, -s_i; j) = \frac{\mathrm{Re}\,\hat{\tau}(s_i,-s_i;j)}{\hat{\tau}(s_i,-s_i;0)}$ and $\mathrm{Im}\,\hat{I}(s_i, -s_i; j) = \frac{\mathrm{Im}\,\hat{\tau}(s_i,-s_i;j)}{\hat{\tau}(s_i,-s_i;0)}$, for $i = 1, \ldots, r$. By applying delta method and Theorem 4 above, we can show that

$$\left[\begin{pmatrix}\mathrm{Re}\,\hat{I}(\mathbf{s},1)\\ \mathrm{Im}\,\hat{I}(\mathbf{s},1)\end{pmatrix}, \begin{pmatrix}\mathrm{Re}\,\hat{I}(\mathbf{s},2)\\ \mathrm{Im}\,\hat{I}(\mathbf{s},2)\end{pmatrix}, \ldots, \begin{pmatrix}\mathrm{Re}\,\hat{I}(\mathbf{s},h)\\ \mathrm{Im}\,\hat{I}(\mathbf{s},h)\end{pmatrix}\right]^T$$

is asymptotically normal distributed with mean

$$\mathbf{g}\left(\left[\begin{pmatrix}\tau(\mathbf{s},0)\\ \mathbf{0}\end{pmatrix}, \begin{pmatrix}\tau(\mathbf{s},1)\\ \mathbf{0}\end{pmatrix}, \ldots, \begin{pmatrix}\tau(\mathbf{s},h)\\ \mathbf{0}\end{pmatrix}\right]^T\right) = \left[\begin{pmatrix}l_r I(1)\\ l_r 0\end{pmatrix}, \begin{pmatrix}l_r I(2)\\ l_r 0\end{pmatrix}, \ldots, \begin{pmatrix}l_r I(h)\\ l_r 0\end{pmatrix}\right]^T$$

and variance $N^{-1}\mathbf{DMD^T}$. Here the matrix $\mathbf{M}$ is as given in Proposition 2, $\mathbf{D}$ is the Jacobian matrix of $\mathbf{g}(\cdot)$ and $l_r = [1, 1, \ldots, 1]^T \in \mathbb{R}^r$. To obtain the elements of matrix $\mathbf{D}$, we proceed as follows. First, note that the codifference function at lag 0 is a real-valued function. Therefore, for $0 \leq j \leq h$, and $\tau(0) \neq 0$, we obtain $\mathrm{Re}\,I(j) = \frac{\mathrm{Re}\,\tau(j)}{\tau(0)} = I(j)$ and $\mathrm{Im}\,I(j) = \frac{\mathrm{Im}\,\tau(j)}{\tau(0)} = 0$. It is straightforward to obtain the Jacobian matrix $\mathbf{D}$ as

$$\mathbf{D} = \begin{bmatrix} D_{11} & D_{12} & \mathbf{0} & \ldots & \mathbf{0} \\ D_{21} & \mathbf{0} & D_{23} & & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ D_{h1} & \mathbf{0} & \mathbf{0} & \ldots & D_{h(h+1)} \end{bmatrix} \tag{40}$$

where

$$D_{l1} = \begin{bmatrix} D_{l1}^{11} & \mathbf{0}_r \\ \mathbf{0}_r & \mathbf{0}_r \end{bmatrix} \tag{41}$$

and

$$D_{l(l+1)} = \begin{bmatrix} D^{11}_{l(l+1)} & \mathbf{0}_r \\ \mathbf{0}_r & D^{11}_{l(l+1)} \end{bmatrix} \tag{42}$$

for $l = 1, ..h$, where

$$D^{11}_{l1} = \mathbf{I}_r \left[\frac{-I(l)}{\tau(s_1, -s_1; 0)}, \frac{-I(l)}{\tau(s_2, -s_2; 0)}, \ldots, \frac{-I(l)}{\tau(s_r, -s_r; 0)} \right]^T$$

and

$$D^{11}_{l(l+1)} = \mathbf{I}_r \left[\frac{1}{\tau(s_1, -s_1; 0)}, \frac{1}{\tau(s_2, -s_2; 0)}, \ldots, \frac{1}{\tau(s_r, -s_r; 0)} \right]^T$$

Here $\mathbf{I}_r$ denotes the matrix identity of size r. Let's denote w_{ij}, for $i, j = 1, \ldots, h$, the (i, j)th block element of $\mathbf{DMD}^T$ and m_{ij}, for $i, j = 0, 1, \ldots, h$, the (i, j)th block element of $\mathbf{M}$. We find that

$$\begin{aligned} w_{ij} &= \begin{bmatrix} \mathrm{cov}(\mathrm{Re}\,\hat{I}(\mathbf{s}, i), \mathrm{Re}\,\hat{I}(\mathbf{s}, j)) & \mathrm{cov}(\mathrm{Re}\,\hat{I}(\mathbf{s}, i), \mathrm{Im}\,\hat{I}(\mathbf{s}, j)) \\ \mathrm{cov}(\mathrm{Im}\,\hat{I}(\mathbf{s}, i), \mathrm{Re}\,\hat{I}(\mathbf{s}, j)) & \mathrm{cov}(\mathrm{Im}\,\hat{I}(\mathbf{s}, i), \mathrm{Im}\,\hat{I}(\mathbf{s}, j)) \end{bmatrix} \\ &= D_{i1} m_{00} D_{j1} + D_{i(i+1)} m_{i0} D_{j1} + D_{i1} m_{0j} D_{j(j+1)} + D_{i(i+1)} m_{ij} D_{j(j+1)} \\ &= \begin{bmatrix} D^{11}_{i1} m^{RR}_{00} D^{11}_{j1} + D^{11}_{i(i+1)} m^{RR}_{i0} D^{11}_{j1} + D^{11}_{i1} m^{RR}_{0j} D^{11}_{j(j+1)} + D^{11}_{i(i+1)} m^{RR}_{ij} D^{11}_{j(j+1)} & \mathbf{0}_r \\ \mathbf{0}_r & D^{11}_{i(i+1)} m^{II}_{ij} D^{11}_{j(j+1)} \end{bmatrix} \end{aligned} \tag{43}$$

Here m^{RR}_{ij} and m^{II}_{ij} denote the partitions of m_{ij} which correspond to the real and the imaginary components, respectively. □

Appendix C: Proof of Corollary 1

Proof (Proof of Corollary 1) As MA(0) is a special case of the linear process (1), by applying Theorem 2, one can conclude the asymptotic normality of

$$\left[\begin{pmatrix} \mathrm{Re}\,\hat{I}(\mathbf{s}, 1) \\ \mathrm{Im}\,\hat{I}(\mathbf{s}, 1) \end{pmatrix}, \begin{pmatrix} \mathrm{Re}\,\hat{I}(\mathbf{s}, 2) \\ \mathrm{Im}\,\hat{I}(\mathbf{s}, 2) \end{pmatrix}, \ldots, \begin{pmatrix} \mathrm{Re}\,\hat{I}(\mathbf{s}, h) \\ \mathrm{Im}\,\hat{I}(\mathbf{s}, h) \end{pmatrix} \right]^T$$

for $h \in \{1, 2, \ldots\}$. The true codifference function of i.i.d. process X_t is

$$\begin{aligned} \tau(s, -s; k) &= -\ln \mathrm{E} \exp(is(X_{t+k} - X_t)) + \ln \mathrm{E} \exp(isX_{t+k}) + \ln \mathrm{E} \exp(-isX_t) \\ &= \begin{cases} -2\sigma^\alpha |s|^\alpha & \text{for } k = 0 \\ 0 & \text{for } k > 0 \end{cases} \end{aligned}$$

which enables us to conclude that the real and the imaginary parts of $I(k) = 0$ whenever $k > 0$.

From (43), we obtain that $w_{kk}, k > 0$ is reduced to

$$w_{kk} = D_{k(k+1)} m_{kk} D_{k(k+1)} \tag{44}$$

where matrix $D_{k(k+1)}$ is as given in (42), with

$$D^{11}_{l(l+1)} = \mathbf{I}_r \left[\frac{1}{-2\sigma^\alpha |s_1|^\alpha}, \frac{1}{-2\sigma^\alpha |s_2|^\alpha}, \ldots, \frac{1}{-2\sigma^\alpha |s_r|^\alpha} \right]^T$$

and where

$$m_{kk} = \lambda \mathbf{L}_2^k \mathbf{V}_{kk} \mathbf{L}_2^k \lambda^T \tag{45}$$

with λ is as given as (27), and the elements of the matrix $\mathbf{L}_2^k$ and the covariance matrix $\mathbf{V}_{kk}$ will be given below. Let us denote

$$\mathbf{V}^{RR}_{kk}(i, j) = [\mathrm{cov}(\mathrm{Re}\, \phi_p(s_i, k), \mathrm{Re}\, \phi_q(s_j, k)))]_{p,q=1,2,3}$$

and

$$\mathbf{V}^{II}_{kk}(i, j) = [\mathrm{cov}(\mathrm{Im}\, \phi_p(s_i, k), \mathrm{Im}\, \phi_q(s_j, k)))]_{p,q=1,2,3}$$

as the (i, j)th block elements of $\mathbf{V}^{RR}_{kk}$ and $\mathbf{V}^{II}_{kk}$, respectively. Using identities (31)–(33) (and the identities for imaginary part afterwards) in p. 23, we can obtain their components as follows

$$\begin{aligned}
\mathrm{cov}(\mathrm{Re}(\phi_1(s_i, p)), \mathrm{Re}(\phi_1(s_j, q))) &= \mathrm{cov}(\cos(-s_i X_t), \cos(-s_j X_t)) \\
&= \frac{1}{2}\{e^{-\sigma^\alpha |s_i + s_j|^\alpha} + e^{-\sigma^\alpha |s_i - s_j|^\alpha}\} \\
&\quad - e^{-\sigma^\alpha (|s_i|^\alpha + |s_j|^\alpha)} \\
\mathrm{cov}(\mathrm{Re}(\phi_1(s_i, p)), \mathrm{Re}(\phi_2(s_j, q))) &= \mathrm{cov}(\mathrm{Re}(\phi_2(s_i, p)), \mathrm{Re}(\phi_1(s_j, q))) \\
&= \mathrm{cov}(\mathrm{Re}(\phi_1(s_i, p)), \mathrm{Re}(\phi_1(s_j, q))) \\
\mathrm{cov}(\mathrm{Re}(\phi_2(s_i, p)), \mathrm{Re}(\phi_2(s_j, q))) &= \mathrm{cov}(\mathrm{Re}(\phi_1(s_i, p)), \mathrm{Re}(\phi_1(s_j, q))) \\
\mathrm{cov}(\mathrm{Re}(\phi_1(s_i, p)), \mathrm{Re}(\phi_3(s_j, q))) &= \mathrm{cov}(\cos(-s_i X_t), \cos(s_j (X_{t+q} - X_t))) \\
&\quad + \mathrm{cov}(\cos(-s_i X_{t+q}), \cos(s_j (X_{t+q} - X_t))) \\
&= e^{-\sigma^\alpha (|s_j|^\alpha + |s_i - s_j|^\alpha)} + e^{-\sigma^\alpha (|s_j|^\alpha + |s_i + s_j|^\alpha)} \\
&\quad - 2e^{-\sigma^\alpha (|s_i|^\alpha + |2s_j|^\alpha)}
\end{aligned}$$

$$\begin{aligned}
\operatorname{cov}(\operatorname{Re}(\phi_3(s_i,p)),\operatorname{Re}(\phi_1(s_j,q))) &= \operatorname{cov}(\cos(-s_jX_t),\cos(s_i(X_{t+p}-X_t)))\\
&\quad+\operatorname{cov}(\cos(-s_jX_{t+p}),\cos(s_i(X_{t+p}-X_t)))\\
&= e^{-\sigma^\alpha(|s_i|^\alpha+|s_i-s_j|^\alpha)}+e^{-\sigma^\alpha(|s_i|^\alpha+|s_i+s_j|^\alpha)}\\
&\quad-2e^{-\sigma^\alpha(|s_j|^\alpha+|2s_i|^\alpha)}\\
\operatorname{cov}(\operatorname{Re}(\phi_2(s_i,p)),\operatorname{Re}(\phi_3(s_j,q))) &= \operatorname{cov}(\cos(s_iX_{t+q}),\cos(s_j(X_{t+q}-X_t)))\\
&\quad+\operatorname{cov}(\cos(s_iX_{t+p}),\cos(s_j(X_{t+p+q}-X_{t+p})))\\
&= e^{-\sigma^\alpha(|s_j|^\alpha+|s_i-s_j|^\alpha)}+e^{-\sigma^\alpha(|s_j|^\alpha+|s_i+s_j|^\alpha)}\\
&\quad-2e^{-\sigma^\alpha(|s_i|^\alpha+|2s_j|^\alpha)}\\
\operatorname{cov}(\operatorname{Re}(\phi_3(s_i,k)),\operatorname{Re}(\phi_2(s_j,k))) &= \operatorname{cov}(\cos(s_jX_{t+k}),\cos(s_i(X_{t+k}-X_t)))\\
&\quad+\operatorname{cov}(\cos(s_jX_{t+k}),\cos(s_i(X_{t+2k}-X_{t+k})))\\
&= e^{-\sigma^\alpha(|s_i|^\alpha+|s_i-s_j|^\alpha)}+e^{-\sigma^\alpha(|s_i|^\alpha+|s_i+s_j|^\alpha)}\\
&\quad-2e^{-\sigma^\alpha(|s_j|^\alpha+|2s_i|^\alpha)}\\
\operatorname{cov}(\operatorname{Re}(\phi_3(s_i,p)),\operatorname{Re}(\phi_3(s_j,q))) &= \operatorname{cov}(\cos(s_i(X_{t+p}-X_t)),\cos(s_j(X_{t+q}-X_t)))\\
&\quad+\operatorname{cov}(\cos(s_i(X_{t+p+q}-X_{t+q})),\\
&\quad\cos(s_j(X_{t+q}-X_t)))\\
&\quad+\operatorname{cov}(\cos(s_i(X_{t+p}-X_t)),\\
&\quad\cos(s_j(X_{t++p+q}-X_{t+p})))+c_{\mathrm{Re}}^{pq}
\end{aligned}$$

where

$$c_{\mathrm{Re}}^{pq}=\begin{cases}0 & \text{if } p=q\\ \operatorname{cov}(\cos(s_i(X_{t+q}-X_{t+q-p})),\cos(s_j(X_{t+q}-X_t))) & \text{if } q>p\\ \operatorname{cov}(\cos(s_i(X_{t+p}-X_t)),\cos(s_j(X_{t+p}-X_{t+p-q}))) & \text{if } p>q\end{cases}$$

yielding for $p=q$

$$\begin{aligned}
\operatorname{cov}(\operatorname{Re}(\phi_3(s_i,p)),\operatorname{Re}(\phi_3(s_j,q))) &= \frac{1}{2}e^{-2\sigma^\alpha|s_i+s_j|^\alpha}+\frac{1}{2}e^{-2\sigma^\alpha|s_i-s_j|^\alpha}\\
&\quad-3e^{-\sigma^\alpha(2|s_i|^\alpha+|2s_j|^\alpha)}\\
&\quad+e^{-\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha+|s_i-s_j|^\alpha)}\\
&\quad+e^{-\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha+|s_i+s_j|^\alpha)}
\end{aligned}$$

and for $p\neq q$

$$\begin{aligned}
\operatorname{cov}(\operatorname{Re}(\phi_3(s_i,p)),\operatorname{Re}(\phi_3(s_j,q))) &= 2e^{-\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha+|s_i-s_j|^\alpha)}\\
&\quad+2e^{-\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha+|s_i+s_j|^\alpha)}\\
&\quad-4e^{-\sigma^\alpha(2|s_i|^\alpha+|2s_j|^\alpha)}
\end{aligned}$$

$$
\begin{aligned}
\operatorname{cov}(\operatorname{Im}(\phi_1(s_i,\,p)),\operatorname{Im}(\phi_1(s_j,q))) &= \operatorname{cov}(\sin(-s_iX_t),\sin(-s_jX_t))\\
&= \frac{1}{2}\{e^{-\sigma^\alpha|s_i-s_j|^\alpha} - e^{-\sigma^\alpha|s_i+s_j|^\alpha}\}\\
\operatorname{cov}(\operatorname{Im}(\phi_1(s_i,\,p)),\operatorname{Im}(\phi_2(s_j,q))) &= \operatorname{cov}(\operatorname{Im}(\phi_2(s_i,\,p)),\operatorname{Im}(\phi_1(s_j,q)))\\
&= -\operatorname{cov}(\operatorname{Im}(\phi_1(s_i,\,p)),\operatorname{Im}(\phi_1(s_j,q)))\\
\operatorname{cov}(\operatorname{Im}(\phi_2(s_i,\,p)),\operatorname{Im}(\phi_2(s_j,q))) &= \operatorname{cov}(\operatorname{Im}(\phi_1(s_i,\,p)),\operatorname{Im}(\phi_1(s_j,q)))\\
\operatorname{cov}(\operatorname{Im}(\phi_3(s_i,\,p)),\operatorname{Im}(\phi_3(s_j,q))) &= \operatorname{cov}(\sin(s_i(X_{t+p}-X_t)),\sin(s_j(X_{t+q}-X_t)))\\
&\quad+\operatorname{cov}(\sin(s_i(X_{t+p+q}-X_{t+q})),\\
&\quad\sin(s_j(X_{t+q}-X_t)))\\
&\quad+\operatorname{cov}(\sin(s_i(X_{t+p+q}-X_{t+p})),\\
&\quad\sin(s_j(X_{t+p}-X_t))) + c^{pq}_{\mathrm{Im}}
\end{aligned}
$$

where

$$
c^{pq}_{\mathrm{Im}} = \begin{cases} 0 & \text{if} \quad p=q\\ \operatorname{cov}(\sin(\,s_i(X_{t+q}-X_{t+q-p})),\sin(\,s_j(X_{t+q}-X_t))) & \text{if} \quad q>p\\ \operatorname{cov}(\sin(\,s_i(X_{t+p}-X_t)),\sin(\,s_j(X_{t+p}-X_{t+p-q}))) & \text{if} \quad p>q \end{cases}
$$

yielding for $p=q$

$$
\begin{aligned}
\operatorname{cov}(\operatorname{Im}(\phi_3(s_i,k)),\operatorname{Im}(\phi_3(s_j,k))) &= \frac{1}{2}e^{-2\sigma^\alpha|s_i-s_j|^\alpha}\\
&\quad-\frac{1}{2}e^{-2\sigma^\alpha|s_i+s_j|^\alpha}+e^{-\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha+|s_i+s_j|^\alpha)}\\
&\quad-e^{-\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha+|s_i-s_j|^\alpha)}
\end{aligned}
$$

and $\operatorname{cov}(\operatorname{Im}(\phi_3(s_i,k)),\operatorname{Im}(\phi_3(s_j,k))) = 0$ for $p\neq q$. The other elements are all zeros. The elements of $\mathbf{L}_2^k$ are as given in (34), where the elements of $\mathbf{d}_i^k$, $i=1,\ldots r$ are

$$
\begin{aligned}
d_i^k(1,1) &= (\operatorname{Re}\Phi(0,-s_i;k))^{-1} = e^{\sigma^\alpha|s_i|^\alpha}\\
d_i^k(2,2) &= (\operatorname{Re}\Phi(s_i,0;k))^{-1} = e^{\sigma^\alpha|s_i|^\alpha}\\
d_i^k(3,3) &= (\operatorname{Re}\Phi(s_i,-s_i;k))^{-1} = e^{2\sigma^\alpha|s_i|^\alpha})
\end{aligned}
$$

As from (45) we obtain

$$
m^{RR}_{kk} = \operatorname{cov}(\operatorname{Re}\hat{\tau}(\mathbf{s},k),\operatorname{Re}\hat{\tau}(\mathbf{s},k)) = (\mathbf{I}_r\otimes\lambda_1)d^kV^{RR}_{kk}d^k(\mathbf{I}_r\otimes\lambda_1^T)
$$

and

$$
m^{II}_{kk} = \operatorname{cov}(\operatorname{Im}\hat{\tau}(\mathbf{s},k),\operatorname{Im}\hat{\tau}(\mathbf{s},k)) = (\mathbf{I}_r\otimes\lambda_1)d^kV^{RR}_{kk}d^k(\mathbf{I}_r\otimes\lambda_1^T)
$$

then the (i,j)th element of m_{kk}^{RR} and m_{kk}^{II} is obtained from

$$m_{kk}^{RR}(i,\, j) = \lambda_1 d_i^k V_{kk}^{RR}(i,\, j) d_j^k \lambda_1^T$$

and

$$m_{kk}^{II}(i,\, j) = \lambda_1 d_i^k V_{kk}^{II}(i,\, j) d_j^k \lambda_1^T$$

which therefore after a simple algebra, we obtain

$$\begin{aligned} m_{kk}^{RR}(i,\, j) &= e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i-s_j|^\alpha)}\left\{\frac{1}{2}e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i-s_j|^\alpha)}-1\right\} \\ &\quad + e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i+s_j|^\alpha)}\left\{\frac{1}{2}e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i+s_j|^\alpha)}-1\right\}+1 \\ m_{kk}^{II}(i,\, j) &= e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i-s_j|^\alpha)}\left\{\frac{1}{2}e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i-s_j|^\alpha)}-1\right\} \\ &\quad + e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i+s_j|^\alpha)}\left\{1-\frac{1}{2}e^{\sigma^\alpha(|s_i|^\alpha+|s_j|^\alpha-|s_i+s_j|^\alpha)}\right\} \end{aligned}$$

The required result follows directly from (44). □

References

Adler RJ, Feldman RE, Gallagher CM (1998) Analyzing stable time series. In: Adler RJ, Feldman RE, Taqqu MS (eds) A practical guide to eavy tails: statistical techniques and applications. Birkhäuser, Boston, pp 133–158

Bhansali RJ (1983) Estimation for the order of MA model from autoregressive and windows estimates of the inverse correlation function. J Time Ser Anal 4:137–162

Brockwell PJ, Davis RA (1987) Time series: theory and methods. Springer, New York

Chambers JM, Mallows CL, Stuck BW (1976) A method for simulating stable random variables. J Am Stat Assoc 71:340–344

Embrechts P, Klüppelberg C, Mikosch T (1997) Modelling extremal events for insurance and finance. Springer, Berlin

Hesse CH (1990) Rates of convergence for the empirical distribution function and the empirical characteristic function of a broad class of linear process. J Multivar Anal 35:186–202

Hong Y (1999) Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach. J Am Stat Assoc 94(448):1201–1214

Janicki A, Weron A (1994) Simulation and chaotic behavior of α-stable stochastic processes. Marcel Dekker, New York

Kogon SM, Williams DB (1998) Characteristic function based estimation of stable distribution parameters. In: Adler RJ, Feldman RE, Taqqu MS (eds) Practical guide to heavy tails: statistical techniques and applications. Birkhäuser, Boston, pp 311–335

Kokoszka PS, Taqqu MS (1994) Infinite variance stable ARMA processes. J Time Ser Anal 15:203–220

Koutrouvelis IA (1980) A goodness-of-fit test of simple hypotheses based on the empirical characteristic function. Biometrika 67:238–240

Lehmann E (1999) Elements of large-sample theory. Springer, New York

Nikias CL, Shao M (1995) Signal processing with α-stable distributions and applications. Wiley, New York

Nolan JP (1999a) Numerical calculation of stable densities and distribution functions. Technical report. Department of Mathematics and Statistics, American University

Nolan JP (1999b) Fitting data and assessing goodness-of-fit with stable distributions. Technical report. Department of Mathematics and Statistics, American University

Nowicka J (1997) Asymptotic behavior of the covariation and the codifference for ARMA models with stable innovations. Stoch Models 13:673–686

Nowicka J, Weron A (1997) Measures of dependence for ARMA models with stable innovations. Annales UMCS Sect A (Mathematica) 51:133–144

Paolella MS (2001) Testing the stable Paretian assumption. Math Comput Model 34:1095–1112

Proakis JG, Manolakis DG (1996) Digital signal processing. Prentice Hall, London

R Development Core Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. ISBN 3-900051-00-3

Rachev ST, Mittnik S (2000) Stable paretian models in finance. Wiley, New York

Rosadi D (2005) Asymptotic behavior of the codifference and the dynamical function for ARMA models with infinite variance. J Ind Math Soc 59–69

Rosadi D (2006) Order identification for gaussian moving averages using the codifference function. J Stat Comput Simul 76:553–559

Rosadi D (2007) Identification of moving average process with infinite variance. Stat Probab Lett 77(14):1490–1496

Samorodnitsky G, Taqqu MS (1994) Stable nonGaussian processes: stochastic models with infinite variance. Chapman and Hall, New York

Yang Q, Petropulu A, Pesquet JC (2001) Estimating long-range dependence in impulsive traffic flows. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp 3413–3416

Yu J (2004) Empirical characteristic function estimation and its applications. Econ Rev 23:93–123