Journal of Multivariate Analysis 172 (2019) 122–146

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

# Model assessment for time series dynamics using copula spectral densities: A graphical tool

Stefan Birr [a], Tobias Kley [b], Stanislav Volgushev [c,*]

[a] Ruhr Universität Bochum, Fakultät für Mathematik, Lehrstuhl für Stochastik, 44780 Bochum, Germany
[b] University of Bristol, School of Mathematics (Faculty of Science), University Walk, Bristol BS8 1TW, United Kingdom
[c] Department of Statistical Sciences, University of Toronto, 100 St. George Street, Toronto, Ontario, Canada M5S 3G3

ARTICLE INFO

*Article history:*
Received 4 April 2018
Available online 18 March 2019

*AMS subject classifications:*
62M10
62M15
62F40

*Keywords:*
Bootstrap
Copula
Frequency domain
Spectral density
Time series

ABSTRACT

Finding parametric models that accurately describe the dependence structure of observed data is a central task in the analysis of time series. Classical frequency domain methods provide a popular set of tools for fitting and diagnostics of time series models, but their applicability is seriously impacted by the limitations of covariances as a measure of dependence. Motivated by recent developments of frequency domain methods that are based on copulas instead of covariances, we propose a novel graphical tool to assess the quality of time series models for describing dependencies that go beyond linearity. We provide a theoretical justification of our approach and show in simulations that it can successfully distinguish between subtle differences in time series dynamics, including non-linear dynamics which result from GARCH and EGARCH models. We also demonstrate the utility of the proposed tools through an application to modeling returns of the S&P 500 stock market index.



## 1. Introduction

Nonparametric methods provide valuable tools for dependence modeling. If a parametric candidate model is available, we can compare the corresponding estimate with a nonparametric one to evaluate how well the chosen model describes the data. If no candidate model is available, nonparametric techniques can be used to get a first impression of the underlying dependence and information about potentially suitable parametric models.

In time series analysis, methods that are based on spectral densities and periodograms have a long and successful history. Priestley [28] suggests the use of spectral densities as a graphical tool for model validation by comparing the spectral shape of a dataset with standard ones from well-known parametric models. Tools based on spectral distributions were considered, among others, by Bartlett [2,3], who proposed using the normalized cumulative periodogram to assess whether a process is uncorrelated and to detect hidden periodicities. Rigorous tests for the hypothesis $\mathcal{H}_0 : f = f_0$ for a fixed $f_0$ were derived by Anderson [1], while the more general testing problem $\mathcal{H}_0 : f \in \mathcal{F}_\theta$, where $\mathcal{F}_\theta$ is a parametric class of spectral densities, was treated by Paparoditis [27]. Fan and Zhang [14] consider generalized likelihood ratio tests for the same hypothesis. There is also a rich literature on nonparametric comparison of the multivariate spectra of two time series; recent references include [9,11,12,21,26], but this list is far from complete.

All of the references cited above deal with classical spectral analysis which is based on the autocovariance function and therefore restricted to the aspects of time series dynamics that can be captured by second-order moments. The

* Corresponding author.
*E-mail address:* stanislav.volgushev@utoronto.ca (S. Volgushev).

https://doi.org/10.1016/j.jmva.2019.03.003
0047-259X/

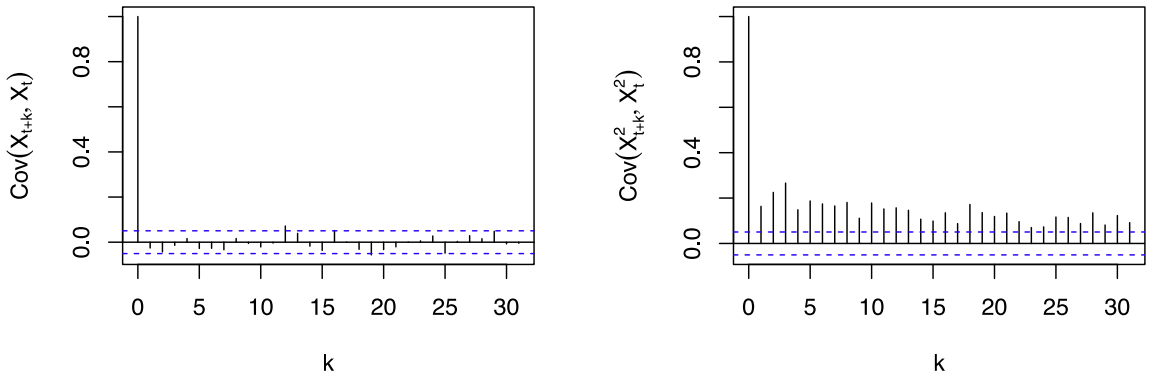


**Fig. 1.** Autocorrelation functions of daily log-returns and squared daily log-returns of the S&P 500 between 2000 and 2005.

autocovariance function does provide a complete description of the dependence of Gaussian processes, but it can completely miss dependencies in a non-Gaussian setting. One such example arises in financial applications when first-order differences of stock market data are analyzed. For illustration, Fig. 1 shows the autocorrelations of the log-returns $X_t$ and the squared log-returns $X_t^2$ calculated from the S&P 500 between 2005 and 2010. While the observations $X_t$ appear to be uncorrelated, we can clearly see positive correlation in the squared observations $X_t^2$. This shows that $X_t$ in fact exhibits strong dependence, which cannot be described through the autocovariance function and therefore completely escapes classical spectral analysis.

This limitation has recently attracted much attention, and new frequency domain tools that can capture non-linear dynamics have been proposed. Pioneering contributions in that direction were made by Hong [17,18] who introduced generalized spectra that are based on joint distributions and joint characteristic functions, respectively. Generalized spectra were later utilized by Hong and Lee [19,20] and Escanciano [13] to test for the validity of various forms of parametric time series models.

More recently, related approaches were taken by Li [24,25], who coined the names of Laplace spectrum and Laplace periodogram. Those ideas were further developed by Hagemann [16] and extended to cross-spectrum and spectral kernel concepts by Dette et al. [10] and Kley et al. [23], who introduced the notion of copula spectral densities, Baruník and Kley [4], who introduced quantile coherency to measure dependence between economic variables, and Birr et al. [6] who consider copula spectra for strictly locally stationary time series.

In the present paper, we utilize copula spectral densities to develop a graphical tool to determine suitable parametric models for time series. We would like to emphasize that our main goal is not to construct yet another test for the hypothesis that a time series is generated by a certain parametric model. Rather, we provide a graphical tool that can indicate whether a chosen model accurately reflects the dependence present in the observed data. By providing useful information about which aspects of the dependence are not captured in case the model is inappropriate, our procedure goes beyond goodness-of-fit tests that merely aim to reject a class of candidate models.

The rest of the paper is organized as follows. Section 2.1 contains a summary of basic properties of copula spectral densities and provides guidelines for their interpretation. In Section 2.2 we provide details on the proposed algorithm. Section 3 gives a theoretical justification for the graphical approach and Section 4 contains a simulation study and a real data example. All proofs are deferred to the Appendix. The Online Supplement contains additional plots and material pertaining to the data analysis.

## 2. Description of the method

### 2.1. Copula spectral densities: Definition and interpretation

We begin by briefly recalling the definition of copula spectral densities. Consider a strictly stationary process $(X_t)_{t\in\mathbb{Z}}$, denote by $F$ its marginal distribution function which we assume to be continuous, by $F_h$ the bivariate distribution function of $(X_{t+h}, X_t)$, and by $C_h$ the corresponding copula. The copula spectral density for the process $X_t$ is then given by

$$f_{\boldsymbol{\tau}}(\omega) = \frac{1}{2\pi}\sum_{k\in\mathbb{Z}} \operatorname{cov}\Big[\mathbf{1}\{F(X_k)\le\tau_1\}, \mathbf{1}\{F(X_0)\le\tau_2\}\Big]\mathrm{e}^{-\mathrm{i}k\omega} = \frac{1}{2\pi}\sum_{k\in\mathbb{Z}}\{C_k(\tau_1,\tau_2)-\tau_1\tau_2\}\mathrm{e}^{-\mathrm{i}k\omega},$$

where $\boldsymbol{\tau} = (\tau_1, \tau_2) \in (0, 1)^2$, $\mathbf{1}(A)$ denotes the indicator function of $A$. We further assume that the terms in the series are absolutely summable. Estimation of copula spectral densities is discussed in the next paragraph, and we begin by providing more insight into their interpretation. Some of the properties mentioned below have been described in [10,16,23], while others are new.

We begin by noting that, being based on copulas, the copula spectral density is invariant under strictly increasing marginal transformations of the time series and thus truly provides information about the temporal dependence structure of the process under consideration. This also implies that the copula spectra of a pair-wise independent time series take the particularly simple form $f_{(\tau,\eta)}(\omega) = (\tau \wedge \eta - \tau\eta)/2\pi$, which is independent of the marginal distribution.

Next we note that $\omega \mapsto f_{\boldsymbol{\tau}}(\omega)$ is $2\pi$-periodic for arbitrary $\boldsymbol{\tau} \in [0,1]^2$ and that $f_{\boldsymbol{\tau}}$ satisfies

$$f_{(\tau_1,\tau_2)}(\omega) = \overline{f_{(\tau_1,\tau_2)}(-\omega)} = \overline{f_{(\tau_2,\tau_1)}(\omega)},$$

where $\bar{a}$ denotes the complex conjugate of $a$. Those properties imply that the values of $\{f_{\boldsymbol{\tau}}(\omega) : \boldsymbol{\tau} \in [0,1]^2, \omega \in [0,\pi]\}$ contain the complete information about the copula spectral density. Even given those restrictions, $f_{\boldsymbol{\tau}}(\omega)$ is still a complex-valued function of three arguments with each argument taking values in an interval, and thus difficult to visualize. One option to get a quick impression of the most important features of the copula spectral density of a given process is to consider all values of $\boldsymbol{\tau} \in \{0.1, 0.5, 0.9\}^2$ and plot the functions $f_{\boldsymbol{\tau}}(\omega)$ against $\omega \in [0,\pi]$. This requires nine plots which are organized as follows:

| | | |
|---|---|---|
| $f_{(0.1,0.1)}(\omega)$ | $\Im f_{(0.5,0.1)}(\omega)$ | $\Im f_{(0.9,0.1)}(\omega)$ |
| $\Re f_{(0.1,0.5)}(\omega)$ | $f_{(0.5,0.5)}(\omega)$ | $\Im f_{(0.9,0.5)}(\omega)$ |
| $\Re f_{(0.1,0.9)}(\omega)$ | $\Re f_{(0.5,0.9)}(\omega)$ | $f_{(0.9,0.9)}(\omega)$. |

In Fig. 2, examples of plots of copula spectral densities corresponding to different parametric models are shown. Those plots will be used to illustrate various properties of copula spectral densities given below.

We begin by observing that $f_{(\tau,\tau)}(\omega)$ is real-valued for any $\tau \in [0,1]$ and $\omega \in \mathbb{R}$. It corresponds to the "classical" $L^2$ spectral density of the clipped process $\mathbf{1}\{F(X_t) \le \tau\}, t \in \mathbb{Z}$ and hence contains information about the dynamics of the level-crossing behavior of the process $\{X_t : t \in \mathbb{Z}\}$. A closer look at Fig. 2 reveals several interesting features. First, for linear Gaussian processes (i. e., AR(1) and MA(1) with i.i.d normal innovations) in (a) and (b), the shape of $f_{(\tau,\tau)}(\omega)$ is similar for all values of $\tau$ and also similar to the corresponding shape of their $L^2$-spectral density. In contrast, the two non-linear models in (c) and (d) have copula spectral densities with shape varying across quantile levels. Both models show no dependence at $\tau = 0.5$, which corresponds to an absence of "central dependence". Yet, both processes show a strong dependence (as indicated by sharp peaks for small values of $\omega$) for more extreme quantiles corresponding to $\tau \in \{0.1, 0.9\}$. Note also that the EGARCH model shows an asymmetric dependence with a higher peak at $\tau = 0.1$ compared to $\tau = 0.9$ indicating a stronger serial dependence in the lower tail. In contrast, the dependence in the GARCH model is completely symmetric.

For $\tau \neq \eta$, $f_{(\tau,\eta)}(\omega)$ can be complex-valued. To interpret the real part of $f_{(\tau,\eta)}(\omega)$, note that after a simple computation we obtain, for $\tau < \eta$,

$$\begin{aligned}
\Re f_{(\tau,\eta)}(\omega) &= -\Re \sum_{k\in\mathbb{Z}} \mathrm{e}^{-\mathrm{i}k\omega}\{\Pr(X_k \le q_\tau, X_0 \ge q_\eta) - \tau(1-\eta)\} \\
&= \tau(1-\eta) - \sum_{k\ge 1} \cos(k\omega)\{\Pr(X_k \le q_\tau, X_0 \ge q_\eta) - \tau(1-\eta)\} \\
&\quad - \sum_{k\ge 1} \cos(k\omega)\{\Pr(X_k \ge q_\eta, X_0 \le q_\tau) - \tau(1-\eta)\}.
\end{aligned}$$

Hence, the function $\omega \mapsto f_{(\tau,\eta)}(\omega)$ contains information about $X_t$ switching between being below $q_\tau$ to above $q_\eta$ and vice versa. In particular, for $\tau$ small and $\eta$ large it can be interpreted as describing the dynamics of the process switching between two extreme states. Here, the negative peak of $\Re f_{(0.1,0.9)}$ at small values of $\omega$ in (c) indicates that the corresponding GARCH process is likely to switch from a high to a low value (or vice versa), which is exactly what happens in periods of high volatility. Similarly, the positive peak in the same function for (a) and (b) corresponds to processes that are unlikely to switch from high to low states immediately, which corresponds to AR(1) or MA(1) dynamics with positive coefficients. It is also interesting to observe that for the two linear processes in (a) and (b), the shapes of $\Re f_{(\tau,\eta)}$ are similar to $f_{(\tau,\tau)}$ for all combinations of $\tau, \eta$.

The imaginary part of $f_{(\tau,\eta)}(\omega)$ for $\tau < \eta$ takes the form

$$\Im f_{(\tau,\eta)}(\omega) = -\sum_{k\ge 1} \sin(\omega k)\{\Pr(X_k \le q_\tau, X_0 \ge q_\eta) - \Pr(X_k \ge q_\eta, X_0 \le q_\tau)\}.$$

Note that $\Im f_{(\tau,\eta)} = 0$ if and only if $\Pr(X_k \le q_\tau, X_0 \ge q_\eta) = \Pr(X_k \ge q_\eta, X_0 \le q_\tau)$ for all $k \in \mathbb{Z}$, which shows that $\Im f_{(\tau,\eta)}$ contains information about asymmetry in going from above $q_\tau$ to below $q_\eta$ and vice versa. Non-zero imaginary parts thus indicate time-irreversibility of the dynamics in the observed time series. In particular, if $\Im f_{(\tau,\eta)} = 0$ for all $\tau, \eta$, then this indicates that the process under consideration is pairwise time-reversible, i.e., $C_k(\tau,\eta) = C_{-k}(\tau,\eta)$ for all $k, \tau, \eta$. The Gaussian linear processes in (a) and (b) of Fig. 2 are time reversible, which is confirmed by the flat imaginary parts of their copula spectra. It is also noteworthy that the imaginary parts of the processes in (c) and (d) show very different behavior, with clear time-irreversibility for the EGARCH process in (d) and no immediately visible evidence of the same for the GARCH process in (c).

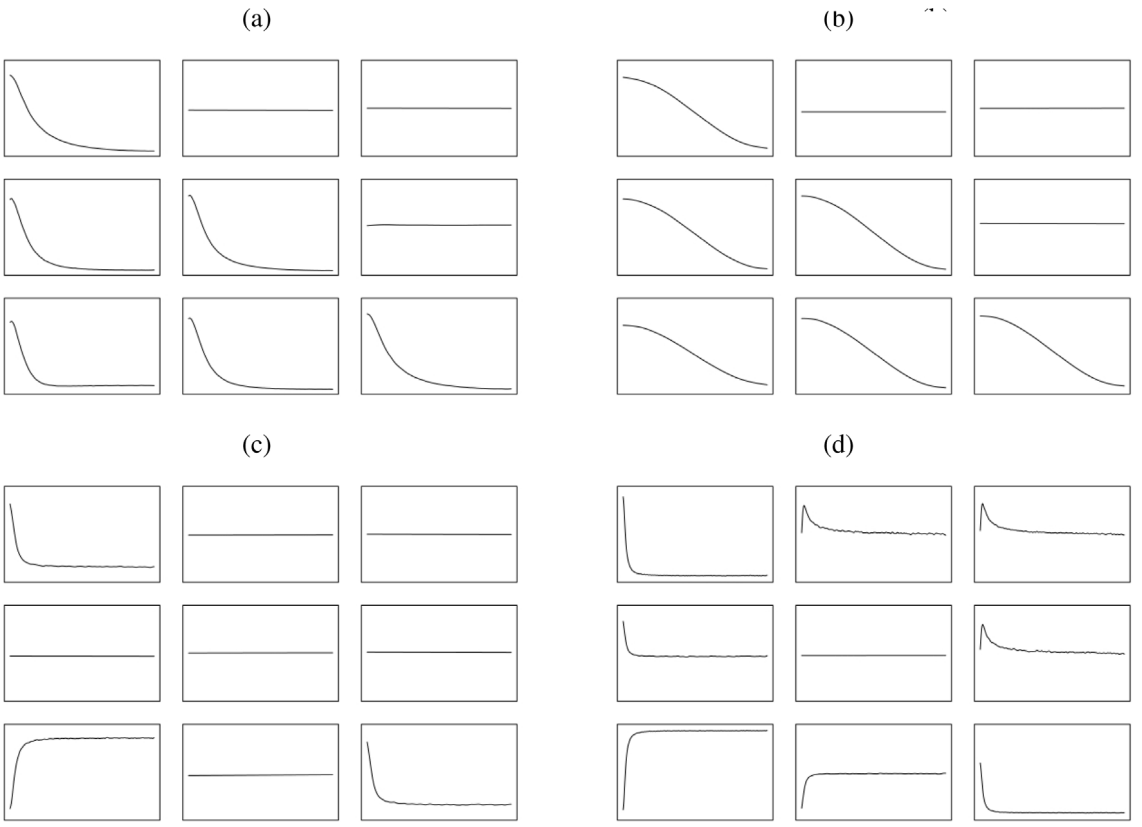


**Fig. 2.** Copula spectral densities for $\boldsymbol{\tau} \in \{0.1, 0.5, 0.9\}^2$ of an (a) AR(1), (b) MA(1), (c) GARCH(1, 1) and (d) EGARCH(1, 1) process; all processes have Gaussian innovations.

### 2.2. Graphical tools for model validation

We begin by briefly reviewing estimation of copula spectral densities as discussed in [23]; see also [16] where the case $\tau_1 = \tau_2$ is considered and [10] for alternative estimators based on ranks and quantile regression. Given the observations $X_0, \ldots, X_{n-1}$, we calculate their empirical distribution function $\hat{F}_n(x) = \sum_{t=0}^{n-1} \mathbf{1}(X_t \leq x)/n$ and the copula periodogram

$$I_{\boldsymbol{\tau},n}(\omega) = \frac{1}{2\pi n} d_{\tau_1,n}(\omega) d_{\tau_2,n}(-\omega),$$

where $\boldsymbol{\tau} = (\tau_1, \tau_2)$ and

$$d_{\tau,n}(\omega) = \sum_{t=0}^{n-1} \mathbf{1}\{\hat{F}_n(X_t) \leq \tau\} \mathrm{e}^{-\mathrm{i}\omega t}.$$

To obtain a consistent estimator we smooth the copula periodograms over frequencies

$$\hat{f}_{\boldsymbol{\tau}}(\omega) = \frac{1}{2\pi n} \sum_{s=1}^{n} W_n(\omega - 2\pi s/n) I_{\boldsymbol{\tau},n}(2\pi s/n), \tag{1}$$

where $W_n$ denotes a sequence of weighting functions which are specified in assumption (W). Kley et al. [23] proved the asymptotic normality of this estimator and computation is possible via the R package `quantspec`; see [22].

Now, given the observations $X_1, \ldots, X_n$ we want to decide whether they could have been produced by a parametric model $P^{\theta_0}$, where $\{P^\theta : \theta \in \Theta\}$ is a collection of candidate models and $\theta_0 \in \Theta$ is an unknown parameter. Here, the specification of parametric models explicitly includes the specification of an innovation density. Indeed, as demonstrated in [10] copula spectral densities of linear processes depend on the underlying innovations. To this end we propose to apply parametric bootstrap ideas in the form of Algorithms 1 and 2 given below.

**Algorithm 1.** To obtain a graphical representation of the copula spectral density estimated from data along with "typical regions" from a parametric model using the parametric bootstrap, proceed as follows.

**Data**: Observations $X_1, \dots, X_n$

**Input**: A class $\{P^\theta : \theta \in \Theta\}$ of parametric models, an estimator $\hat\theta$, a collection of frequencies $\omega_1, \dots, \omega_K \in [0, \pi]$, and a quantile level $\boldsymbol{\tau} = (\tau_1, \tau_2)$

**Output**: Plot comparing copula spectral density estimated from data with 'typical regions' created by a parametric bootstrap

**begin**

Estimate $\hat\theta$ from $X_1, \dots, X_n$

**for** *r in 1:R* **do** /* parametric bootstrap */

$X_1^{\hat\theta,r}, \dots, X_n^{\hat\theta,r}$ = simulate from the model $P^{\hat\theta}$

$\hat f_{\boldsymbol{\tau}}^{\hat\theta,r}(\omega_k)$ = estimated copula spectral density from $X_1^{\hat\theta,r}, \dots, X_n^{\hat\theta,r}$

**end**

/* Calculate lower and upper bounds, separately for real and imaginary parts: */

$$\ell^{\Re}_{\boldsymbol{\tau},R}(\omega_k) = \alpha/2 - \text{quantile}\{\Re\hat f_{\boldsymbol{\tau}}^{\hat\theta,1}(\omega_k), \dots, \Re\hat f_{\boldsymbol{\tau}}^{\hat\theta,R}(\omega_k)\},$$

$$\ell^{\Im}_{\boldsymbol{\tau},R}(\omega_k) = \alpha/2 - \text{quantile}\{\Im\hat f_{\boldsymbol{\tau}}^{\hat\theta,1}(\omega_k), \dots, \Im\hat f_{\boldsymbol{\tau}}^{\hat\theta,R}(\omega_k)\}$$

$$u^{\Re}_{\boldsymbol{\tau},R}(\omega_k) = (1-\alpha/2) - \text{quantile}\{\Re\hat f_{\boldsymbol{\tau}}^{\hat\theta,1}(\omega_k), \dots, \Re\hat f_{\boldsymbol{\tau}}^{\hat\theta,R}(\omega_k)\}$$

$$u^{\Im}_{\boldsymbol{\tau},R}(\omega_k) = (1-\alpha/2) - \text{quantile}\{\Im\hat f_{\boldsymbol{\tau}}^{\hat\theta,1}(\omega_k), \dots, \Im\hat f_{\boldsymbol{\tau}}^{\hat\theta,R}(\omega_k)\}$$

/* Estimate the Copula Spectral Density for the data: */

$\hat f_{\boldsymbol{\tau}}(\omega)$ = estimate the copula spectral density from $X_1, \dots, X_n$

/* Plot the result */

plot $\{\hat f_{\boldsymbol{\tau}}(\omega_k)\}_{k=1,\dots,K}$

plot intervals computed from $\{\ell_{\boldsymbol{\tau},R}(\omega_k), u_{\boldsymbol{\tau},R}(\omega_k)\}_{k=1,\dots,K}$ (separately for real and imaginary parts)

**end**

Algorithm 1 provides a graphical summary of the copula spectral density estimated from data (blue lines) for a few distinct combinations of quantile levels (in the present paper, $(\tau_1, \tau_2) \in \{0.1, 0.5, 0.9\}^2$) as a function of $\omega$ together with typical regions (gray areas) that would contain this estimator with probability $1-\alpha$ if the corresponding class of parametric candidate models were specified correctly. For a formal statement, see Proposition 1 . One potential concern with Algorithm 1 is that the graphics can become overwhelming if many different quantile levels need to be considered simultaneously. Algorithm 2 provides a useful supplement to Algorithm 1 which allows us to consider many quantile levels at the same time. This necessitates a different graphical representation. The results from Algorithm 2 can be displayed in two different ways. The first provides a summary over all quantile levels $(\tau_1, \tau_2) \in M$ (in the present paper, we choose $M = \{0.05, \dots, 0.95\}^2$) indicating whether the candidate model class produces spectral densities which are compatible with the data for a given frequency but uniformly over quantile levels. If a deviation is detected for a given frequency, a second plot for that particular frequency can be used to determine at which quantile levels the mismatch between the data and the parametric candidate model appears.

Fig. 3 illustrates Algorithm 1 in a simple example. Here the data is a single path simulated from a GARCH(1, 1) model with Gaussian innovations, and we considered two classes of parametric models, namely Gaussian AR(3) and GARCH(1, 1) with Gaussian innovations. The true model is in the latter class, but the parameter was not specified. The blue line shows the estimated copula spectral density $\hat f_{\boldsymbol{\tau}}$ and the gray area corresponds to the typical regions for $\alpha = 0.05$; see Algorithm 1 for details. The plot is organized as discussed in Section 2.1 with $\Re\hat f_{\boldsymbol{\tau}}$ on/below and $\Im\hat f_{\boldsymbol{\tau}}$ above the diagonal, respectively. We clearly see that an AR(3) model is unable to describe the dynamics of a GARCH model, as it fails to capture the dependency in the extreme quantiles $\boldsymbol{\tau} = (0.1, 0.1), (0.9, 0.9), (0.1, 0.9)$, especially at low frequencies. Considering the true model class in the right panel, we see that the typical regions now almost completely contain the estimated spectrum. Note that typical regions are computed pointwise in $\boldsymbol{\tau}$, $\omega$, so the estimator can occasionally be just outside of the boundary of typical regions.

The output of Algorithm 2 for the same data and the same models is illustrated in Figs. 4 and 5. First, consider the summary plots in Fig. 4 with frequencies on the $x$-axis and $p_{\min,R}(\omega)$ defined in Algorithm 2 on the $y$-axis; the $y$-axis is on the log-scale for better visibility of very low values. By Proposition 1, proved below, the values on the $y$-axis can be interpreted as $p$-values (uniform in $\boldsymbol{\tau}$ and pointwise in $\omega$) of a test for the null hypothesis that the data were generated from a model in the given parametric class against a non-parametric alternative. The left panel of Fig. 4 shows the plot corresponding to a Gaussian AR(3) model class. Here, the candidate model class obviously does not match the data; this is particularly visible at the lowest frequencies where several $p$-values in a row are below 0.001. In contrast, the right panel which uses the true model class does not show evidence of model misspecification.

**Algorithm 2.** To obtain a graphical representation of 'critical $\boldsymbol{\tau}$s' from a parametric model (with parameter estimated from the data) together with the estimator based on observations, proceed as follows.

**Data**: Observations $X_1, \dots, X_n$

**Input**: A class $\{P^\theta : \theta \in \Theta\}$ of parametric models, an estimator $\hat{\theta}$, a frequency $\omega \in [0, \pi]$, a number $K$ of how many equally spaced quantile levels should be used and a quantile level $\beta$

**Output**: Heat-plot of signed $p$-values indicating whether estimated copula spectral density is within 'typical regions' created by a parametric bootstrap

**begin**

Estimate $\hat{\theta}$ from $X_1, \dots, X_n$

**for** $r$ *in 1:R* **do** /* parametric bootstrap */

$X_1^{\hat{\theta},r}, \dots, X_n^{\hat{\theta},r}$ = simulate from the model $P^{\hat{\theta}}$

$\hat{f}_{\boldsymbol{\tau}}^{\hat{\theta},r}(\omega)$ = estimated copula spectral density from $X_1^{\hat{\theta},r}, \dots, X_n^{\hat{\theta},r}$, $\forall \boldsymbol{\tau} \in M = \{1/(K+1), \dots, K/(K+1)\}^2$

**end**

/* Calculate scaling factors, separately for real & imaginary parts: */

Let $\ell_{\boldsymbol{\tau},R}^{\Re}(\omega), u_{\boldsymbol{\tau},R}^{\Re}(\omega)$ denote $\beta/2$ and $1-\beta/2$ quantile of $\Re\hat{f}_{\boldsymbol{\tau}}^{\hat{\theta},1}(\omega), \dots, \Re\hat{f}_{\boldsymbol{\tau}}^{\hat{\theta},R}(\omega)$, respectively (same for $\Im$)

Define

$$c_{\boldsymbol{\tau},R}^{\Re}(\omega) = \{u_{\boldsymbol{\tau},R}^{\Re}(\omega) + \ell_{\boldsymbol{\tau},R}^{\Re}(\omega)\}/2, \quad c_{\boldsymbol{\tau},R}^{\Im}(\omega) = \{u_{\boldsymbol{\tau},R}^{\Im}(\omega) + \ell_{\boldsymbol{\tau},R}^{\Im}(\omega)\}/2$$

and

$$\Delta_{\boldsymbol{\tau},R}^{\Re}(\omega) = \{u_{\boldsymbol{\tau},R}^{\Re}(\omega) - \ell_{\boldsymbol{\tau},R}^{\Re}(\omega)\}/2, \quad \Delta_{\boldsymbol{\tau},R}^{\Im}(\omega) = \{u_{\boldsymbol{\tau},R}^{\Im}(\omega) - \ell_{\boldsymbol{\tau},R}^{\Im}(\omega)\}/2 + 10^{-6}\mathbf{1}\{u_{\boldsymbol{\tau},R}^{\Im}(\omega) = \ell_{\boldsymbol{\tau},R}^{\Im}(\omega)\}$$

The scaled and centred bootstrap replicate is

$$A_r^{\Re}(\omega) = \max_{\boldsymbol{\tau}=(\tau_1,\tau_2)\in M} |\Re\hat{f}_{\boldsymbol{\tau}}^{\hat{\theta},R}(\omega) - c_{\boldsymbol{\tau},R}^{\Re}(\omega)|/\Delta_{\boldsymbol{\tau},R}^{\Re}(\omega), \quad A_r^{\Im}(\omega) = \max_{\boldsymbol{\tau}=(\tau_1,\tau_2)\in M} |\Im\hat{f}_{\boldsymbol{\tau}}^{\hat{\theta},R}(\omega) - c_{\boldsymbol{\tau},R}^{\Im}(\omega)|/\Delta_{\boldsymbol{\tau},R}^{\Im}(\omega).$$

/* Estimate the Copula Spectral Density for the data: */

$\hat{f}_{\boldsymbol{\tau}}(\omega)$ = estimate the copula spectral density from $X_1, \dots, X_n$. Define

$$E_{\boldsymbol{\tau}}^{\Re}(\omega) = |\Re\hat{f}_{\boldsymbol{\tau}}(\omega) - c_{\boldsymbol{\tau},R}^{\Re}(\omega)|/\Delta_{\boldsymbol{\tau},R}^{\Re}(\omega), \quad E_{\boldsymbol{\tau}}^{\Im}(\omega) = |\Im\hat{f}_{\boldsymbol{\tau}}(\omega) - c_{\boldsymbol{\tau},R}^{\Im}(\omega)|/\Delta_{\boldsymbol{\tau},R}^{\Im}(\omega)$$

/* Calculate $p$-values, separately for real and imaginary parts: */

Define $\hat{F}_R$ as the empirical cdf of $\max\{A_1^{\Re}(\omega), A_1^{\Im}(\omega)\}, \dots, \max\{A_R^{\Re}(\omega), A_R^{\Im}(\omega)\}$ and compute

$$p_{\boldsymbol{\tau},R}^{\Re}(\omega) = 1 - \hat{F}_R\{E_{\boldsymbol{\tau}}^{\Re}(\omega)-\}, \quad p_{\boldsymbol{\tau},R}^{\Im}(\omega) = 1 - \hat{F}_R\{E_{\boldsymbol{\tau}}^{\Im}(\omega)-\}, \quad p_{\min,R}(\omega) = \min_{\boldsymbol{\tau}\in M} \min\{p_{\boldsymbol{\tau},R}^{\Re}(\omega), p_{\boldsymbol{\tau},R}^{\Im}(\omega)\}.$$

/* Plot the result */

plot 1: $\omega \mapsto p_{\min,R}(\omega)$; $x$-axis from $1/R$ to 1, on the log-scale; $p_{\min,R}(\omega) = 0$ is indicated by a red circle on the $x$-axis.

plot 2: $K \times K$ panels for each $\omega$. The position within each panel corresponds to $\boldsymbol{\tau} \in M$, the symbols used correspond to the magnitude of $p_{\boldsymbol{\tau},R}^{\Re}(\omega), p_{\boldsymbol{\tau},R}^{\Im}(\omega)$ (1, 2 and 3 triangles correspond to $p_{\boldsymbol{\tau},R}^{\cdot}(\omega) < 0.05$, 0.01 and 0.001, respectively), and sign of $\Re\hat{f}_{\boldsymbol{\tau}}(\omega) - c_{\boldsymbol{\tau},R}^{\Re}(\omega)$, $\Im\hat{f}_{\boldsymbol{\tau}}(\omega) - c_{\boldsymbol{\tau},R}^{\Im}(\omega)$ (red triangles facing up indicate a positive and blue triangles facing down indicate a negative value). Information corresponding to $p_{(\tau_i,\tau_j),R}^{\Re}(\omega)$ for $i \geq j$ is in row $i$ column $j$ and information on $p_{(\tau_i,\tau_j),R}^{\Im}(\omega)$ for $i < j$ in row $i$ column $j$.

**end**

The plots in Fig. 5 provide more detailed information about the quantile levels at which a mismatch between the data-based spectrum and a spectrum from the candidate parametric model is detected for the frequencies $\omega = 0$ and $\omega = 4\pi/64$. Here, the blue triangles facing down indicate that the data-based spectrum is smaller compared to the candidate model spectrum (with 1, 2 and 3 triangles indicating significance at the 5%, 1% and 0.1% level, respectively) while the red triangles facing up indicate a data-based spectrum which is larger compared to the candidate model spectrum. The corresponding plots reveal that most of the disagreement between the data and the model dynamics happens in the real parts of spectra corresponding to quantile levels $(\tau_1, \tau_2)$ where both $\tau_1$ and $\tau_2$ are either unexpectedly small or unexpectedly large. This confirms the first impression obtained from the plots in Fig. 3 and provides a more detailed view of the quantiles where the data and the model spectra disagree.

## 3. Formal justification of the graphical tools

In this section, we present a formal justification for the graphical approaches introduced in Section 2.2. Denote by $\Theta \subset \mathbb{R}^d$ a candidate parameter space. For any $\theta \in \Theta$, let $(X_t^\theta)_{t\in\mathbb{Z}}$ be a strictly stationary process distributed according to

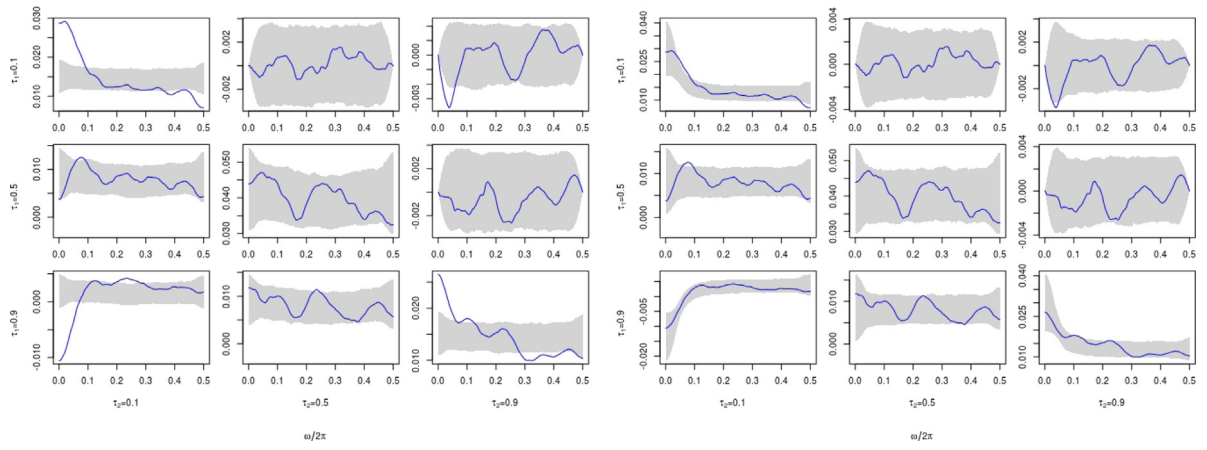

**Fig. 3.** Example in which our graphical tool (Algorithm 1) is applied to data generated from a Gaussian GARCH(1, 1) model with $n = 1024$ observations. We fit a Gaussian AR(3) model (left) and a Gaussian GARCH(1, 1) model (right). Blue lines are the estimated copula spectral densities, gray areas correspond to typical regions from Algorithm 1 with $\alpha = 0.05$. The plot is organized as discussed in Section 2.1; $\Re\hat{f}_{\boldsymbol{\tau}}$ on/below and $\Im\hat{f}_{\boldsymbol{\tau}}$ above the diagonal, respectively.

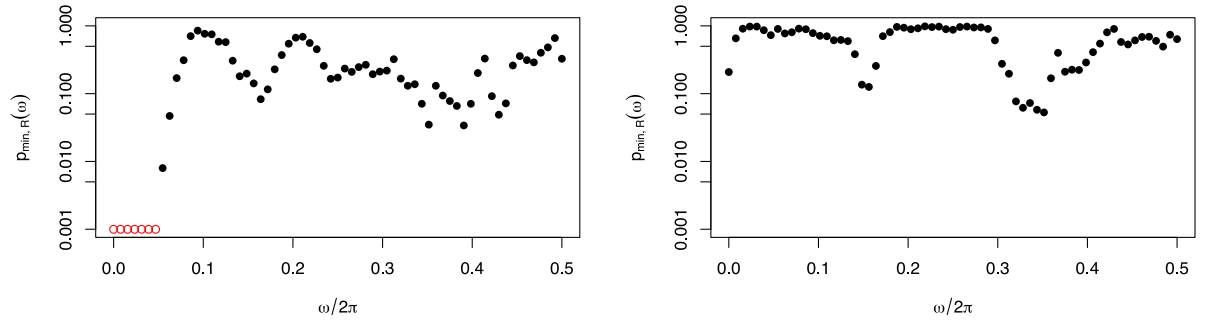


**Fig. 4.** Summary plot produced by the application of Algorithm 2 to data generated from a GARCH(1, 1) with Gaussian innovations with $n = 1024$ observations. We fit a Gaussian AR(3) model (left) and a GARCH(1, 1) model with Gaussian innovations (right). Frequencies are on the $x$-axis and $p_{\min,R}(\omega)$ defined in Algorithm 2 on the $y$-axis; the $y$-axis is on the log-scale for better visibility of very low values. Red circles correspond to $p_{\min,R}(\omega) < 0.001$.

$P^\theta$. Furthermore, let $F^\theta$ be the cumulative marginal distribution function of $X_t^\theta$, let $F_h^\theta$ be the bivariate distribution function of $(X_{t+h}^\theta, X_t^\theta)$, and let $C_h^\theta$ be the copula of $(X_{t+h}^\theta, X_t^\theta)$. We denote the copula spectral density of the process $(X_t^\theta)_{t\in\mathbb{Z}}$ by

$$f_{\boldsymbol{\tau}}^\theta(\omega) = \frac{1}{2\pi}\sum_{k\in\mathbb{Z}}\{C_k^\theta(\tau_1,\tau_2) - \tau_1\tau_2\}\mathrm{e}^{-\mathrm{i}k\omega}.$$

The corresponding estimator $\hat{f}^\theta$, which is computed from $X_1^\theta, \ldots, X_n^\theta$, is denoted by $\hat{f}_{\boldsymbol{\tau}}^\theta$. We make the following technical assumptions.

**(LC)** The copulas $C_h^\theta$ are Lipschitz continuous with respect to the parameter $\theta$ in a neighborhood of $\theta_0$ uniformly in $\mathcal{T} \subseteq [0, 1]^2$, i.e., there exist constants $\varepsilon > 0, L < \infty$ such that whenever $\|\theta - \theta_0\| \le \varepsilon$,

$$\sup_{h\in\mathbb{Z}}\sup_{\boldsymbol{\tau}\in\mathcal{T}}|C_h^\theta(\boldsymbol{\tau}) - C_h^{\theta_0}(\boldsymbol{\tau})| \le L\,\|\theta - \theta_0\|.$$

**(C)** For any $p \in \mathbf{N}$ there exist constants $\rho_p \in (0, 1)$ and $K_p < \infty$ such that, for arbitrary intervals $A_1, \ldots, A_p \subset \mathbb{R}$ and arbitrary $t_1, \ldots, t_p \in \mathbb{Z}$,

$$\sup_{\|\theta-\theta_0\|\le\varepsilon}|\mathrm{cum}\{\mathbf{1}(X_{t_1}^\theta \in A_1), \ldots, \mathbf{1}(X_{t_p}^\theta \in A_p)\}| \le K_p\rho_p^{\max_{i,j}|t_i - t_j|}.$$

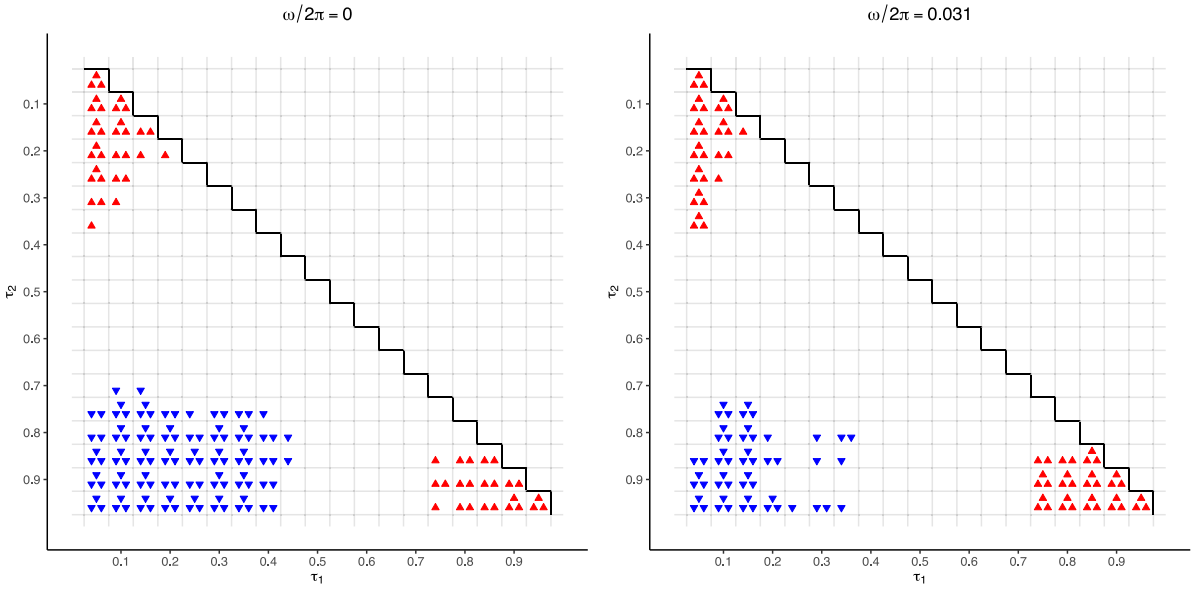


**Fig. 5.** Detailed plots for two frequencies produced by Algorithm 2 on data generated from a GARCH(1, 1) model with Gaussian innovations with $n = 1024$ observations with fitted Gaussian AR(3) model (compare left panel of Fig. 4). Blue triangles facing down indicate that the data-based spectrum is smaller compared to the candidate model spectrum (with 1, 2 and 3 triangles indicating significance at the 5%, 1% and 0.1% level, respectively) while red triangles facing up indicate a data-based spectrum which is larger compared to the candidate model spectrum . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**(W)** The weight function $W$ is real-valued and even with support $[-\pi, \pi]$; moreover, it has bounded variation and satisfies $\int W(u)\mathrm{d}u = 1$. We denote by $b_n$ a sequence of scaling parameters such that $b_n \to 0$ and $nb_n \to \infty$, and assume that $W_n$ in Eq. (1) takes the form

$$W_n(u) = \sum_{j\in\mathbb{Z}} b_n^{-1} W\{b_n^{-1}(u + 2\pi j)\}.$$

**Remark 1.** Assumption (C) is fulfilled under certain mixing assumptions; see Propositions 3.1 and 3.2 in [23]. Assumption (W) places restrictions on the smoothing function which are standard in time series analysis; see, e.g., p. 147 in [7]. Assumption (LC) ensures that if $\theta_n$ is close to $\theta_0$, the corresponding copula spectral densities are also close. Below we show that this assumption is satisfied for ARMA$(p, q)$ processes with normal innovations.

**Example 1.** Let $\{X_t^\theta : t \in \mathbb{Z}\}$ be a strictly stationary ARMA$(p, q)$ process where $\theta = (a_1, \ldots, a_p, b_1, \ldots, b_q)$ denotes the AR and MA coefficients, i.e., $X_t^\theta$ solves

$$X_t^\theta - \sum_{j=1}^{p} a_j X_{t-j}^\theta = \epsilon_t + \sum_{i=1}^{q} b_i \epsilon_{t-i},$$

where and $\epsilon_t$ are centred i.i.d $\mathcal{N}(0, 1)$ random variables. Using the backshift operator $B$, we can write this as $P^\theta(B)X_t^\theta = Q_\theta(B)\epsilon_t$, where $P^\theta$ and $Q^\theta$ are polynomials given, for all $z \in \mathbb{C}$, by

$$P^\theta(z) = 1 - a_1 z - \cdots - a_p z^p, \quad Q^\theta(z) = 1 + b_1 z + \cdots + b_q z^q.$$

To guarantee the existence of a unique strictly stationary and causal solution, see [8], we assume that $\Theta$ is a set such that for all $\theta \in \Theta$ the polynomials $P^\theta$ and $Q^\theta$ have no common roots and $P^\theta$ only has roots outside of the unit circle $\{z \in \mathbb{C} : |z| \leq 1\}$. Under these conditions, (LC) holds for any open $\mathcal{T} \subset [0, 1]^2$ and any $\theta_0$ in the interior of $\Theta$. This statement will be proved in the Appendix.

The main result in this section is Proposition 1. It implies that if the parametric model is specified correctly, the intervals $[\ell^{\Re}_{\boldsymbol{\tau},R}(\omega), u^{\Re}_{\boldsymbol{\tau},R}(\omega)]$ and $[\ell^{\Im}_{\boldsymbol{\tau},R}(\omega), u^{\Im}_{\boldsymbol{\tau},R}(\omega)]$ will contain the real and imaginary parts of the estimator $\hat{f}_{\boldsymbol{\tau}}(\omega)$ with given probability $\alpha$ asymptotically. This provides a formal justification for the graphical approach introduced in Algorithm 1. The second part of Proposition 1 shows that the output of Algorithm 2 can indeed be interpreted as $p$-values for the null hypothesis that the class of candidate models contains the true model.

**Proposition 1.** *Assume that the data $X_1, \ldots, X_n$ are generated from the model $P^{\theta_0}$ and let $\hat{\theta}_n$ be a $\sqrt{n}$-consistent estimator of $\theta_0$. Suppose that Assumptions (LC), (C), (W) hold and assume that $R = R_n \to \infty$ as $n \to \infty$. Further assume that there exist*

*constants* $k \in \mathbb{N}$ *and* $\kappa > 0$ *with*

$$b_n = o(n^{-1/(2k+1)}), \quad b_n n^{1-\kappa} \to \infty.$$

*Then, for* $\ell_{\boldsymbol{\tau},R_n}(\omega), u_{\boldsymbol{\tau},R_n}(\omega)$ *defined in Algorithm 1 we have, for any* $\boldsymbol{\tau} \in \mathcal{T}, \omega \in \mathbb{R}$ *with* $\operatorname{var}\{\Re H_0(\boldsymbol{\tau},\omega)\} \neq 0$,

$$\lim_{n\to\infty} \Pr\{\ell^{\Re}_{\boldsymbol{\tau},R_n}(\omega) \leq \Re\hat{f}_{\boldsymbol{\tau}}(\omega) \leq u^{\Re}_{\boldsymbol{\tau},R_n}(\omega)\} = 1-\alpha. \tag{2}$$

*The same holds for the imaginary parts. If additionally* $\min_{\boldsymbol{\tau}\in M} \operatorname{var}\{\Re H_0(\boldsymbol{\tau},\omega)\} > 0$ *and* $\min_{\boldsymbol{\tau}\in M} \operatorname{var}\{\Im H_0(\boldsymbol{\tau},\omega)\} > 0$ *then, for* $p^{\Re}_{\boldsymbol{\tau},R}(\omega), p^{\Im}_{\boldsymbol{\tau},R}(\omega)$ *defined in Algorithm 2, we have, for any* $\omega \in \mathbb{R}$,

$$\lim_{n\to\infty} \Pr\Big[\min_{\boldsymbol{\tau}\in M} \min\{p^{\Re}_{\boldsymbol{\tau},R}(\omega), p^{\Im}_{\boldsymbol{\tau},R}(\omega)\} < \alpha\Big] = \alpha. \tag{3}$$

The key technical ingredient for proving Proposition 1 is given by the following theorem. It provides a generalization of Theorem 3.6 in [23] to a particular kind of triangular array asymptotics. This result is of independent interest, and hence we choose to state it separately.

**Theorem 1.** *Suppose that Assumptions* $(LC), (C), (W)$ *hold, and assume that there exist constants* $k \in \mathbb{N}$ *and* $\kappa > 0$ *with*

$$b_n = o\{n^{-1/(2k+1)}\}, \quad b_n n^{1-\kappa} \to \infty.$$

*If* $\theta_n = \theta_0 + O(n^{-1/2})$ *then, for* $\mathcal{T}$ *from Assumption (LC), as* $n \to \infty$,

$$\sqrt{nb_n}\,\{\hat{f}^{\theta_n}_{\boldsymbol{\tau}}(\omega) - f^{\theta_0}_{\boldsymbol{\tau}}(\omega) - B^{(k)}_n(\boldsymbol{\tau},\omega)\}_{\boldsymbol{\tau}\in\mathcal{T}} \rightsquigarrow H_0(\cdot;\omega)$$

*in* $\ell^\infty(\mathcal{T})$, *where* $\rightsquigarrow$ *denotes weak convergence in* $\ell^\infty(\mathcal{T})$ *and*

$$B^{(k)}_n(\boldsymbol{\tau},\omega) = \begin{cases} \displaystyle\sum_{j=2}^{k} b_n^j \int_{-\pi}^{\pi} u^j W(u)\mathrm{d}u (f^{\theta_0}_{\boldsymbol{\tau}})^{(j)}(\omega)/j! & \text{if} \quad \omega \neq 0 \pmod{2\pi}, \\ n\tau_1\tau_2/(2\pi) & \text{if} \quad \omega = 0 \pmod{2\pi}, \end{cases}$$

*and* $H_0(\cdot;\omega)$ *is a complex-valued, centered Gaussian process characterized by*

$$\operatorname{cov}\Big[H_0\{(u_1,v_1);\omega\}, H_0\{(u_2,v_2);\omega\}\Big] = 2\pi \int_{-\pi}^{\pi} W^2(u)\mathrm{d}u \times \Big[f^{\theta_0}_{(u_1,u_2)}(\omega) f^{\theta_0}_{(v_1,v_2)}(-\omega) + f^{\theta_0}_{(u_1,v_2)}(\omega) f^{\theta_0}_{(v_1,u_2)}(-\omega)\mathbf{1}\{\omega = 0 \pmod{\pi}\}\Big].$$

## 4. Simulation study and data example

In this section, we present a simulation study and an application to the returns of the S&P 500 stock index between 2000 and 2005 and 1966 and 1970.

### 4.1. Real data example: S&P 500 returns

In this section we demonstrate how the proposed graphical tools can be utilized to find an appropriate parametric model for a given time series and further provide an example where none of the standard models seem to work. To this end we consider the daily log-returns of the S&P 500 between 03.01.2000 and 30.12.2005 (corresponding to $n = 1508$ observations) as well as between 03.01.1966 and 31.12.1970 (corresponding to $n = 1233$ observations). Throughout this section we use the Epanechnikov kernel for $W_n$, a moderate bandwidth $b_n = 0.1$ and set $\alpha = 0.05$ in Algorithm 1.

We first consider the daily log-returns of the S&P 500 between 03.01.2000 and 30.12.2005. Assuming for the moment that we have little knowledge about financial time series, we first attempt to fit a Gaussian AR(3) model. Algorithm 1 with this model class produces Fig. 6(a). This clearly shows that a Gaussian AR(3) manages to capture the "median dependence" but cannot account for the strong dependencies observed at $\boldsymbol{\tau} = (0.1, 0.1), (0.9, 0.9)$ and $(0.9, 0.1)$. This is further confirmed in the output produced by Algorithm 2; see Fig. 7(a). The most basic model which has the potential to model such dependencies is an ARCH(1) model with Gaussian innovations, which is our next candidate. Plot (b) in Fig. 6 indeed shows that an ARCH(1) model with Gaussian innovations produces the peaks around frequency $\omega = 0$, but those peaks are not high enough to match the data; this is again confirmed by the summary plot from Algorithm 2 provided in Fig. 7(b).

Our next try is a GARCH(1, 1) model with Gaussian innovations which was specifically designed to model the types of dependence observed in financial data. Fig. 6(c) shows that this model is well suited to reproduce the peaks for $\boldsymbol{\tau} = (0.1, 0.1), (0.9, 0.9)$ and $(0.9, 0.1)$. However, the imaginary parts of the spectra still do not match the data, as can be seen from the part of Fig. 6(c) corresponding to $\boldsymbol{\tau} = (0.1, 0.9)$; the mismatch between the model and the data dynamics is confirmed in the summary plot from Fig. 7(c). A closer look at the corresponding detailed plot in the top row of Fig. 8 sheds light on the specific combinations of quantile levels for which a significant mismatch occurs.

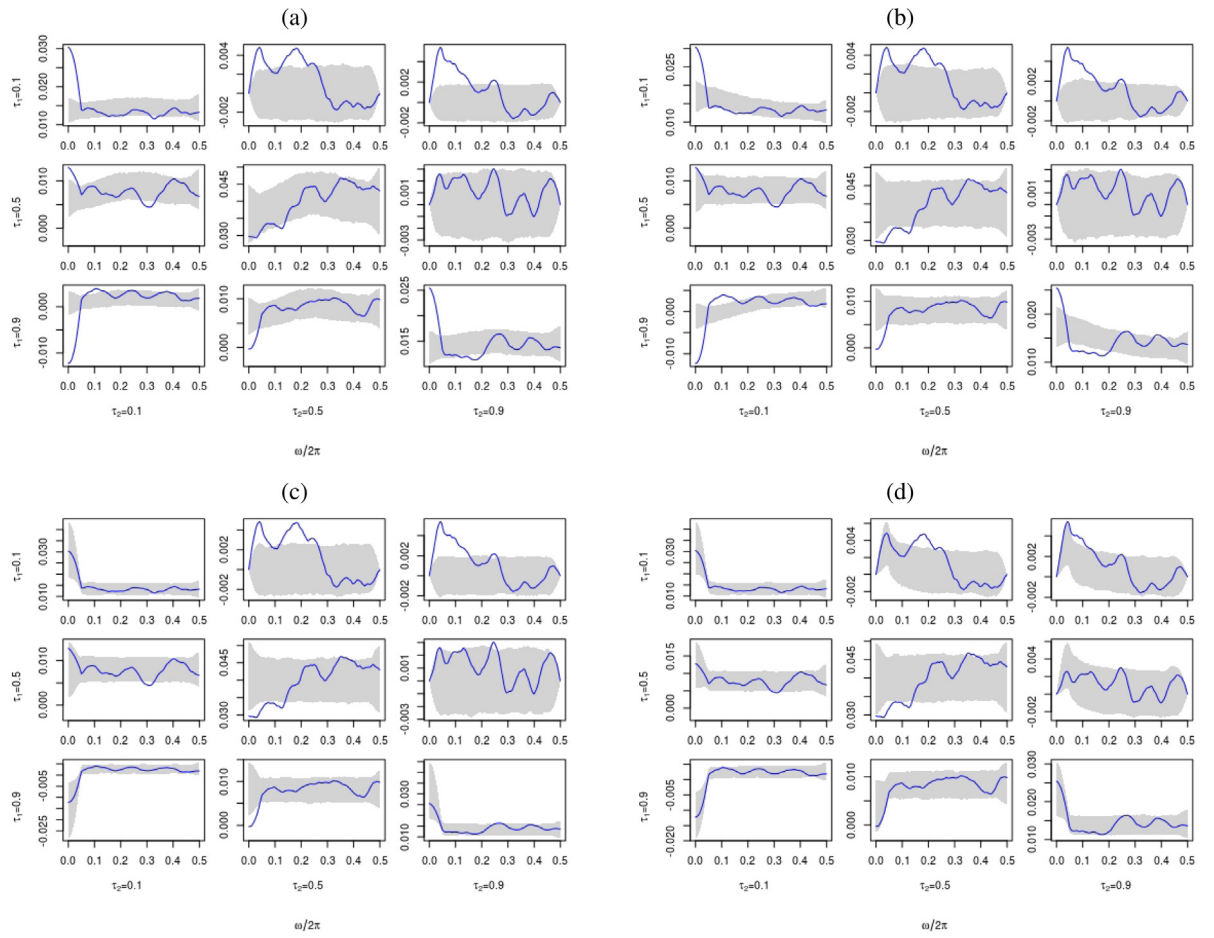


**Fig. 6.** Estimated copula spectral densities based on the daily log-returns of the S&P 500 between 2000 and 2005. Displayed are the plots produced by Algorithm 1 for the following model classes (a) AR(3), (b) ARCH(1), (c) GARCH(1, 1) and (d) EGARCH(1, 1); all models with Gaussian innovations. Blue lines are estimated copula spectral densities, gray areas correspond to typical regions from Algorithm 1 with $\alpha = 0.05$. The plot is organized as discussed in Section 2.1; $\Re\hat{f}_{\boldsymbol{\tau}}$ on/below and $\Im\hat{f}_{\boldsymbol{\tau}}$ above the diagonal, respectively.

Based on the discussion in Section 2.1 about asymmetric dependence a reasonable model to try is an EGARCH(1, 1) model with Gaussian innovations. The output of Algorithm 1 in Fig. 6(d) indeed indicates that among all models considered so far, this one has the best performance, although we still detect slight deviations for some of the imaginary parts. The impression that this model still does not provide a perfect fit is further strengthened by the summary plot in Fig. 7(d) where we see a fairly high proportion of $p$-values below 5% which is much higher than we would expect even after adjusting for multiple testing across frequencies.

In the final part of this section, we consider the daily log-returns of the S&P 500 between 1966 and 1970. The output of Algorithm 1 for the same four model classes as considered above is depicted in Fig. 9(a)–(d). Interestingly, we find that none of the four model classes provide an adequate description of the dynamics observed in the data since the data contain both linear type dynamics at the median level, but also strong GARCH-like tail dependencies and EGARCH-like imaginary parts that appear to be too steep to be captured by an EGARCH(1, 1) model with Gaussian innovations, indicating a strongly asymmetric behavior of the process going forward and backward in time.

The inability of all considered models to capture the dynamics in the data is further confirmed by summary plots from Algorithm 2 as depicted in Fig. 10. Additional detailed plots from Algorithm 2 corresponding to specific frequencies are provided in Figure 4 in the Online Supplement. The middle column corresponding to $\omega = 4\pi/64$ confirms that none of the considered models, including the EGARCH model, are able to produce a sufficiently sharp peak in the imaginary part which is observed in the spectrum of the data. The ARCH, GARCH and EGARCH models with Gaussian innovations further struggle to produce the right amount of dependence at central quantile values while the AR(3) process does not have the right kind of dependence in low quantiles. At this point it is natural to wonder whether specifying a different innovation distribution would change any of the finding described above. A natural candidate class of innovations is given by skewed t distributions, and we also considered GARCH(1, 1) and EGARCH(1, 1) processes with such innovation distributions. The

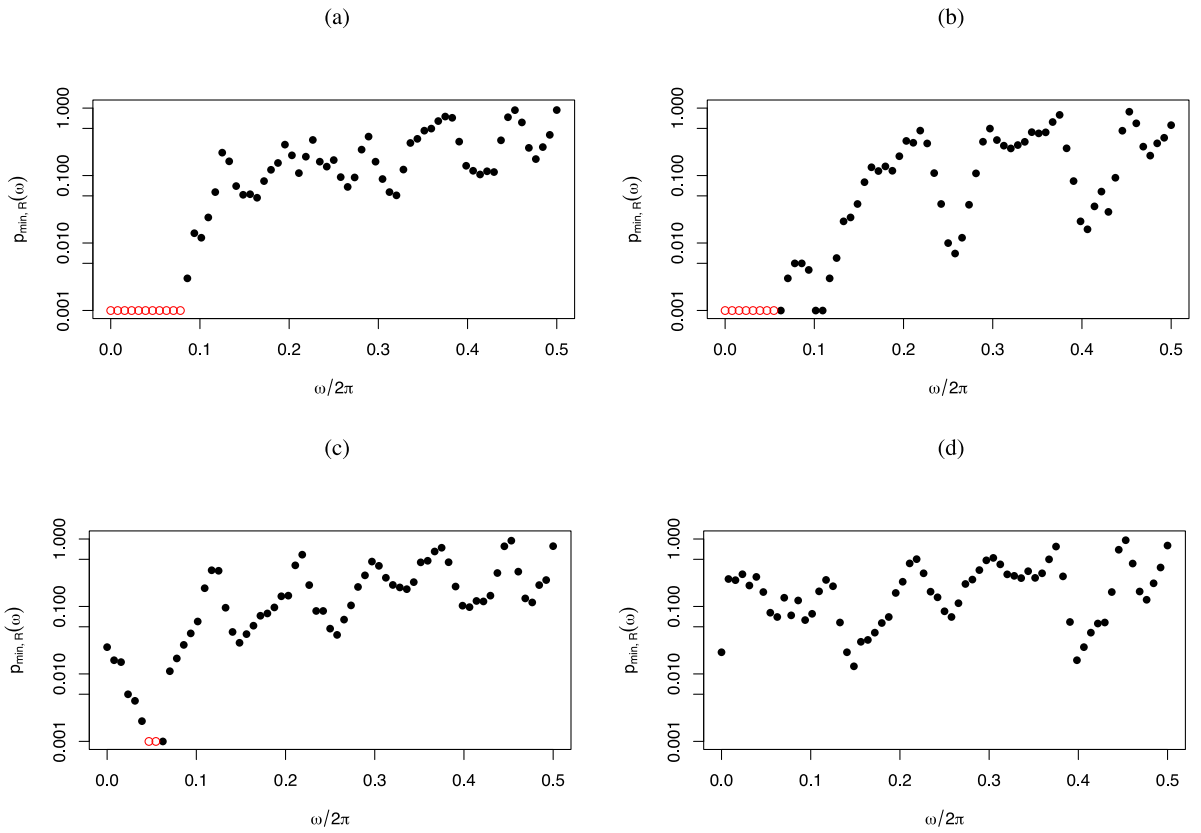


**Fig. 7.** Summary plots produced by Algorithm 2 based on the daily log-returns of the S&P 500 between 2000 and 2005. Plots from top left to bottom right correspond to the following candidate model classes (all with Gaussian innovations) (a) AR(3), (b) ARCH(1), (c) GARCH(1, 1) and (d) EGARCH(1, 1). Frequencies are on the $x$-axis and $p_{\min,R}(\omega)$ defined in Algorithm 2 on the $y$-axis; for better visibility of very low values the $y$-axis is on the log-scale. Red circles correspond to $p_{\min,R}(\omega) < 0.001$.

results did not change much; a more detailed account of our findings is given in Section 1 in the Online Supplement. To what extent other types of innovations could lead to a better model fit is an interesting question that we leave to future research.

### 4.2. Simulation study

In this section we illustrate the finite-sample properties of Algorithms 1 and 2 with simulated data. First, we show that in settings where the data are generated from a model that is contained in the candidate parametric class, the simulated typical regions from Algorithm 1 contain the estimated spectral densities with probability $1-\alpha$ across a range of models, sample sizes and bandwidth parameters; note that this is counted point-wise in $\tau, \omega$. To this end, we consider the following three data generating processes:

$$(a_0)\quad X_t = 0.1X_{t-1} + 0.8Z_{t-1} + Z_t, \quad Z_t \sim \mathcal{N}(0,1)$$
$$(b_0)\quad X_t = 0.2X_{t-1} - 0.4X_{t-2} + 0.2X_{t-3} + Z_t, \quad Z_t \sim \mathcal{N}(0,1)$$
$$(c_0)\quad X_t = \sigma_t Z_t, \text{ where } \sigma_t^2 = 0.01 + 0.4X_{t-1}^2 + 0.5\sigma_{t-1}^2, \quad Z_t \sim \mathcal{N}(0,1)$$

In each case we simulate a time series of length $n \in \{256, 512, 1024\}$ and consider the fixed bandwidth parameters $b_n \in \{0.01, 0.02, 0.05, 0.1, 0.4\}$. For each possible combination, we simulate 1000 repetitions of our algorithm with $\alpha = 0.05$ and the following candidate classes of parametric models:

$$(P_a)\quad X_t = \theta_1 X_{t-1} + \theta_2 Z_{t-1} + Z_t, \quad Z_t \sim \mathcal{N}(0,1)$$
$$(P_b)\quad X_t = \theta_1 X_{t-1} + \theta_2 X_{t-2} + \theta_3 X_{t-3} + Z_t, \quad Z_t \sim \mathcal{N}(0,1)$$
$$(P_c)\quad X_t = \sigma_t Z_t, \text{ where } \sigma_t^2 = \theta_0 + \theta_1 X_{t-1}^2 + \theta_2 \sigma_{t-1}^2, \quad Z_t \sim \mathcal{N}(0,1)$$

Here, $\theta_j$ denote unknown parameters of the models.

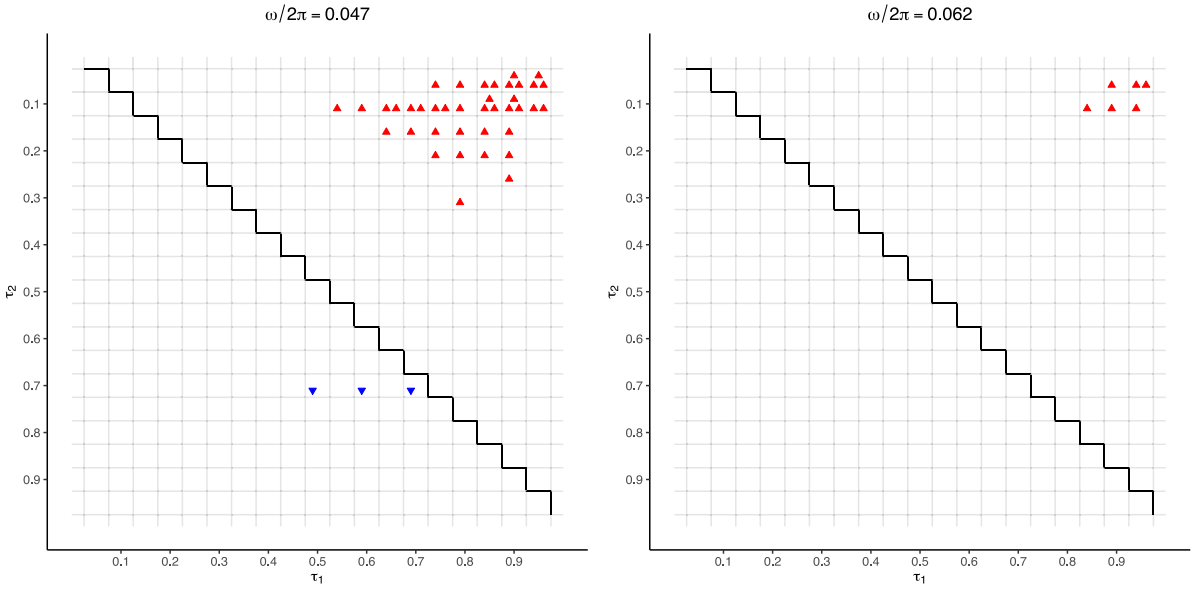


**Fig. 8.** Detailed plots produced by Algorithm 2 at two particular frequencies based on the daily log-returns of the S&P 500 between 2000 and 2005 with GARCH(1, 1) (Gaussian innovations) as candidate model class. Blue triangles facing down indicate that the data-based spectrum is smaller compared to the candidate model spectrum (with 1, 2 and 3 triangles indicating significance at the 5%, 1% and 0.1% level, respectively) while red triangles facing up indicate a data-based spectrum which is larger compared to the candidate model spectrum . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We use the R packages `QPBoot` [5] which contains useful functions for parametric bootstrap procedures for quantile spectra, `quantspec` [22] to compute the estimators for the copula spectral densities, and `rugarch` [15] to estimate and simulate the GARCH-type models. For each frequency $\omega$, we count the number of times the estimated spectral density $\hat{f}_{\boldsymbol{\tau}}(\omega)$ does not lie in the interval $(\ell_{\tau,R}(\omega), u_{\tau,R}(\omega))$ (separately for real and imaginary parts). The resulting counts normalized by 1000 are shown in the top row of Figs. 11 ($n = 256$), 12 ($n = 256$) and 13 ($n = 1024$), respectively. Additional plots for other sample sizes are provided in Section 3 in the Online Supplement. We can see that the simulated typical regions contain the estimator $\hat{f}_{\boldsymbol{\tau}}$ with prescribed probability across a wide range of scenarios.

Next, we show that the aggregated $p$-values obtained from Algorithm 2 are calibrated properly. To this end, we consider the same models and bandwidth parameters as described above and use 1000 simulation replications to approximate the probabilities

$$\Pr\Big[\min_{\boldsymbol{\tau}\in M} \min\{p^{\Re}_{\tau,R}(\omega), p^{\Im}_{\tau,R}(\omega)\} \le \alpha\Big], \tag{4}$$

where we use $M = \{0.05, \ldots, 0.95\}^2$ and $R = 1000$. The results corresponding to models $(a_0)$–$(c_0)$ with candidate model classes $(P_a)$–$(P_c)$ are depicted in the top three rows of Fig. A.14 with frequencies on the $x$-axis and simulated values for the probabilities in (4), with $\alpha = 0.05$, on the $y$-axis. The plots suggest that the $p$-values perform as specified in all settings considered.

Next, consider the case when the observations are created by the following models with $Z_t \sim$ i.i.d. $\mathcal{N}(0, 1)$:

$$(a_1)\quad X_t = 0.2X_{t-1} - 0.4X_{t-2} + 0.2X_{t-3} + Z_t$$
$$(b_1)\quad X_t = \sigma_t Z_t, \text{ where } \sigma_t^2 = 0.01 + 0.4X_{t-1}^2 + 0.5\sigma_{t-1}^2$$
$$(c_1)\quad X_t = \sigma_t Z_t, \text{ where } \ln(\sigma_t^2) = 0.1 + 0.21(|X_{t-1}| - \mathrm{E}|X_{t-1}|) - 0.2X_{t-1} + 0.8\ln(\sigma_{t-1}^2),$$

while the candidate parametric model classes are still $(P_a)$, $(P_b)$, $(P_c)$ and thus are misspecified. The results for Algorithm 1 are depicted in the bottom panels of Figs. 11 ($n = 256$), 12 ($n = 256$), and 13 ($n = 1024$), respectively. Additional plots for other sample sizes are provided in Section 3 in the Online Supplement. The plots corresponding to Algorithm 2 are in rows 4–6 of Fig. A.14.

The results in Fig. 11 and row 4 of Fig. A.14 show that copula spectral densities are informative for distinguishing different types of linear dynamics. Note that in this setting, any of the classical tests that are tailored to linear models would also be applicable and have excellent power properties. Fig. 11 indicates that in this setting, copula spectral densities corresponding to $\boldsymbol{\tau} = (0.5, 0.5), (0.1, 0.5), (0.5, 0.9)$ are most informative. This is not surprising since linear dynamics act similarly in all real parts of copula spectral densities and copula spectral densities corresponding to the quantile values mentioned above are easier to estimate. Note that for more extreme quantiles only a smaller proportion

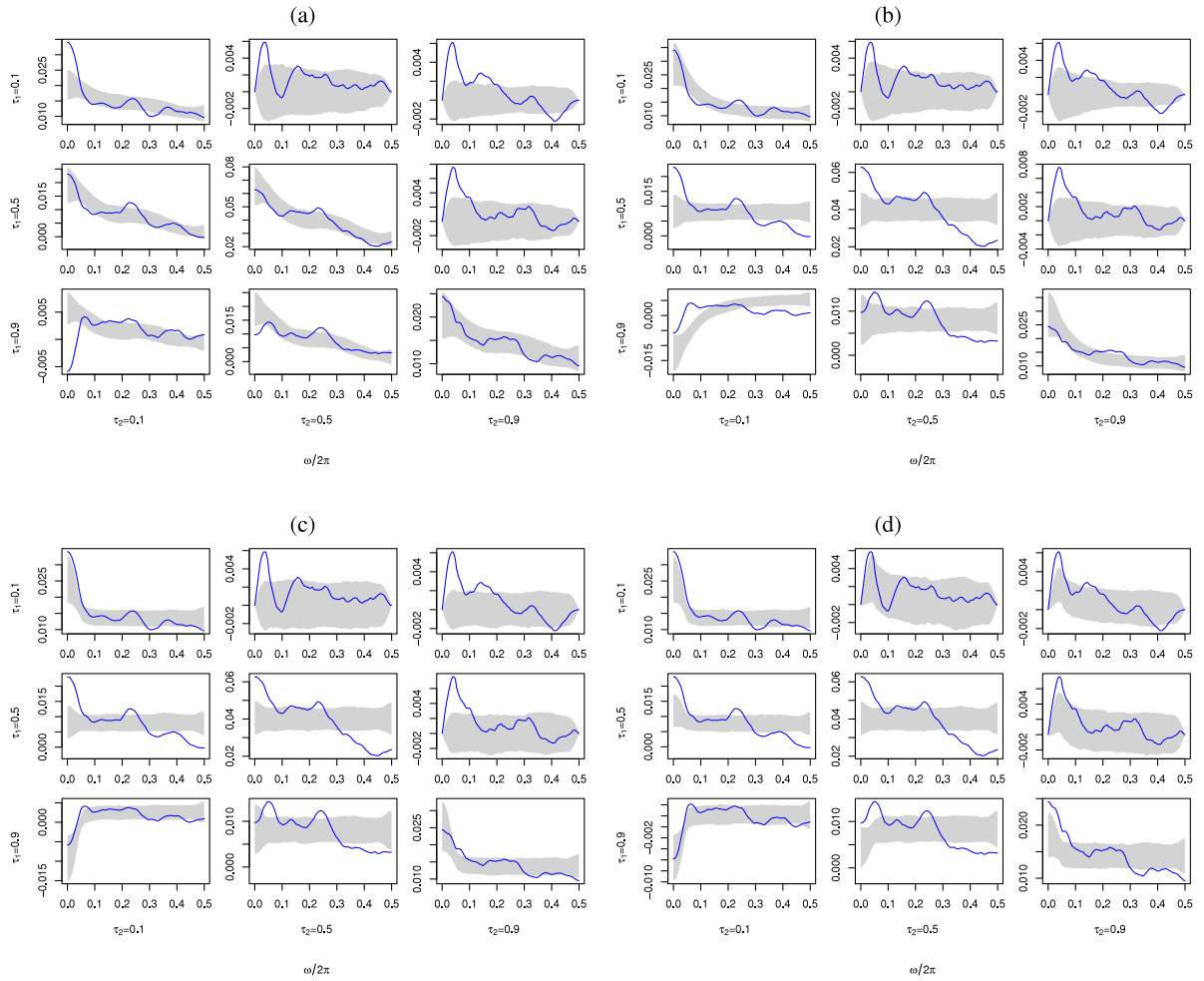


**Fig. 9.** Estimated copula spectral densities based on the daily log-returns of the S&P 500 between 1966 and 1970. Figure displays the plots produced by Algorithm 1 for the following model classes (a) AR(3), (b) ARCH(1), (c) GARCH(1, 1) and (d) EGARCH(1, 1); all models with Gaussian innovations. Blue lines are estimated copula spectral densities, gray areas correspond to typical regions from Algorithm 1 with $\alpha = 0.05$. The plot is organized as discussed in Section 2.1; $\Re\hat{f}_{\boldsymbol{\tau}}$ on/below and $\Im\hat{f}_{\boldsymbol{\tau}}$ above the diagonal, respectively.

of the data carry relevant information. Since linear Gaussian processes are time-reversible, the imaginary parts of copula spectra carry no relevant information in this case. Finally, we remark that for this particular data generation process intermediate bandwidth values lead to the most informative results in Fig. 11. Row 4 of Fig. A.14 additionally shows that aggregating over different frequencies does not lead to a loss in power (despite the uniformity over $\boldsymbol{\tau}$) and in fact improves power for the largest bandwidth $b_n = 0.4$.

Fig. 12 and row 5 of Fig. A.14 show what happens if data are generated by a GARCH model but we attempt to fit their dependence structure by an AR(3) process. In this case the AR(3) model tries to capture the serial correlation of the data, which is zero, so the AR(3) model essentially results in i.i.d data without any serial dependence. This does capture the median dynamics corresponding to $\boldsymbol{\tau} = (0.5, 0.5)$, but completely fails to account for dependence in the more extreme quantiles. This is clearly visible for the real parts of the copula spectral densities corresponding to $\boldsymbol{\tau} = (0.1, 0.1), (0.1, 0.9), (0.9, 0.9)$, with $\boldsymbol{\tau} = (0.1, 0.9)$ leading to the clearest distinction. It is also interesting to observe how different bandwidth values pick up different aspects of the deviation between data and model dynamics. While smaller bandwidth values mainly pick up the sharp peak near zero frequencies, larger bandwidth values also find differences for intermediate frequency values.

The most complicated case that we investigate in our study is to differentiate between a GARCH and an EGARCH process. Results for this are shown in the right panel of Fig. 13 and in the bottom row of Fig. A.14. Both processes have a very similar serial dependence structure as they are uncorrelated but dependent in the extreme quantiles. The difference is that the EGARCH process is asymmetric in the sense, that the dependence is higher in the lower quantiles due to the negative leverage parameter of $-0.2$. This difference is subtle and only present in the dependence at large quantiles and hence difficult to pick up and large sample sizes are needed to reliably pick up this distinction. It also turns out that

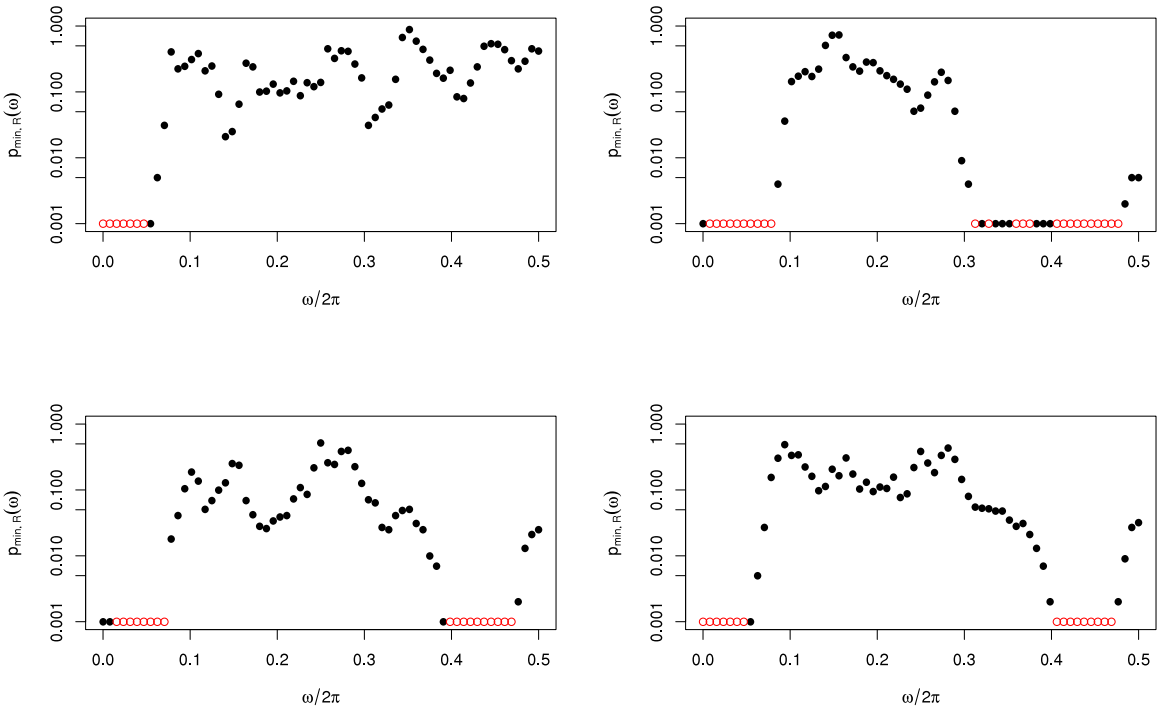


**Fig. 10.** Summary plots produced by Algorithm 2 based on the daily log-returns of the S&P 500 between 1966 and 1970 with AR(3), ARCH(1), GARCH(1, 1) and EGARCH(1, 1) as candidate model class, all models with Gaussian innovations. Frequencies are on the $x$-axis and $p_{\min,R}(\omega)$ (see Algorithm 2) on the $y$-axis; for better visibility of very low values the $y$-axis is in log-scale. Red circles correspond to $p_{\min,R}(\omega) < 0.001$.

the imaginary part corresponding to $\boldsymbol{\tau} = (0.1, 0.9)$ carries the most information here, with larger bandwidth parameters leading to higher probabilities of detecting relevant differences. The results in the bottom row of Fig. A.14 additionally show that by aggregating over different quantile levels we are likely to detect deviations between GARCH and EGARCH processes across a wider range of frequencies. This is due to the fact that for different quantile levels the deviations between the two models are most pronounced at different frequencies.

## Acknowledgments

The authors thank the Guest Editor-in-Chief, Johanna G. Nešlehová, an Associate Editor, three referees, and the participants in the workshop "Dependence Modeling Tools for Risk Management" (Montréal, 2017) for their helpful comments. The research of Stefan Birr was partially supported by project C1 of SFB 823 of the DFG. The research of Tobias Kley was partially supported by the Engineering and Physical Sciences Research Council, UK Grant No. EP/L014246/1. The research of Stanislav Volgushev was partially supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada.

## Appendix. Proofs

**Proof of Example 1.** Note that under the assumptions made $X_t^\theta$ has the representation

$$X_t^\theta = \sum_{k=0}^{\infty} \psi_k^\theta \epsilon_{t-k}$$

where the coefficients are defined, for $|z| \leq 1$, by

$$\sum_{k=0}^{\infty} \psi_k^\theta z^k = Q^\theta(z)/P^\theta(z).$$

By properties of the multivariate normal distribution it suffices to show that, for some $L$,

$$\sup_{\boldsymbol{u}\in\mathbb{R}^2} |F_h^\theta(\boldsymbol{u}) - F_h^{\theta_0}(\boldsymbol{u})| \leq \|\theta - \theta_0\|.$$

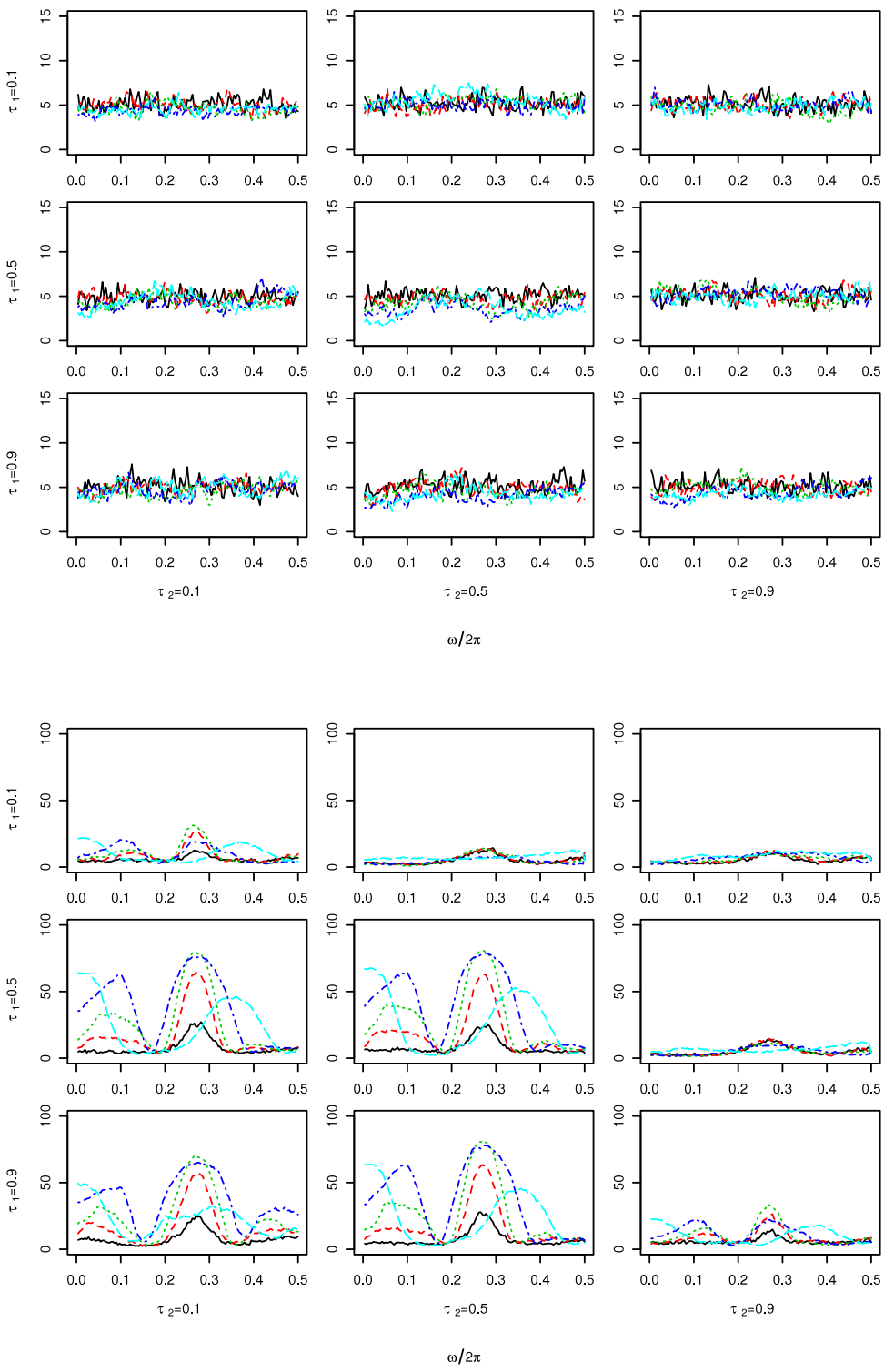


**Fig. 11.** Coverage of the estimator $\hat{f}$ by the critical regions obtained by Algorithm 1. Model class used for the critical regions: $P_a$ (ARMA(1,1)). Data generated according to $(a_0)$ (ARMA(1,1), top row) or $(a_1)$ (AR(3), bottom row). We use $n = 256$ observations. Different bandwidth choices are shown using different colors and line types. The solid line (black), the lines with short (red), medium (green), alternating-length (blue) and long (cyan) dashes correspond to $b_n = 0.01, 0.02, 0.05, 0.1, 0.4$, respectively . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

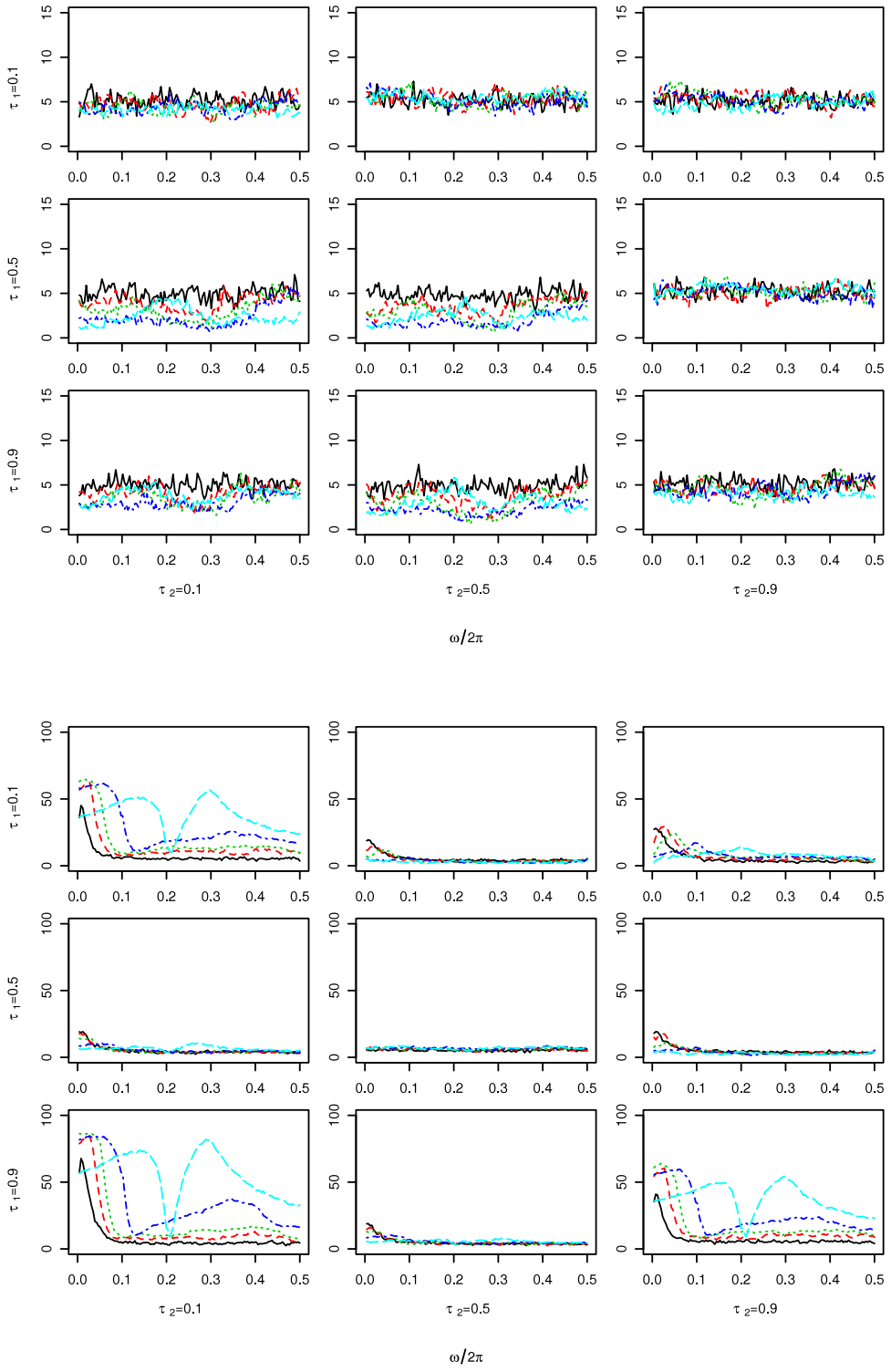


**Fig. 12.** Coverage of the estimator $\hat{f}$ by the critical regions obtained by Algorithm 1. Model class used for the critical regions: $P_b$ (AR(3)). Data generated according to $(b_0)$ (AR(3), top row) and $(b_1)$ (GARCH(1, 1), bottom row). We use $n = 256$ observations. Different bandwidth choices are shown using different colors and line types. The solid line (black), the lines with short (red), medium (green), alternating-length (blue) and long (cyan) dashes correspond to $b_n = 0.01, 0.02, 0.05, 0.1, 0.4$, respectively . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

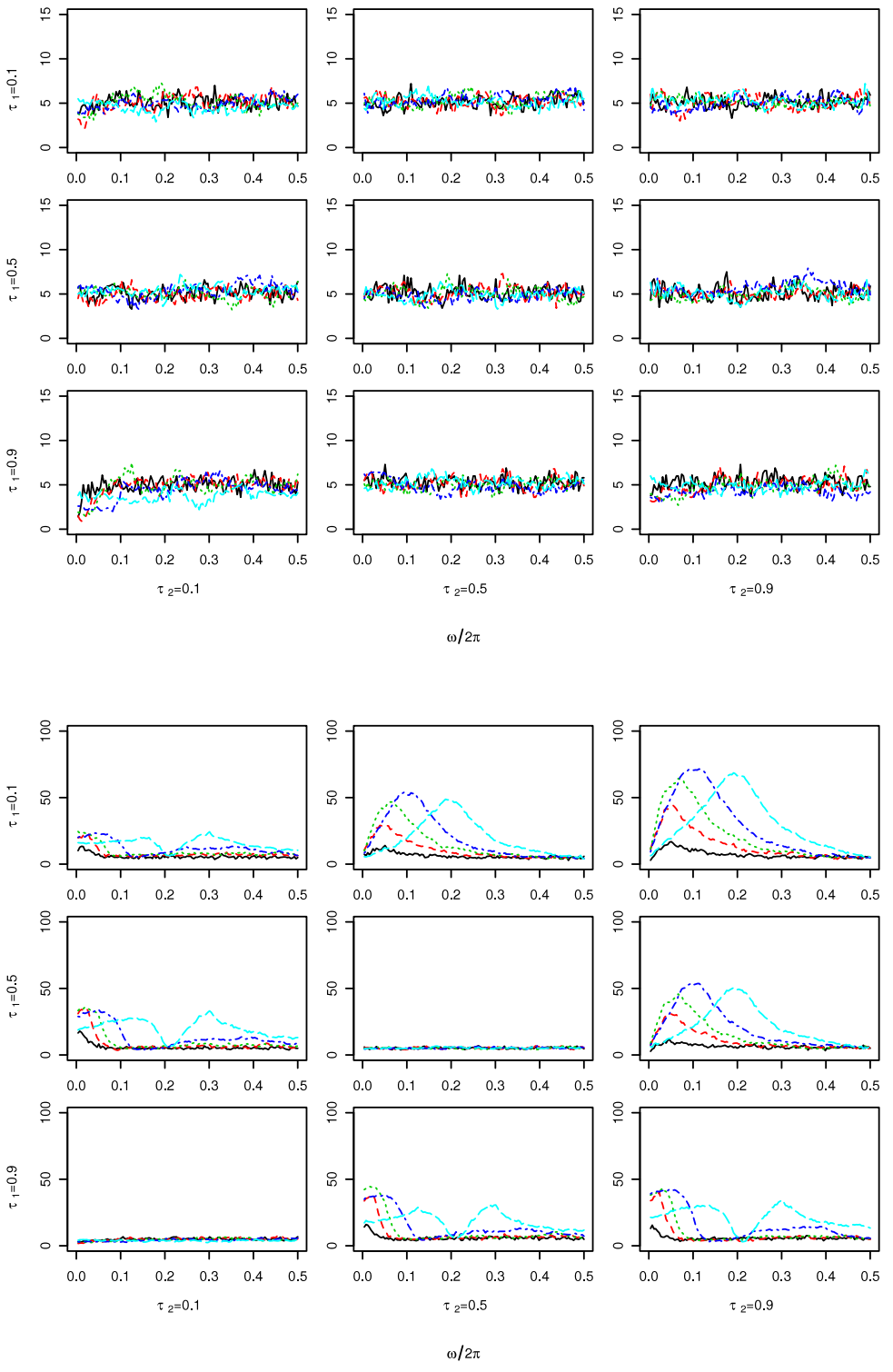


**Fig. 13.** Coverage of the estimator $\hat{f}$ by the critical regions obtained by Algorithm 1. Model class used for the critical regions: $P_c$ (GARCH(1, 1)). Data generated according to ($c_0$) (GARCH(1, 1), top row) and ($c_1$) (EGARCH(1, 1), bottom row). We use $n = 1024$ observations. Different bandwidth choices are shown using different colors and line types. The solid line (black), the lines with short (red), medium (green), alternating-length (blue) and long (cyan) dashes correspond to $b_n = 0.01, 0.02, 0.05, 0.1, 0.4$, respectively . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Applying the triangle inequality, we find

$$\sup_{\boldsymbol{u}\in\mathbb{R}^2} |F_h^\theta(\boldsymbol{u}) - F_h^{\theta_0}(\boldsymbol{u})| \le 2 \sup_{x\in\mathbb{R}} \mathrm{E}|\mathbf{1}(X_0^\theta \le x) - \mathbf{1}(X_0^{\theta_0} \le x)|$$

and hence it suffices to show that

$$\sup_{x\in\mathbb{R}} \mathrm{E}|\mathbf{1}(X_0^\theta \le x) - \mathbf{1}(X_0^{\theta_0} \le x)| \le L\,\|\theta - \theta_0\|.$$

Denote by $\mathcal{A}_t$ the $\sigma$-field generated by $\{\epsilon_s : s < t\}$, and by $F_\epsilon$ the distribution function of $\epsilon$. This yields

$$\begin{aligned}
\sup_{x\in\mathbb{R}} \mathrm{E}|\mathbf{1}(X_0^\theta \le x) - \mathbf{1}(X_0^{\theta_0} \le x)| &= \sup_{x\in\mathbb{R}} \mathrm{E}\Big[\mathrm{E}\{|\mathbf{1}(X_t^\theta \le x) - \mathbf{1}(X_t^{\theta_0} \le x)| \mid \mathcal{A}_t\}\Big] \\
&= \sup_{x\in\mathbb{R}} \mathrm{E}\left[\mathrm{E}\left\{\left|\mathbf{1}\left(\epsilon_t \le x - \sum_{j=1}^\infty \psi_j^\theta \epsilon_{t-j}\right) - \mathbf{1}\left(\epsilon_t \le x - \sum_{j=1}^\infty \psi_j^{\theta_0} \epsilon_{t-j}\right)\right| \,\middle|\, \mathcal{A}_t\right\}\right] \\
&= \sup_{x\in\mathbb{R}} \mathrm{E}\left|F_\epsilon\left(x - \sum_{j=1}^\infty \psi_j^\theta \epsilon_{t-j}\right) - F_\epsilon\left(x - \sum_{j=1}^\infty \psi_j^{\theta_0} \epsilon_{t-j}\right)\right| \le C_1 \mathrm{E}\left|\sum_{j=1}^\infty \psi_j^\theta \epsilon_{t-j} - \sum_{j=1}^\infty \psi_j^{\theta_0} \epsilon_{t-j}\right| \\
&\le C_2 \sum_{j=1}^\infty |\psi_j^\theta - \psi_j^{\theta_0}|.
\end{aligned}$$

Finally, we bound the last term above. To shorten notation we write $p_\theta(z) = Q_\theta(z)/P_\theta(z)$. As $P_{\theta_0}(z)$ has no roots on the unit circle, there exist $\eta, \delta > 0$ such that, for all $\|\theta - \theta_0\| \le \eta$,

$$\forall_{z\in\mathbb{C}:|z|<1+2\delta} \quad P_{\theta_n}(z) \ne 0.$$

Otherwise we could derive a contradiction by using the fact that on $\mathbb{C}$ the locations of roots of a polynomial are a continuous function of the coefficients. Therefore, $p_\theta$ is a holomorphic function on $\{z \in \mathbb{C} : |z| \le 1 + 2\delta\}$ and, by Cauchy's differentiation formula, we can expand $p_\theta(z) = \sum_{j=0}^\infty \psi_j^\theta z^j$ and $p_{\theta_0}(z) = \sum_{j=0}^\infty \psi_j^{\theta_0} z^j$ with

$$\psi_j^\theta = \frac{1}{2\pi i} \oint_{|\zeta|=1+\delta} \frac{p_\theta(\zeta)}{\zeta^{j+1}} d\zeta$$

for all $j \in \mathbb{N}_0$. This implies

$$|\psi_j^\theta - \psi_j^{\theta_0}| = \frac{1}{2\pi} \left| \oint_{|\zeta|=1+\delta} \frac{p_\theta(\zeta) - p_{\theta_0}(\zeta)}{\zeta^{j+1}} d\zeta \right|.$$

And with $p_\theta(z) = Q_\theta(z)/P_\theta(z)$ we have that

$$\sup_{|z|=1+\delta} |\{p_\theta(z) - p_{\theta_0}(z)\}/z^{j+1}| \le \|\theta - \theta_0\|/(1+\delta)^{j+1},$$

which leads to

$$\sum_{j=1}^\infty |\psi_j^\theta - \psi_j^\theta| \le C_3 \|\theta - \theta_0\| \sum_{j=1}^\infty \frac{1}{(1+\delta)^j} = L\|\theta - \theta_0\|. \quad \square$$

**Proof of Proposition 1.** We begin by stating a useful lemma.

**Lemma A.** *Consider a sequence $r_n = o(1)$ and collection of distribution functions $F_{n,\xi}$ indexed by $\xi \in \Xi$, $n \in \mathbb{N}$ such that for any deterministic sequence $\xi_n$ in $\Xi$ with $\xi_n = \xi_0 + O(r_n)$, we have $F_{n,\xi_n} \rightsquigarrow F$ for some distribution function $F$. Then, for any sequence of random variables $\hat{\xi}_n$ in $\Xi$ with $\hat{\xi}_n = \xi_0 + O_p(r_n)$ we have: if $Y_1, \ldots, Y_{m_n}$ are iid $F_{n,\hat{\xi}_n}$ conditional on $\hat{\xi}_n$, $m_n \to \infty$ and $q_n$ denotes the $\alpha$th sample quantile of $Y_1, \ldots, Y_{m_n}$ then $q_n = F^{-1}(\alpha) + o_p(1)$ for any continuity point $\alpha$ of $F^{-1}$.*

**Proof of Lemma A.** Let $\hat{F}_n$ denote the empirical cdf of $Y_1, \ldots, Y_{m_n}$. For any fixed $t \in \mathbb{R}$, we have by the conditional Chebycheff inequality

$$\Pr\left\{|\hat{F}_n(t) - F_{n,\hat{\xi}_n}(t)| \ge \varepsilon | \hat{\xi}_n\right\} \le (4 m_n \varepsilon^2)^{-1} \quad \text{a.s.}$$

Taking the expectation with respect to $\hat{\xi}_n$ shows that $\hat{F}_n(t) - F_{n,\hat{\xi}_n}(t) = o_p(1)$ since by assumption $m_n \to \infty$. Next note that, for arbitrary $C > 0$,

$$\Pr\Big\{|F(t) - F_{n,\hat{\xi}_n}(t)| \ge \varepsilon\Big\} \le \mathbf{1}\Big\{\sup_{|\xi-\xi_0|\le Cr_n} |F(t) - F_{n,\xi}(t)| \ge \varepsilon\Big\} + \Pr(|\hat{\xi}_n - \xi_0| \ge C r_n).$$

We will first show that the first term on the right-hand side converges to zero (for $n \to \infty$) for arbitrary $0 < C < \infty$. Suppose this were not true. Then there exist $\delta > 0$, a subsequence $(n_k)_{k\in\mathbb{N}}$, and $\xi_{n_k} \in \Xi$ with $|\xi_{n_k} - \xi_0| \leq Cr_{n_k}$ and $|F(t) - F_{n_k,\xi_{n_k}}(t)| \geq \delta$ for all $k \in \mathbb{N}$. However, by construction $\xi_{n_k} = \xi_0 + O(r_n)$ (for $k \to \infty$) which contradicts the assumption. Thus for all $C > 0$

$$\limsup_{n\to\infty} \Pr\left\{|F(t) - F_{n,\hat{\xi}_n}(t)| \geq \varepsilon\right\} \leq \limsup_{n\to\infty} \Pr(|\hat{\xi}_n - \xi_0| \geq Cr_n).$$

The right-hand side can be made arbitrarily small by choosing $C$ large since $\hat{\xi}_n = \xi_0 + O_p(r_n)$. Thus we have proved $\hat{F}_n(t) = F(t) + o_p(1)$ for all $t \in \mathbb{R}$.

To complete the proof, observe that $\hat{F}_n$ is a sequence of distribution functions and $F$ is a distribution function. Thus a standard argument implies that $\sup_{t\in\mathbb{R}} |\hat{F}_n(t) - F(t)| = o_p(1)$. This implies $\hat{F}_n^{-1}(\alpha) = F^{-1}(\alpha) + o_p(1)$ for all $\alpha$ that are continuity points of $F^{-1}$; the latter statement follows by the characterization of convergence in probability in terms of a.s. convergence along subsequences and Lemma 21.2 in [29]. □

With the preparations above we are ready to prove Proposition 1.

We begin with the proof of Eq. (2). Recall the setting and notation introduced at the beginning of Section 3. Let $q(\alpha, \theta_0)$ denote the $\alpha$-quantile of the distribution of $\Re H_0(\boldsymbol{\tau}; \omega)$ (where $H_0$ denotes the weak limit in Theorem 1). Define

$$g_n = \Re\{f_{\boldsymbol{\tau}}^{\theta_0}(\omega) + B_n^{(k)}(\boldsymbol{\tau}, \omega)\}$$

for $B_n^{(k)}$ from Theorem 1 and let

$$Z_n = \sqrt{nb_n}\{\Re\hat{f}_{\boldsymbol{\tau}}(\omega) - g_n\}.$$

By Theorem 1 applied to the sequence $\theta_n \equiv \theta_0$, $Z_n \rightsquigarrow \Re H_0(\boldsymbol{\tau}; \omega)$ with the limit being a centered normal random variable with non-zero variance.

Now consider the setting of Lemma A with $m_n = R_n, r_n = n^{-1/2}, \hat{\xi}_n = \hat{\theta}$, $F$ the cdf of $\Re H_0(\boldsymbol{\tau}; \omega)$, $F_{n,\theta}$ the cdf of $\sqrt{nb_n}\{\Re\hat{f}_{\boldsymbol{\tau}}^{\theta}(\omega) - g_n\}$ and $Y_i = \sqrt{nb_n}\{\Re\hat{f}_{\boldsymbol{\tau}}^{\hat{\theta},i}(\omega) - g_n\}$ for all $i \in \{1, \ldots, R_n\}$. Note that $\hat{\theta}_n = \theta_0 + O_p(n^{-1/2})$ by assumption and $F_{n,\theta_n} \rightsquigarrow F$ for any sequence $\theta_n = \theta_0 + O(n^{-1/2})$ by Theorem 1. Hence, all conditions of Lemma A are satisfied and we obtain

$$\sqrt{nb_n}\{\ell_{\boldsymbol{\tau},R_n}^{\Re}(\omega) - g_n\} = q(\alpha/2, \theta_0) + o_p(1).$$

Similarly,

$$\sqrt{nb_n}\{u_{\boldsymbol{\tau},R_n}^{\Re}(\omega) - g_n\} = q(1 - \alpha/2, \theta_0) + o_p(1).$$

By Slutsky's Lemma

$$Z_n - \sqrt{nb_n}\{u_{\boldsymbol{\tau},R_n}^{\Re}(\omega) - g_n\} \rightsquigarrow \Re H_0(\boldsymbol{\tau}; \omega) - q(1 - \alpha/2, \theta_0),$$

and since the distribution of the limit is continuous

$$\Pr\{\Re\hat{f}_{\boldsymbol{\tau}}(\omega) \leq u_{\boldsymbol{\tau},R_n}^{\Re}(\omega)\} = \Pr[Z_n \leq \sqrt{nb_n}\{u_{\boldsymbol{\tau},R_n}^{\Re}(\omega) - g_n\}] \to 1 - \alpha/2.$$

Similarly,

$$\Pr\{\Re\hat{f}_{\boldsymbol{\tau}}(\omega) < \ell_{\boldsymbol{\tau},R_n}^{\Re}(\omega)\} = \Pr[Z_n < \sqrt{nb_n}\{\ell_{\boldsymbol{\tau},R_n}^{\Re}(\omega) - g_n\}] \to \alpha/2.$$

This completes the proof of Eq. (2).

Next let us prove Eq. (3). Begin by observing that

$$x \mapsto 1 - \hat{F}_R(x-) = \frac{1}{R}\sum_{r=1}^{R} \mathbf{1}[x \leq \max\{A_r^{\Re}(\omega), A_r^{\Im}(\omega)\}]$$

is non-increasing, so

$$\min_{\boldsymbol{\tau}\in M} \min\{p_{\boldsymbol{\tau},R}^{\Re}(\omega), p_{\boldsymbol{\tau},R}^{\Im}(\omega)\} = \frac{1}{R}\sum_{r=1}^{R} \mathbf{1}\left[\max_{\boldsymbol{\tau}\in M} \max\{E_{\boldsymbol{\tau}}^{\Re}(\omega), E_{\boldsymbol{\tau}}^{\Im}(\omega)\} \leq \max\{A_r^{\Re}(\omega), A_r^{\Im}(\omega)\}\right].$$

Let

$$Z_{r,n}(\boldsymbol{\tau}, \omega) = \sqrt{nb_n}\left\{\hat{f}_{\boldsymbol{\tau}}^{\hat{\theta},r}(\omega) - f_{\boldsymbol{\tau}}^{\theta_0}(\omega) - B_n^{(k)}(\boldsymbol{\tau}, \omega)\right\}$$

where $B_n^{(k)}$ is defined in Theorem 1. Define

$$\widetilde{\ell}_{\boldsymbol{\tau},R}^{\Re}(\omega) = \beta/2 - \text{quantile}\{\Re Z_{r,n}(\boldsymbol{\tau}, \omega), \ldots, \Re Z_{r,n}(\boldsymbol{\tau}, \omega)\},$$
$$\widetilde{u}_{\boldsymbol{\tau},R}^{\Re}(\omega) = (1 - \beta/2) - \text{quantile}\{\Re Z_{r,n}(\boldsymbol{\tau}, \omega), \ldots, \Re Z_{r,n}(\boldsymbol{\tau}, \omega)\}$$

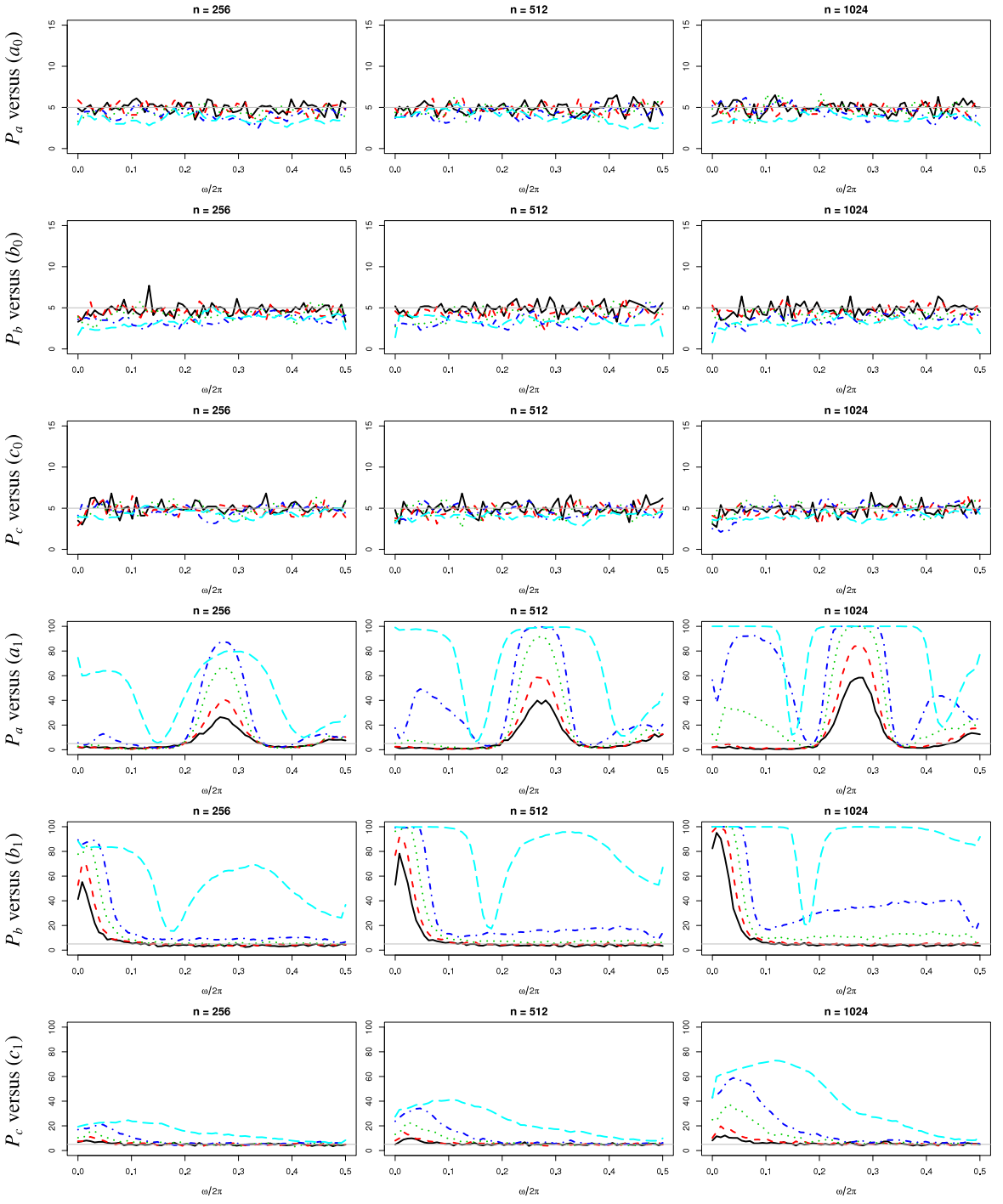


**Fig. A.14.** Proportion of cases, per frequency, where $p$-values obtained from Algorithm 2 are below $\alpha = 0.05$, for at least one quantile level. First, second and third rows show $P_a$ versus $(a_0)$ (ARMA(1,1)), $P_b$ versus $(b_0)$ (AR(3)), and $P_c$ versus $(c_0)$ (GARCH(1, 1)), respectively. Fourth, fifth and sixth rows show $P_a$ versus $(a_1)$ (ARMA(1,1) and AR(3)), $P_b$ versus $(b_1)$ (AR(3) and GARCH(1, 1)), and $P_c$ versus $(c_1)$ (GARCH(1, 1) versus EGARCH(1, 1)), respectively. We use $n = 256, 512$ and 1024 observations in the first, second and third columns, respectively. Different bandwidth choices are shown using different colors and line types. The solid line (black), the lines with short (red), medium (green), alternating-length (blue) and long (cyan) dashes correspond to $b_n = 0.01, 0.02, 0.05, 0.1, 0.4$, respectively . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and similarly for imaginary parts. Let $\widetilde{\Delta}^{\Re}_{\boldsymbol{\tau},R}, \widetilde{c}^{\Re}_{\boldsymbol{\tau},R}, \widetilde{A}^{\Re}_r(\omega)$ denote the corresponding versions of $\Delta^{\Re}_{\boldsymbol{\tau},R}, c^{\Re}_{\boldsymbol{\tau},R}, A^{\Re}_r(\omega)$ with $\widetilde{u}^{\Re}_{\boldsymbol{\tau},R}, \widetilde{\ell}^{\Re}_{\boldsymbol{\tau},R}, \Re Z_{r,n}(\boldsymbol{\tau},\omega)$ instead of $u^{\Re}_{\boldsymbol{\tau},R}, \ell^{\Re}_{\boldsymbol{\tau},R}, \hat{f}^{\theta,r}_{\boldsymbol{\tau}}(\omega)$ and note that by equivariance of quantiles under the given transformations, we have

$$\widetilde{u}^{\Re}_{\boldsymbol{\tau},R} = \sqrt{nb_n}\,\{u^{\Re}_{\boldsymbol{\tau},R} - \Re f^{\theta_0}_{\boldsymbol{\tau}}(\omega) - \Re B^{(k)}_n(\boldsymbol{\tau},\omega)\}, \quad \widetilde{\ell}^{\Re}_{\boldsymbol{\tau},R} = \sqrt{nb_n}\,\{\ell^{\Re}_{\boldsymbol{\tau},R} - \Re f^{\theta_0}_{\boldsymbol{\tau}}(\omega) - \Re B^{(k)}_n(\boldsymbol{\tau},\omega)\}$$

which implies $\widetilde{A}^{\Re}_r(\omega) \equiv A^{\Re}_r(\omega)$ after some simple algebra. From Lemma A we obtain by similar arguments as above (noting that $M$ is finite) that $\widetilde{\ell}^{\Re}_{\boldsymbol{\tau},R}(\omega)$ converges to the $\beta/2$-quantile of the distribution of $\Re H_0(\boldsymbol{\tau},\omega)$ and $\widetilde{u}^{\Re}_{\boldsymbol{\tau},R}(\omega)$ converges to the $1-\beta/2$-quantile of the distribution of $\Re H_0(\boldsymbol{\tau},\omega)$ (both convergences are in probability). Since $\Re H_0(\boldsymbol{\tau},\omega)$ follows a normal distribution with non-zero variance this implies

$$\max_{\boldsymbol{\tau}\in M} |\widetilde{\Delta}^{\Re}_{\boldsymbol{\tau},R}(\omega) - \Delta^{\Re}_{\boldsymbol{\tau}}(\omega)| = o_p(1), \quad \max_{\boldsymbol{\tau}\in M} |\widetilde{c}^{\Re}_{\boldsymbol{\tau},R}(\omega)| = o_p(1)$$

where $\Delta^{\Re}_{\boldsymbol{\tau}}(\omega) = \sigma^{\Re}_{\boldsymbol{\tau}}(\omega)\{1/2 - \Phi^{-1}(\beta/2)\}$ and $\sigma^{\Re}_{\boldsymbol{\tau}}(\omega)$ denotes the standard deviation of $\Re H_0(\boldsymbol{\tau},\omega)$. Similar results hold for the imaginary parts. By a combination of Slutsky's Lemma and the Continuous Mapping Theorem, we now obtain from Theorem 1 that

$$\max\{\widetilde{A}^{\Re}_1(\omega), \widetilde{A}^{\Im}_1(\omega)\} \rightsquigarrow \frac{1}{1/2 - \Phi^{-1}(\beta/2)} \max_{\boldsymbol{\tau}\in M} \max\left\{\frac{\Re H_0(\boldsymbol{\tau};\omega)}{\sigma^{\Re}_{\boldsymbol{\tau}}(\omega)}, \frac{\Im H_0(\boldsymbol{\tau};\omega)}{\sigma^{\Im}_{\boldsymbol{\tau}}(\omega)}\right\}. \tag{A.1}$$

By similar arguments it follows that

$$\max_{\boldsymbol{\tau}\in M} \max\{E^{\Re}_{\boldsymbol{\tau}}(\omega), E^{\Im}_{\boldsymbol{\tau}}(\omega)\} \rightsquigarrow \max_{\boldsymbol{\tau}\in M} \max\left\{\frac{\Re H_0(\boldsymbol{\tau};\omega)}{\sigma^{\Re}_{\boldsymbol{\tau}}(\omega)}, \frac{\Im H_0(\boldsymbol{\tau};\omega)}{\sigma^{\Im}_{\boldsymbol{\tau}}(\omega)}\right\} \sim F. \tag{A.2}$$

Denoting by $F_R$ the cdf of the random variable $\max\{\widetilde{A}^{\Re}_1(\omega), \widetilde{A}^{\Im}_1(\omega)\}$, the uniform Glivenko–Cantelli Theorem (see Theorem 2.8.1 in [30]) implies that $\sup_{x\in\mathbb{R}} |\hat{F}_R(x) - F_R(x)| = o_p(1)$. Together with (A.1) and continuity of the cdf, say $F$, of the random variable $\max_{\boldsymbol{\tau}\in M} \max\{\Re H_0(\boldsymbol{\tau};\omega)/\sigma^{\Re}_{\boldsymbol{\tau}}(\omega), \Im H_0(\boldsymbol{\tau};\omega)/\sigma^{\Im}_{\boldsymbol{\tau}}(\omega)\}$ (note that the latter is a maximum over a finite number of (dependent) standard normal random variables, hence has a continuous distribution), it follows that $\sup_{x\in\mathbb{R}} |F_R(x) - F(x)| = o(1)$, and hence

$$\frac{1}{R}\sum_{r=1}^{R} \mathbf{1}\left[\max_{\boldsymbol{\tau}\in M} \max\{E^{\Re}_{\boldsymbol{\tau}}(\omega), E^{\Im}_{\boldsymbol{\tau}}(\omega)\} \le \max\{A^{\Re}_r(\omega), A^{\Im}_r(\omega)\}\right] = 1 - F\left[\max_{\boldsymbol{\tau}\in M} \max\{E^{\Re}_{\boldsymbol{\tau}}(\omega), E^{\Re}_{\boldsymbol{\tau}}(\omega)\}\right] + o_p(1).$$

Now, by (A.2) and by continuity of $F$ combined with the Continuous Mapping Theorem and Slutsky's Lemma, we finally obtain

$$\min_{\boldsymbol{\tau}\in M} \min\left\{p^{\Re}_{\boldsymbol{\tau},R}(\omega), p^{\Im}_{\boldsymbol{\tau},R}(\omega)\right\} = \frac{1}{R}\sum_{r=1}^{R} \mathbf{1}\left[\max_{\boldsymbol{\tau}\in M} \max\{E^{\Re}_{\boldsymbol{\tau}}(\omega), E^{\Im}_{\boldsymbol{\tau}}(\omega)\} \le \max\{A^{\Re}_r(\omega), A^{\Im}_r(\omega)\}\right] \rightsquigarrow 1 - \mathcal{U}[0,1].$$

This completes the proof of (3). □

**Proof of Theorem 1.** We will make use of the following notation: $U^{\theta}_t = F^{\theta}(X^{\theta}_t)$

$$d^{U,\theta}_{\tau,n}(\omega) = \sum_{t=0}^{n-1} \mathbf{1}(U^{\theta}_t \le \tau) e^{-i\omega t}, \quad I^{U,\theta}_{(\tau_1,\tau_2),n}(\omega) = \frac{1}{2\pi n} d^{U,\theta}_{\tau_1,n}(\omega) d^{U,\theta}_{\tau_2,n}(-\omega),$$

and

$$\hat{f}^{U,\theta}_{\boldsymbol{\tau},n}(\omega) = \frac{1}{2\pi n}\sum_{s=1}^{n-1} W_n(\omega - 2\pi s/n) I^{U,\theta}_{\boldsymbol{\tau},n}(2\pi s/n).$$

Theorem 1 follows from the following four statements.

(i) For any fixed $\omega \in \mathbb{R}$ and an arbitrary sequence $\theta_n$ in $\Theta$ with $\theta_n = \theta_0 + o(1)$ we have, as $n \to \infty$, in $\ell^{\infty}(\mathcal{T})$,

$$\sqrt{nb_n}\,\{\hat{f}^{U,\theta_n}_{\boldsymbol{\tau},n}(\omega) - \mathrm{E}\hat{f}^{U,\theta_n}_{\boldsymbol{\tau},n}(\omega)\}_{\tau\in\mathcal{T}} \rightsquigarrow H_0(\cdot;\omega).$$

(ii) As $n \to \infty$, we obtain the following result for the expectation:

$$\sup_{\substack{\tau\in[0,1]^2\\ \omega\in\mathbb{R}}} |\mathrm{E}\hat{f}^{U,\theta_0}_{\boldsymbol{\tau},n}(\omega) - f^{\theta_0}_{\boldsymbol{\tau}}(\omega) - B^{(k)}_n(\boldsymbol{\tau},\omega)| = O\{(nb_n)^{-1}\} + o(b^k_n),$$

(iii) For any fixed $\omega$

$$\sup_{\boldsymbol{\tau}\in[0,1]^2} |\hat{f}^{\theta_n}_{\boldsymbol{\tau}}(\omega) - \hat{f}^{U,\theta_n}_{\boldsymbol{\tau},n}(\omega)| = o_p\{(nb_n)^{-1/2} + b^k_n\},$$

(iv) For any sequence $\theta_n$ in $\Theta$ with $\theta_n = \theta_0 + O(n^{-1/2})$ we have, as $n \to \infty$,

$$\sqrt{nb_n} \sup_{\boldsymbol{\tau}\in\mathcal{T},\omega\in\mathbb{R}} \left|\mathrm{E}\hat{f}^{U,\theta_n}_{\boldsymbol{\tau},n}(\omega) - \mathrm{E}\hat{f}^{U,\theta_0}_{\boldsymbol{\tau},n}(\omega)\right| = o(1).$$

Note that (ii) is proved in Theorem 3.6(ii) in [23] so that it remains to prove (i), (iii), (iv).

**Convergence as a process and the proof of (i)**. In what follows, let

$$\Delta_n(\omega) = \sum_{t=0}^{n-1} e^{i\omega t}, \quad \mathcal{F}_n = \{2\pi j/n : j \in \{1,\dots,n-1\}\}.$$

For intervals $A \subset [0,1]$ and values $\omega \in \mathbb{R}$ define

$$d^{U,\theta_n}_{A,n}(\omega) = \sum_{t=0}^{n-1} \mathbf{1}(U^{\theta_n}_t \in A)e^{-it\omega}.$$

Let

$$H_n(\boldsymbol{\tau},\omega) = \sqrt{nb_n}\{\hat{f}^{U,\theta_n}_{\boldsymbol{\tau},n}(\omega) - \mathrm{E}\hat{f}^{U,\theta_n}_{\boldsymbol{\tau},n}(\omega)\}$$

and denote by $\hat{F}^{U,\theta_n}_n$ the empirical cdf of $U^{\theta_n}_1,\dots,U^{\theta_n}_n$.

We begin by stating several generalizations of the results in [23]. The proofs of those results are very similar to the corresponding proofs in [23] and are omitted for the sake of brevity.

The following statement can be proved similarly to Lemma A.4 in [23]: for arbitrary intervals $A_1,\dots,A_p \subset [0,1]$ define $\varepsilon = \min_{1\le j\le p}\lambda(A_j)$. Then there exist constants $C$ and $d$ that depend only on $p$ and $K_2,\dots,K_p$ and $\rho_2,\dots,\rho_p$ from assumption (C) such that, for all $\omega_1,\dots,\omega_p \in \mathbb{R}$,

$$|\mathrm{cum}\{d^{U,\theta_n}_{A_1,n}(\omega_1),\dots,d^{U,\theta_n}_{A_p,n}(\omega_p)\}| \le C\{|\Delta_n(\omega_1+\cdots+\omega_p)|+1\}\varepsilon\{|\ln(\varepsilon)|+1\}^d. \tag{A.3}$$

Utilizing this statement and following the proof of Lemma 1.6 in the Online Supplement of [23], we find that, for any $k \in \mathbb{N}$, there exists a constant $d_k$ such that, as $\delta_n \to 0$,

$$\sup_{\substack{x,y\in[0,1]\\|x-y|\le\delta_n}} \sqrt{n}|\hat{F}^{U,\theta_n}_n(x) - \hat{F}^{U,\theta_n}_n(y) - x + y| = O_p\{(n^2\delta_n+n)^{1/2k}(\delta_n|\ln\delta_n|^{d_k}+n^{-1})^{1/2}\}. \tag{A.4}$$

This equation combined with the arguments in the proof of Lemma A.6 in [23] shows that, for any $k \in \mathbb{N}$,

$$\sup_{\omega\in\mathcal{F}_n}\sup_{\tau\in[0,1]} |d^{U,\theta_n}_{\tau,n}(\omega)| = O_p(n^{1/2+1/k}). \tag{A.5}$$

Now Eqs. (A.3), (A.4) and (A.5) can be used to replace Lemma A.2, Lemma 1.6 (Online Supplement), and Lemma A.6 from [23] in the proof of Lemma A.7 in [23] to show the following: if $\delta_n = O\{(nb_n)^{-1/\gamma}\}$ for some $\gamma \in (0,1)$, then

$$\sup_{\omega\in\mathbb{R}}\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^2,\,\|\boldsymbol{u}-\boldsymbol{v}\|\le\delta_n} |H_n(\boldsymbol{u},\omega) - H_n(\boldsymbol{v},\omega)| = o_p(1). \tag{A.6}$$

We are now ready for the proof of (i). In view of Theorems 1.5.4 and 1.5.7 in [30], it suffices to show

(i1) Convergence of the finite-dimensional distributions

$$\{H_n(\boldsymbol{\tau}_j,\omega_j)\}_{j=1,\dots,k} \rightsquigarrow \{H_0(\boldsymbol{\tau}_j,\omega_j)\}_{j=1,\dots,k}$$

for any $(\boldsymbol{\tau}_j,\omega_j) \in \mathcal{T}\times\mathbb{R}$ with $j \in \{1,\dots,k\}$ and $k \in \mathbb{N}$.

(i2) Stochastic equicontinuity: For any $x > 0$ and any $\omega \in \mathbb{R}$,

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} \Pr\left\{\sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^2,\|\boldsymbol{u}-\boldsymbol{v}\|<\delta} |H_n(\boldsymbol{u},\omega) - H_n(\boldsymbol{v},\omega)| > x\right\} = 0.$$

**Proof of (i2).** Apply Lemma A.2 from [23] with $L = 3$ to obtain

$$\sup_{\omega\in\mathbb{R}}\sup_{\|\boldsymbol{u}-\boldsymbol{v}\|_1\le\epsilon} \mathrm{E}|H_n(\boldsymbol{u},\omega) - H_n(\boldsymbol{v},\omega)|^6 \le K\sum_{j=0}^{2} g(\epsilon)^{3-j}/(nb_n)^j.$$

Here, condition (A.2) from Lemma A.2 in [23] is satisfied with $g(x) = x\{|\ln x|+1\}^d$ by (A.3). With $\Psi(x) = x^6$ the Orlicz norm $\|X\|_\Psi$ coincides with the $L_6$ norm $\|X\|_6 = (\mathrm{E}|X|^6)^{1/6}$ so that we have, for any $\kappa \in (0,1)$ and sufficiently small $\|a-b\|_1$,

$$\|H_n(\boldsymbol{u},\omega) - H_n(\boldsymbol{v},\omega)\|_\Psi \le C\{\|\boldsymbol{u}-\boldsymbol{v}\|_1^\kappa/(nb_n)^2 + \|\boldsymbol{u}-\boldsymbol{v}\|_1^{2\kappa}/(nb_n)^1 + \|\boldsymbol{u}-\boldsymbol{v}\|_1^{3\kappa}\}.$$

To complete the proof of (i2), follow the arguments in the proof of Theorem 3.6, step (i2), in [23]. Replace Lemma A.7 therein by (A.6) to obtain for all $x, \mu > 0$ and $\gamma \in (2/3, 1)$,

$$\lim_{\delta \downarrow 0} \limsup_{n\to\infty} \Pr\Big\{ \sup_{\boldsymbol{u},\boldsymbol{v}\in[0,1]^2} \sup_{\|\boldsymbol{u}-\boldsymbol{v}\|_1<\delta} |H_n(\boldsymbol{u},\omega) - H_n(\boldsymbol{v},\omega)| > x \Big\} \le \Big\{ \frac{8C}{x} \int_0^\mu z^{-2/(3\gamma)} dz \Big\}^6.$$

Then (i2) follows since the integral tends to zero for $\mu \to 0$.

**Proof of (i1).** We have to show, that for any $\boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_k \in \mathcal{T}$, $k \in \mathbb{N}$ where $\boldsymbol{\tau}_i = (\tau_{i1}, \tau_{i2})$ and $\omega_1, \dots, \omega_k \neq 0 \mod 2\pi$ all cumulants of $\{H_n(\boldsymbol{\tau}_j, \omega_j)\}_{j=1,\dots,k}$ converge to the corresponding cumulants of $\{H_0(\boldsymbol{\tau}_j, \omega_j)\}_{j=1,\dots,k}$, which by Lemma P4.5 in [7] gives the desired result. By construction,

$$\mathrm{E}\{H_n(\boldsymbol{\tau}, \omega)\} = 0$$

and

$$\mathrm{cov}\{H_n(\boldsymbol{\tau}_1, \omega_1), H_n(\boldsymbol{\tau}_2, \omega_2)\} = nb_n \mathrm{cov}\{\hat{f}^{U,\theta_n}_{\boldsymbol{\tau}_1,n}(\omega_1), \hat{f}^{U,\theta_n}_{\boldsymbol{\tau}_2,n}(\omega_2)\}.$$

Under assumption (C) the random processes $\{\mathbf{1}(U_t^{\theta_n} \le \tau_{11}), \dots, \mathbf{1}(U_t^{\theta_n} \le \tau_{k2})\}_{t\in\mathbb{Z}}$ satisfy a uniform version of Assumption 2.6.2(2) in [7] while the weight functions $W_n$ satisfy Assumption 5.6.1 in [7]. A close look at the proof of Theorem 7.4.3 and Corollary 7.4.3 in [7] shows that all proofs go through without change and lead to the representation

$$\begin{aligned} nb_n \mathrm{cov}\{\hat{f}^{U,\theta_n}_{\boldsymbol{\tau}_1,n}(\omega_1), \hat{f}^{U,\theta_n}_{\boldsymbol{\tau}_2,n}(\omega_2)\} = {} & 2\pi \int_{-\pi}^{\pi} W^2(u)\mathrm{d}u \Big\{ f^{\theta_n}_{(\tau_{11},\tau_{21})}(\omega_1) f^{\theta_n}_{(\tau_{12},\tau_{22})}(-\omega_1)\mathbf{1}(\omega_1 = \omega_2) \\ & + f^{\theta_n}_{(\tau_{11},\tau_{22})}(\omega_1) f^{\theta_n}_{(\tau_{12},\tau_{21})}(-\omega_1)\mathbf{1}(\omega_1 = 2\pi - \omega_2) \Big\} + O(b_n) + O\{(nb_n)^{-1}\}. \end{aligned}$$

Next we note that

$$\sup_{\boldsymbol{\tau}\in\mathcal{T},\omega\in\mathbb{R}} |f^{\theta_n}_{\boldsymbol{\tau}}(\omega) - f^{\theta_0}_{\boldsymbol{\tau}}(\omega)| \le \sum_{h\in\mathbb{Z}} \sup_{\boldsymbol{\tau}\in\mathcal{T}} |C^{\theta_n}_h(\boldsymbol{\tau}) - C^{\theta_0}_h(\boldsymbol{\tau})| = o(1)$$

by dominated convergence. Hence

$$\begin{aligned} nb_n \mathrm{cov}\{\hat{f}^{U,\theta_n}_{\boldsymbol{\tau}_1,n}(\omega_1), \hat{f}^{U,\theta_n}_{\boldsymbol{\tau}_2,n}(\omega_2)\} \to {} & 2\pi \int_{-\pi}^{\pi} W^2(u)\mathrm{d}u \Big\{ f^{\theta_0}_{(\tau_{11},\tau_{21})}(\omega_1) f^{\theta_0}_{(\tau_{12},\tau_{22})}(-\omega_1)\mathbf{1}(\omega_1 = \omega_2) \\ & + f^{\theta_0}_{(\tau_{11},\tau_{22})}(\omega_1) f^{\theta_0}_{(\tau_{12},\tau_{21})}(-\omega_1)\mathbf{1}(\omega_1 = 2\pi - \omega_2) \Big\}. \end{aligned}$$

To complete the proof, it remains to show that the cumulants of order $K \ge 3$ vanish as $n \to \infty$. With $\boldsymbol{\tau}_i = (\tau_{i1}, \tau_{i2})$ for all $i \in \{1, \dots, K\}$, we have

$$\begin{aligned} \mathrm{cum}\{H_n(\boldsymbol{\tau}_1, \omega_1), \dots, H_n(\boldsymbol{\tau}_K, \omega_K)\} = {} & (nb_n)^{K/2} \mathrm{cum}\{\hat{f}^{U,\theta_n}_{\boldsymbol{\tau}_1,n}(\omega_1), \dots, \hat{f}^{U,\theta_n}_{\boldsymbol{\tau}_K,n}(\omega_K)\} \\ = {} & (2\pi)^{2K} n^{-3K/2} (b_n)^{K/2} \sum_{s_1=1}^{n-1} \cdots \sum_{s_K=1}^{n-1} W_n(\omega_1 - 2\pi s_1/n) \cdots W_n(\omega_k - 2\pi s_K/n) \\ & \times \mathrm{cum}\Big\{ d^{U,\theta_n}_{\tau_{11},n}(2\pi s_1/n) d^{U,\theta_n}_{\tau_{12},n}(-2\pi s_1/n), \dots, d^{U,\theta_n}_{\tau_{K1},n}(2\pi s_K/n) d^{U,\theta_n}_{\tau_{K2},n}(-2\pi s_K/n) \Big\}. \end{aligned}$$

To see that these cumulants tend to zero we will need arguments similar to those used in Step 2 of the proof of Lemma A.2 in [23]. Applying the product Theorem 2.3.2 in [7] to the last cumulant leads to

$$\begin{aligned} & \mathrm{cum}\Big\{ d^{U,\theta_n}_{\tau_{11},n}(2\pi s_1/n) d^{U,\theta_n}_{\tau_{12},n}(-2\pi s_1/n), \dots, d^{U,\theta_n}_{\tau_{K1},n}(2\pi s_K/n) d^{U,\theta_n}_{\tau_{K2},n}(-2\pi s_K/n) \Big\} \\ & = \sum_{(\nu_1,\dots,\nu_N)} \prod_{k=1}^{N} \mathrm{cum}\Big[ d^{U,\theta_n}_{\tau_{ij},n}\{(-1)^{j+1} s_i 2\pi/n\}; (i,j) \in \nu_k \Big], \end{aligned} \tag{A.7}$$

where the sum runs over all indecomposable partitions $(\nu_1, \dots, \nu_N)$ (see [7], p. 20) of

$$\begin{matrix} (1,1) & (1,2) \\ \vdots & \vdots \\ (K,1) & (K,2). \end{matrix}$$

Note that an indecomposable partition consists of at most $N \le K + 1$ sets. Now by (A.3) the absolute values of those cumulants are bounded by

$$\text{(A.7)} \le C \sum_{(\nu_1,\dots,\nu_N)} \prod_{k=1}^{N} \Big[ \Delta_n \Big\{ \frac{2\pi}{n} \sum_{(i,j)\in\nu_k} (-1)^{j+1} s_i \Big\} + 1 \Big] = C \sum_{(\nu_1,\dots,\nu_N)} \sum_{I\subset\{1,\dots,N\}} \prod_{k\in I} \Delta_n \Big\{ \frac{2\pi}{n} \sum_{(i,j)\in\nu_k} (-1)^{j+1} s_i \Big\}$$

where $C$ is some constant that depends on $K$ and the constants $K_p$ and $\rho_p$ with $p \in \{1, \dots, 2K\}$ from assumption (C) only. Furthermore, since

$$\Delta_n(2\pi\omega/n) = \begin{cases} n & \text{if } \omega \in n\mathbb{Z}, \\ 0 & \text{if } \omega \notin n\mathbb{Z}, \end{cases}$$

we have that for each combination of $\nu = \{\nu_1, \dots, \nu_N\}$ and $I \subset \{1, \dots, N\}$

$$\prod_{k\in I} \Delta_n\Big\{\frac{2\pi}{n} \sum_{(i,j)\in\nu_k} (-1)^{j+1} s_i\Big\} = 0$$

unless

$$\forall_{\nu_k\in\nu: k\in I} \quad \sum_{(i,j)\in\nu_k} (-1)^{j+1} s_i \in n\mathbb{Z}.$$

In the latter case,

$$\prod_{k\in I} \Delta_n\Big\{\frac{2\pi}{n} \sum_{(i,j)\in\nu_k} (-1)^{j+1} s_i\Big\} = n^{|I|}.$$

Now we can restrict the sum over the indices $(s_1, \dots, s_K)$ to the set

$$S(\nu, I) = \Big\{(s_1, \dots, s_K) \in \{1, \dots, n-1\}^K : \forall_{\nu_k\in\nu: k\in I} \quad \sum_{(i,j)\in\nu_k} (-1)^{j+1} s_i \in n\mathbb{Z}\Big\}.$$

To complete the proof, follow the arguments starting at the bottom of p. 16 of the Online Supplement in the proof of Lemma A.2 in [23] (note that the supplement states this as "proof of Lemma 7.2"). First, note that $S(\nu, I)$ is empty for $|I| = K + 1$ and

$$\begin{aligned}\sum_{s_1,\dots,s_K=1}^{n-1} \prod_{m=1}^{K} \big|W_n(\omega_m - 2\pi s_m/n)\big| \prod_{k\in I} \Delta_n\Big\{\frac{2\pi}{n} \sum_{(i,j)\in\nu_k} (-1)^{j+1} s_i\Big\} &= \sum_{(s_1,\dots,s_K)\in S(\nu,I)} \prod_{m=1}^{K} \big|W_n(\omega_m - 2\pi s_m/n)\big| n^{|I|} \\ &= O\{(b_n^{-1})^{|I|-\lfloor |I|/N\rfloor} n^{K-(|I|-\lfloor |I|/N\rfloor)} n^{|I|}\},\end{aligned}$$

where the last equality follows from the arguments around Eq. (1.26) in the Online Supplement of [23]. Finally, the number $N$ of indecomposable partitions does not depend on $n$, so that $\operatorname{cum}\{H_n(\boldsymbol{\tau}_1, \omega_1), \dots, H_n(\boldsymbol{\tau}_K, \omega_K)\}$ is of order

$$n^{-3K/2}(b_n)^{K/2} \max_{N\le K} \max_{|I|\le N} (b_n^{-1})^{|I|-\lfloor |I|/N\rfloor} n^{K-(|I|-\lfloor |I|/N\rfloor)} n^{|I|} = O\{(nb_n)^{1-K/2}\},$$

which tends to zero for $K \ge 3$. □

**Proof of (iii).** Following the proof of Lemma A.3 and the arguments in the end of the proof of Lemma A.4 in [23] and using (C) it is straightforward to prove that $\omega \mapsto f^\theta_{\boldsymbol{\tau}}(\omega)$ is infinitely often continuously differentiable (for any $\boldsymbol{\tau} \in (0, 1)^2$ and $\theta \in U_\varepsilon(\theta_0)$) and that there exist constants $C, d$ that are independent of $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \theta$ with

$$\sup_{\omega\in\mathbb{R}, \theta\in U_\varepsilon(\theta_0)} \Big|\frac{d^j}{d\omega^j} f^\theta_{\boldsymbol{\tau}_1}(\omega) - \frac{d^j}{d\omega^j} f^\theta_{\boldsymbol{\tau}_2}(\omega)\Big| \le C\|\boldsymbol{\tau}_1 - \boldsymbol{\tau}_2\|_1\{1 + \ln\|\boldsymbol{\tau}_1 - \boldsymbol{\tau}_2\|_1\}^d. \tag{A.8}$$

Moreover, the proof of Lemma A.5 [23] can be modified to obtain (recall the definition of $\hat{F}_n^{U,\theta_n}$ at the beginning of the proof of (i)):

$$\sup_{\tau\in[0,1]} |(\hat{F}_n^{U,\theta_n})^{-1}(\tau) - \tau| = O_p(n^{-1/2}). \tag{A.9}$$

As in [23] (A.4), it follows that, for any $k \in \mathbb{N}$, we have

$$\sup_{\omega\in\mathbb{R}} \sup_{\tau\in[0,1]} |d_{\tau,n}(\omega) - d^U_{(\hat{F}_n^{U,\theta_n})^{-1}(\tau),n}| \le n \sup_{\tau\in[0,1]} |\hat{F}_n^{U,\theta_n}(\tau) - \hat{F}_n^{U,\theta_n}(\tau-)| \le O_p(n^{1/(2k)}),$$

where $\hat{F}_n^{U,\theta_n}(\tau_j-) = \lim_{\xi\uparrow 0} \hat{F}_n^{U,\theta_n}(\tau - \xi)$ and the last inequality follows from (A.4). The remaining part of the proof is analogous to the arguments given in Section A.3 of [23] and details are omitted for the sake of brevity. □

**Proof of (iv).** Begin by observing that for some constant $C_W$ that depends on $W$ only, we have, for any $\boldsymbol{\tau} \in [0,1]^2$,

$$|\mathrm{E}\hat{f}^{U,\theta_n}_{\boldsymbol{\tau},n}(\omega) - \mathrm{E}\hat{f}^{U,\theta_0}_{\boldsymbol{\tau},n}(\omega)| \leq \frac{1}{2\pi n}\sum_{s=1}^{n-1}|W_n(\omega - 2\pi s/n)| \times |\mathrm{E}\{I^{U,\theta_n}_{\boldsymbol{\tau},n}(2\pi s/n)\} - \mathrm{E}\{I^{U,\theta_0}_{\boldsymbol{\tau},n}(2\pi s/n)\}|$$

$$\leq C_W \max_{\omega\in\mathcal{F}_n}|\mathrm{E}\{I^{U,\theta_n}_{\boldsymbol{\tau},n}(\omega)\} - \mathrm{E}\{I^{U,\theta_0}_{\boldsymbol{\tau},n}(\omega)\}| = C_W \max_{\omega\in\mathcal{F}_n}\left|\frac{1}{2\pi n}\sum_{t_1,t_2=0}^{n-1} e^{-i(t_1-t_2)\omega}\{C^{\theta_n}_{t_2-t_1}(\boldsymbol{\tau}) - C^{\theta_0}_{t_2-t_1}(\boldsymbol{\tau})\}\right|$$

$$\leq \frac{C_W}{2\pi n}\sum_{|t_1|\leq n}\sum_{k\in\mathbb{Z}}|C^{\theta_n}_k(\boldsymbol{\tau}) - C^{\theta_0}_k(\boldsymbol{\tau})| \leq C_W\sum_{k\in\mathbb{Z}}|C^{\theta_n}_k(\boldsymbol{\tau}) - C^{\theta_0}_k(\boldsymbol{\tau})|.$$

Now under (C) we have $|C^{\theta_n}_k(\boldsymbol{\tau}) - C^{\theta_0}_k(\boldsymbol{\tau})| \leq 2K_2\rho_2^{|k|}$. Hence, for any fixed $N \in \mathbb{N}$ we have by (LC)

$$\sup_{\boldsymbol{\tau}\in\mathcal{T}}\sum_{k\in\mathbb{Z}}|C^{\theta_n}_k(\boldsymbol{\tau}) - C^{\theta_0}_k(\boldsymbol{\tau})| \leq \sum_{|k|\leq N}\sup_{\boldsymbol{\tau}\in\mathcal{T}}|C^{\theta_n}_k(\boldsymbol{\tau}) - C^{\theta_0}_k(\boldsymbol{\tau})| + 4K_2\sum_{k>N}\rho_2^{|k|} \leq (2N+1)\|\theta_n - \theta_0\| + \frac{4K_2\rho_2^{N+2}}{1-\rho_2}.$$

Now by assumption $\|\theta_n - \theta_0\| = O(n^{-1/2})$, so picking $N = C\ln(n)$ for a constant $C > 0$ such that $4K_2\rho_2^{C\ln(n)} = o(n^{-1})$, we obtain

$$\sup_{\boldsymbol{\tau}\in\mathcal{T}}|\mathrm{E}\hat{f}^{U,\theta_n}_{\boldsymbol{\tau},n}(\omega) - \mathrm{E}\hat{f}^{U,\theta_0}_{\boldsymbol{\tau},n}(\omega)| = O(\ln n)\|\theta_n - \theta_0\| = o(\sqrt{nb_n})$$

since by assumption $nb_n = o(n^{1-1/(2k+1)})$ for some $k \in \mathbb{N}$. □

## Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmva.2019.03.003.

## References


[1] T.W. Anderson, Goodness of fit tests for spectral distributions, Ann. Statist. (1993) 830–847.
[2] M.S. Bartlett, Problèmes de l'analyse spectral des séries temporelles stationnaires, Publ. Inst. Statist. Univ. Paris 3 (1954) 119–134.
[3] M.S. Bartlett, An Introduction to Stochastic Processes: with Special Reference to Methods and Applications, CUP Archive, 1978.
[4] J. Baruník, T. Kley, Quantile coherency: A general measure for dependence between cyclical economic variables, Econom. J. forthcoming (2019).
[5] S. Birr, QPBoot: Model Validation using quantile spectral analysis and parametric bootstrap, 2017, R package version 0.2.
[6] S. Birr, S. Volgushev, T. Kley, H. Dette, M. Hallin, Quantile spectral analysis for locally stationary time series, J. R. Stat. Soc. Ser. B 79 (5) (2017) 1619–1643.
[7] D.R. Brillinger, Time Series. Data Analysis and Theory, Holt, Rinehart and Winston, New York-Montréal-London, 1975.
[8] P.J. Brockwell, A. Lindner, Strictly stationary solutions of autoregressive moving average equations, Biometrika 97 (2010) 765–772.
[9] J. Chau, H. Ombao, R. von Sachs, Data depth and rank-based tests for covariance and spectral density matrices, Available at arXiv:1706.08289 2017.
[10] H. Dette, M. Hallin, T. Kley, S. Volgushev, Of copulas, quantiles, ranks and spectra: An $L_1$ approach to spectral analysis, Bernoulli 21 (2015) 781–831.
[11] H. Dette, E. Paparoditis, Bootstrapping frequency domain tests in multivariate time series with an application to comparing spectral densities, J. R. Stat. Soc. Ser. B Stat. Methodol. 71 (4) (2009) 831–857.
[12] P.J. Diggle, N.I. Fisher, Nonparametric comparison of cumulative periodograms, Appl. Stat. (1991) 423–434.
[13] J.C. Escanciano, Goodness-of-fit tests for linear and nonlinear time series models, J. Amer. Statist. Assoc. 101 (474) (2006) 531–541.
[14] J. Fan, W. Zhang, Generalised likelihood ratio tests for spectral density, Biometrika 91 (1) (2004) 195–209.
[15] A. Ghalanos, Rugarch: univariate garch models., 2018, R package version 1.4-0..
[16] A. Hagemann, Robust spectral analysis, Available at arXiv:1111.1965v2 2013.
[17] Y. Hong, Hypothesis testing in time series via the empirical characteristic function: A generalized spectral density approach., J. Amer. Statist. Assoc. 94 (1999) 1201–1220.
[18] Y. Hong, Generalized spectral tests for serial dependence., J. R. Stat. Soc. Ser. B 62 (2000) 557–574.
[19] Y. Hong, T.-H. Lee, Diagnostic checking for the adequacy of nonlinear time series models, Econ. Theory 19 (6) (2003) 1065–1121.
[20] Y. Hong, Y. Lee, Generalized spectral tests for conditional mean models in time series with conditional heteroscedasticity of unknown form, Rev. Econom. Stud. 72 (2) (2005) 499–541.
[21] C. Jentsch, M. Pauly, Testing equality of spectral densities using randomization techniques, Bernoulli 21 (2) (2015) 697–739.
[22] T. Kley, Quantile-based spectral analysis in an object-oriented framework and a reference implementation in R: the quantspec package., J. Stat. Softw. 70(3) (2016) 1–27.
[23] T. Kley, S. Volgushev, H. Dette, M. Hallin, Quantile spectral processes: asymptotic analysis and inference, Bernoulli 22 (2016) 1770–1807.
[24] T.-H. Li, Laplace periodogram for time series analysis, J. Amer. Statist. Assoc. 103 (2008) 757–768.
[25] T.-H. Li, Quantile periodograms, J. Amer. Statist. Assoc. 107 (2012) 765–776.
[26] T. McElroy, S. Holan, A local spectral approach for assessing time series model misspecification, J. Multivariate Anal. 100 (4) (2009) 604–621.
[27] E. Paparoditis, Spectral density based goodness-of-fit tests for time series models, Scand. J. Stat. 27 (1) (2000) 143–176.
[28] M.B. Priestley, Spectral Analysis and Time Series: Univariate series, Academic Press, London-New York, 1981.
[29] A.W. van der Vaart, Asymptotic statistics, 3, Cambridge University Press, 1998.
[30] A.W. van der Vaart, J. Wellner, Weak convergence and empirical processes: with applications to statistics, Springer Science & Business Media, 1996.