Nonlinear cross-spectrum analysis via the local Gaussian correlation

Lars Arne Jordanger Dag Tjøstheim

Abstract

Spectrum analysis can detect frequency related structures in a time series $\{\boldsymbol{Y}_t\}_{t\in\mathbb{Z}}$, but may in general be an inadequate tool if asymmetries or other nonlinear phenomena are present. This limitation is a consequence of the way the spectrum is based on the second order moments (auto and cross-covariances), and alternative approaches to spectrum analysis have thus been investigated based on other measures of dependence. One such approach was developed for univariate time series in Jordanger and Tjøstheim (2017), where it was seen that a *local Gaussian autospectrum* $f_{\boldsymbol{v}}(\omega)$, based on the *local Gaussian autocorrelations* $\rho_{\boldsymbol{v}}(h)$ from Tjøstheim and Hufthammer (2013), could detect local structures in time series that looked like white noise when investigated by the ordinary auto-spectrum $f(\omega)$. The *local Gaussian approach* in this paper is extended to a *local Gaussian cross-spectrum* $f_{k\ell:\boldsymbol{v}}(\omega)$ for multivariate time series. The local cross-spectrum $f_{k\ell:\boldsymbol{v}}(\omega)$ has the desirable property that it coincides with the ordinary cross-spectrum $f_{k\ell}(\omega)$ for Gaussian time series, which implies that $f_{k\ell:\boldsymbol{v}}(\omega)$ can be used to detect non-Gaussian traits in the time series under investigation. In particular: If the ordinary spectrum is flat, then peaks and troughs of *the local Gaussian spectrum* can indicate nonlinear traits, which potentially might discover *local periodic phenomena* that goes undetected in an ordinary spectral analysis.

1 Introduction

The auto and cross-covariances $\big\{\{\gamma_{k\ell}(h)\}_{h\in\mathbb{Z}}\big\}_{k,\ell=1}^{d}$ from a time series $\big\{\boldsymbol{Y}_t=\big(Y_{1,t},\ldots,Y_{d,t}\big)\big\}_{t\in\mathbb{Z}}$, can range from determining it completely (Gaussian time series) to containing no information at all (GARCH-type models). The auto- and cross-spectral densities $\{f_{k\ell}(\omega)\}_{k,\ell=1}^{d}$ (based on these second order moments) inherit these features, and they may thus be inadequate tools when the task of interest is to investigate non-Gaussian time series containing asymmetries or other nonlinear structures – like those observed in stock returns, cf. e.g. Hong et al. (2007).

In ordinary spectral analysis, if $\big\{Y_{k,t}\big\}_{t\in\mathbb{Z}}$ and $\big\{Y_{\ell,t}\big\}_{t\in\mathbb{Z}}$ are jointly weakly stationary, and if the cross-covariances $\gamma_{k\ell}(h) := \mathrm{Cov}\big(Y_{k,t+h},Y_{\ell,t}\big)$ are absolutely summable, then the cross-spectrum $f_{k\ell}(\omega)$ is defined as the Fourier transform of the autocovariances, .i.e.

$$f_{k\ell}(\omega)=\sum_{h\in\mathbb{Z}}\gamma_{k\ell}(h)\cdot e^{-2\pi i\omega h}. \tag{1.1}$$

The expression for the inverse Fourier transform reveals, when $h = 0$, that the covariance $\text{Cov}\big(Y_{k,t}, Y_{\ell,t}\big) = \gamma_{k\ell}(0)$ can be expressed as the integral $\int_{-1/2}^{1/2} f_{k\ell}(\omega)\,\mathrm{d}\omega$. This makes it possible to inspect how the interaction between the marginal time series varies with the frequency ω. An inspection of the cross-spectrum $f_{k\ell}(\omega)$ is a bit more complicated than that of the auto-spectrum, since $f_{k\ell}(\omega)$ in general will be a complex-valued function. It is thus usually the following real valued functions that are investigated,

$$c_{k\ell}(\omega) = \text{Re}\left(f_{k\ell}(\omega)\right), \qquad q_{k\ell}(\omega) = -\,\text{Im}\left(f_{k\ell}(\omega)\right), \tag{1.2a}$$

$$\alpha_{k\ell}(\omega) = \text{Mod}\left(f_{k\ell}(\omega)\right), \qquad \phi_{k\ell}(\omega) = \quad \text{Arg}\left(f_{k\ell}(\omega)\right), \tag{1.2b}$$

where $c_{k\ell}(\omega)$, $q_{k\ell}(\omega)$, $\alpha_{k\ell}(\omega)$ and $\phi_{k\ell}(\omega)$, respectively, are referred to as the cospectrum, quadrature spectrum, amplitude spectrum and phase spectrum. Note that $c_{k\ell}(\omega)$ always integrates to one over one period, whereas $q_{k\ell}(\omega)$ always integrates to zero.

The *coherence* $\mathcal{K}_{k\ell}(\omega) \coloneqq f_{k\ell}(\omega)/\sqrt{f_{kk}(\omega)f_{\ell\ell}(\omega)}$ is an important tool when a spectral analysis is performed on a multivariate time series, in particular since $\mathcal{K}_{k\ell}(\omega)$ can be realised as the correlation of $\mathrm{d}Z_k(\omega)$ and $\mathrm{d}Z_\ell(\omega)$, where $Z_k(\omega)$ and $Z_\ell(\omega)$ are the right continuous orthogonal-increment processes that by the Spectral Representation Theorem correspond to the weakly stationary time series $\big\{Y_{k,t}\big\}_{t\in\mathbb{Z}}$ and $\big\{Y_{\ell,t}\big\}_{t\in\mathbb{Z}}$, see e.g. Brockwell and Davis (1986, p. 436) for details. The *squared coherence* $|\mathcal{K}_{k\ell}(\omega)|^2$ is of interest since its value (in the interval $[0,1]$) reveals to what extent the two time series $\big\{Y_{k,t}\big\}_{t\in\mathbb{Z}}$ and $\big\{Y_{\ell,t}\big\}_{t\in\mathbb{Z}}$ can be related by a linear filter.

Other spectral approaches, involving different generalisations of the auto-spectrum $f(\omega)$ were discussed in Jordanger and Tjøstheim (2017, section 1), and the majority of the approaches were based on the following idea: The second order moments captured by the autocovariances $\{\gamma(h)\}_{h\in\mathbb{Z}}$ can be replaced by alternative dependence measures ξ_h computed from the bivariate random variables $\big(Y_{t+h}, Y_t\big)$, and a spectral density approach can then (under suitable regularity conditions) be defined as the Fourier transform of $\{\xi_h\}_{h\in\mathbb{Z}}$. For multivariate time series, the natural extension is then to define similar measures $\xi_{k\ell:h}$ for the bivariate random variables $\big(Y_{k,t+h}, Y_{\ell,t}\big)$, and then use the corresponding Fourier-transform as an alternative to the cross-spectrum $f_{k\ell}(\omega)$.

It does not seem to be the case (yet) that multivariate versions have been investigated for all of the possible generalisations of the auto-spectrum $f(\omega)$, but some generalisations do exist. The first extension of the cross-spectrum $f_{k\ell}(\omega)$ along these lines is the *polyspectra* introduced in Brillinger (1965), which is the multivariate version of the higher order moments/cumulants approach to spectral analysis, see Brillinger (1984, 1991); Tukey (1959). Another generalisation of $f_{k\ell}(\omega)$ is given in Chung and Hong (2007), where the generalised function approach introduced in Hong (1999) is used to set up a cross-spectrum that can be used for the testing of directional predictability in foreign exchange markets.

A local Gaussian spectral density $f_{\boldsymbol{v}}(\omega)$ for univariate strictly stationary time series was defined in Jordanger and Tjøstheim (2017), based on the local Gaussian auto-correlations $\rho_{\boldsymbol{v}}(h)$ from Tjøstheim and Hufthammer (2013). A simple adjustment gives the local Gaussian cross-correlations $\rho_{k\ell:\boldsymbol{v}}(h)$ for multivariate strictly stationary time series, from which a local Gaussian analogue $f_{k\ell:\boldsymbol{v}}(\omega)$ of the cross-spectrum $f_{k\ell}(\omega)$ can be constructed using the Fourier transform. The local Gaussian version of the cross-spectrum enables local Gaussian alternatives to be defined of the cospectrum, quadrature spectrum, amplitude spectrum, and phase spectrum, by

simply copying the setup used in the ordinary (global) case. Local Gaussian analogues of the coherence and squared coherence were investigated in the preparation for this paper, but then discarded, see remark 2.5 for further details.

An overview of the paper is as follows: Section 2 defines the *local Gaussian cross-spectrum* $f_{k\ell:\boldsymbol{v}}(\omega)$, which immediately gives the related local Gaussian variants of the cospectrum, quadrature spectrum, amplitude spectrum and phase spectrum from eq. (1.2). The asymptotic theory for the estimators are then presented (some technical details and proofs are postponed to the appendices). The real and simulated examples in section 3 shows that estimates of $f_{k\ell:\boldsymbol{v}}(\omega)$ can be used to detect and investigate nonlinear structures in non-Gaussian white noise, and in particular that $f_{k\ell:\boldsymbol{v}}(\omega)$ can detect local periodic phenomena that goes undetected in an ordinary spectral analysis. Note that the scripts needed for the reproduction of these examples are contained in the R-package `localgaussSpec`,[1] where it in addition is possible to use an interactive solution to see how adjustments of the input parameters (used in the estimation algorithms) influence the estimates of $f_{k\ell:\boldsymbol{v}}(\omega)$. A discussion is given in section 4, and section 5 presents the conclusion.

2 Definitions

This section will present the formal definitions of the local Gaussian versions of the cross-correlation $f_{k\ell}(\omega)$ and its derived entities. The details are almost identical to those encountered when the local Gaussian spectral density was introdued in Jordanger and Tjøstheim (2017), and the present discussion will thus only give short summaries of descriptions and arguments already undertaken in (Jordanger and Tjøstheim, 2017).

2.1 The local Gaussian correlations

At the core of the generalisation of eq. (1.1) lies the local Gaussian correlation $\rho_{\boldsymbol{v}}$ from Tjøstheim and Hufthammer (2013). The theoretical treatment can be based directly on $\rho_{\boldsymbol{v}}$, but the numerical convergence of the estimation-algorithm might then sometimes fail, in particular if the samples contains outliers. As noted in (Jordanger and Tjøstheim, 2017, sections 2.1.2 and 2.1.3), this estimation-problem can be countered by the help of the two revised versions $\rho_{\boldsymbol{v}|5}$ and $\rho_{\boldsymbol{v}|1}$, where a *normalisation of the marginals* (see definition 2.1 below) are performed before the estimation-algorithm is used. The numbers 5 and 1 refer to whether the estimation-algorithm use a local Gaussian correlation that originates from a *five-parameter* or a *one-parameter* local Gaussian approximation, as shown below,

$$\psi_5(\boldsymbol{w};\mu_1,\mu_2,\sigma_1,\sigma_2,\rho) \coloneqq \frac{1}{2\pi\cdot\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{\sigma_1^2\big(w_1-\mu_1\big)^2-2\sigma_1\sigma_2\rho\big(w_1-\mu_1\big)\big(w_2-\mu_2\big)+\sigma_2^2\big(w_2-\mu_2\big)^2}{\sigma_1^2\sigma_2^2(1-\rho^2)}\right\}, \tag{2.1a}$$

$$\psi_1(\boldsymbol{w};\rho) \coloneqq \frac{1}{2\pi\cdot\sqrt{1-\rho^2}} \exp\left\{-\frac{w_1^2-2\rho w_1 w_2+w_2^2}{1-\rho^2}\right\}. \tag{2.1b}$$

Both $\rho_{\boldsymbol{v}|5}$ and $\rho_{\boldsymbol{v}|1}$ can be used as the starting point for the theoretical investigation, and the notation $\rho_{\boldsymbol{v}|p}$ will be used to indicate that both of the alternatives are discussed simultaneously. Although both of the alternatives can be used, only the constructions based on $\rho_{\boldsymbol{v}|5}$ will in

[1] Use `devtools::install_github("LAJordanger/localgaussSpec")` to install the package. See the documentation of the function `LG_extract_scripts` for further details.

general be able to properly detect local properties of the investigated time series. The discussion in (Jordanger and Tjøstheim, 2017, appendix C.6)) gives some reasons for the failure of the one-parameter local Gaussian approach in this context.

Remark 2.1. In order to have a unified notation, the two functions in eq. (2.1) will later on be denoted by $\psi_p\big(\boldsymbol{w};\boldsymbol{\theta}_p\big)$, where additional indices will be added to $\boldsymbol{\theta}_p$ in order to identify the targeted density and the point of investigation.

2.2 The local Gaussian cross-spectrum

The definition of the local Gaussian cross-spectrum density is almost identical to the definition of the local Gaussian spectral density from (Jordanger and Tjøstheim, 2017, section 2.2), which in this paper henceforth will be referred to as the local Gaussian auto-spectrum.

Definition 2.1. *For a strictly stationary multivariate time series $\{\boldsymbol{Y}_t\}_{t\in\mathbb{Z}}$, where $\boldsymbol{Y}_t = \big(Y_{1,t},\ldots,Y_{d,t}\big)$, the local Gaussian cross-spectrum of the marginal time series $\big\{Y_{k,t}\big\}_{t\in\mathbb{Z}}$ and $\big\{Y_{\ell,t}\big\}_{t\in\mathbb{Z}}$ is constructed in the following manner.*

(a) With G_k and G_ℓ the univariate marginal cumulative distribution of respectively $\big\{Y_{k,t}\big\}_{t\in\mathbb{Z}}$ and $\big\{Y_{\ell,t}\big\}_{t\in\mathbb{Z}}$, and Φ the cumulative distribution of the univariate standard normal distribution, define normalised versions $\big\{Z_{k,t}\big\}_{t\in\mathbb{Z}}$ and $\big\{Z_{\ell,t}\big\}_{t\in\mathbb{Z}}$ by

$$\big\{Z_{k,t} \coloneqq \Phi^{-1}\big(G_k\big(Y_{k,t}\big)\big)\big\}_{t\in\mathbb{Z}}, \qquad \big\{Z_{\ell,t} \coloneqq \Phi^{-1}\big(G_\ell\big(Y_{\ell,t}\big)\big)\big\}_{t\in\mathbb{Z}}. \tag{2.2}$$

(b) For a given point $\boldsymbol{v} = (v_1, v_2)$ and for each bivariate lag h pair $\boldsymbol{Z}_{k\ell:h:t} \coloneqq \big(Z_{k:t+h}, Z_{\ell:t}\big)$, a local Gaussian cross-correlation $\rho_{k\ell:\boldsymbol{v}|p}(h)$ can be computed, where the p specifies if the correlations stems from a one or a five parameter local Gaussian approximation of the bivariate density of $\boldsymbol{Z}_{k\ell:h:t}$ at (v_1, v_2).

(c) When $\sum_{h\in\mathbb{Z}}\big|\rho_{k\ell:\boldsymbol{v}|p}(h)\big| < \infty$, the local Gaussian cross-spectrum at the point $\boldsymbol{v}$ is defined as

$$f_{k\ell:\boldsymbol{v}|p}(\omega) \coloneqq \sum_{h=-\infty}^{\infty} \rho_{k\ell:\boldsymbol{v}|p}(h)\cdot e^{-2\pi i\omega h}, \qquad \omega\in\left[-\tfrac{1}{2},\tfrac{1}{2}\right]. \tag{2.3}$$

Remark 2.2. The definition of the local Gaussian auto-spectrum is in essence the same as the one given here for the local Gaussian cross-spectrum, with the minor adjustment that $k = \ell$ in the auto-spectrum case – which requires the added convention that $\rho_{kk:\boldsymbol{v}|p}(0) \equiv 1$ for all points $\boldsymbol{v}$.

The basic properties of the local Gaussian cross-spectrum are quite similar to those encountered for the local Gaussian auto-spectrum in (Jordanger and Tjøstheim, 2017, lemma 2.3).

Lemma 2.2. *The following properties holds for $f_{k\ell:\boldsymbol{v}|p}(\omega)$.*

(a) $f_{k\ell:\boldsymbol{v}|p}(\omega)$ coincides with $f_{k\ell}(\omega)$ for all $\boldsymbol{v}\in\mathbb{R}^2$ when $\{\boldsymbol{Y}_t\}_{t\in\mathbb{Z}}$ is a multivariate Gaussian time series.

(b) The following holds when $\breve{\boldsymbol{v}} \coloneqq (v_2, v_1)$ is the diagonal reflection of $\boldsymbol{v} = (v_1, v_2)$;

$$f_{k\ell:\boldsymbol{v}|p}(\omega) = \overline{f_{\ell k:\breve{\boldsymbol{v}}|p}(\omega)}, \tag{2.4a}$$

$$f_{k\ell:\boldsymbol{v}|p}(\omega) = \rho_{k\ell:\boldsymbol{v}|p}(0) + \sum_{h=1}^{\infty}\rho_{\ell k:\breve{\boldsymbol{v}}|p}(h)\cdot e^{+2\pi i\omega h} + \sum_{h=1}^{\infty}\rho_{k\ell:\boldsymbol{v}|p}(h)\cdot e^{-2\pi i\omega h}. \tag{2.4b}$$

Proof. Item (a) follows since the local Gaussian cross-correlations $\rho_{k\ell:\boldsymbol{v}|p}(h)$ by construction coincides with the ordinary (global) cross-correlations $\rho(h)$ in the Gaussian case. For the proof of item (b), the key observation is that the *diagonal folding property* that was observed for the local Gaussian auto-spectrum, see (Jordanger and Tjøstheim, 2017, lemma C.1), extends directly to the present case, i.e. $\rho_{k\ell|\boldsymbol{v}}(-h) = \rho_{\ell k|\breve{\boldsymbol{v}}}(h)$, where $\breve{\boldsymbol{v}} = (v_2, v_1)$ is the *diagonally reflected point corresponding to* $\boldsymbol{v}$. This implies that $f_{k\ell:\boldsymbol{v}|p}(\omega) = \overline{f_{k\ell:\boldsymbol{v}|p}(-\omega)} = \overline{f_{\ell k:\breve{\boldsymbol{v}}|p}(\omega)}$, and it also follows that eq. (2.3) can be reexpressed as eq. (2.4b). □

2.3 Related local Gaussian entities

From the definition of the local Gaussian cross-spectrum, it is possible to define related spectra in the same manner as those mentioned for the ordinary spectrum in eq. (1.2).

Definition 2.3. *The local Gaussian versions of the cospectrum $c_{k\ell}(\omega)$, the quadrature spectrum $q_{k\ell}(\omega)$, the amplitude spectrum $\alpha_{k\ell}(\omega)$ and the phase spectrum $\phi_{k\ell}(\omega)$, are given by*

$$c_{k\ell:\boldsymbol{v}|p}(\omega) := \operatorname{Re}\left(f_{k\ell:\boldsymbol{v}|p}(\omega)\right) = \rho_{k\ell:\boldsymbol{v}|p}(0) + \sum_{h=1}^{\infty} \cos\left(2\pi\omega h\right)\left[\rho_{k\ell:\boldsymbol{v}|p}(h) + \rho_{k\ell:\breve{\boldsymbol{v}}|p}(h)\right], \tag{2.5a}$$

$$q_{k\ell:\boldsymbol{v}|p}(\omega) := -\operatorname{Im}\left(f_{k\ell:\boldsymbol{v}|p}(\omega)\right) = \sum_{h=1}^{\infty} \sin\left(2\pi\omega h\right)\left[\rho_{k\ell:\boldsymbol{v}|p}(h) - \rho_{k\ell:\breve{\boldsymbol{v}}|p}(h)\right], \tag{2.5b}$$

$$\alpha_{k\ell:\boldsymbol{v}|p}(\omega) := \operatorname{Mod}\left(f_{k\ell:\boldsymbol{v}|p}(\omega)\right) = \sqrt{c^2_{k\ell:\boldsymbol{v}|p}(\omega) + q^2_{k\ell:\boldsymbol{v}|p}(\omega)}, \tag{2.5c}$$

$$\phi_{k\ell:\boldsymbol{v}|p}(\omega) := \operatorname{Arg}\left(f_{k\ell:\boldsymbol{v}|p}(\omega)\right) \in (-\pi, \pi]. \tag{2.5d}$$

Remark 2.3. The sums occurring in eqs. (2.5a) and (2.5b) follows from eq. (2.4b). Equation (2.4a) gives $c_{k\ell:\boldsymbol{v}|p}(\omega) = c_{\ell k|\breve{\boldsymbol{v}}}(\omega)$, $q_{k\ell:\boldsymbol{v}|p}(\omega) = -q_{\ell k|\breve{\boldsymbol{v}}}(\omega)$, $\alpha_{k\ell:\boldsymbol{v}|p}(\omega) = \alpha_{\ell k|\breve{\boldsymbol{v}}}(\omega)$ and $\phi_{k\ell:\boldsymbol{v}|p}(\omega) = -\phi_{\ell k|\breve{\boldsymbol{v}}}(\omega)$.

Remark 2.4. For Gaussian distributions, the local Gaussian correlations will always be equal to the ordinary (global) correlations,[2] and the local Gaussian constructions in definitions 2.1 and 2.3 will thus coincide with the ordinary (global) versions for multivariate Gaussian time series. A comparison of the local and global estimates in the same plot is thus of interest when a given sample is considered, since this could detect nonlinear interactions of the time series under investigation.

Remark 2.5. It is possible to define a local Gaussian anaologue of the squared coherence mentioned in section 1 by replacing the ordinary cross- and auto-spectra with the corresponding local Gaussian versions, i.e. the object of interest would be $\mathcal{Q}_{k\ell:\boldsymbol{v}|p}(\omega) := f_{k\ell:\boldsymbol{v}|p}(\omega) f_{\ell k:\boldsymbol{v}|p}(\omega) / f_{kk:\boldsymbol{v}|p}(\omega) f_{\ell\ell:\boldsymbol{v}|p}(\omega)$. This approach was investigated in the preparation of this paper, but it has not been included here since $\mathcal{Q}_{k\ell:\boldsymbol{v}|p}(\omega)$ in general lacked the nice properties known from the ordinary global case. In particular, the local Gaussian auto-spectra $f_{kk:\boldsymbol{v}|p}(\omega)$ and $f_{\ell\ell:\boldsymbol{v}|p}(\omega)$ will in general be complex valued functions, so an inspection of $\mathcal{Q}_{k\ell:\boldsymbol{v}|p}(\omega)$ must thus be based on plots of its real and imaginary parts (or its amplitude and phase). Moreover, these plots did more often than not turn out to be rather hard to investigate, since the estimates of $f_{kk:\boldsymbol{v}|p}(\omega)$ and $f_{\ell\ell:\boldsymbol{v}|p}(\omega)$ (for some distributions and some frequencies ω) gave values very close to zero in the denominator.

[2] This is due to the way the local Gaussian correlation is defined, see Tjøstheim and Hufthammer (2013) for details.

2.4 Estimation

The estimation of the local Gaussian cross-spectrum $f_{k\ell:\boldsymbol{v}|p}(h)$ from section 2.2 follows the same setup that was used in (Jordanger and Tjøstheim, 2017, section 2.3) for the estimation of the local Gaussian auto-spectrum, with the obvious difference that some extra indices are needed in the present case. The estimation of the related spectra $c_{k\ell:\boldsymbol{v}|p}(h)$, $q_{k\ell:\boldsymbol{v}|p}(h)$, $\alpha_{k\ell:\boldsymbol{v}|p}(h)$ and $\phi_{k\ell:\boldsymbol{v}|p}(h)$ from section 2.3 is then obtained from the estimate of $f_{k\ell:\boldsymbol{v}|p}(h)$ in an obvious manner.

Definition 2.4. *For a sample $\{\boldsymbol{y}_t = (y_{1,t}, \ldots, y_{d,t})\}_{t=1}^n$ of size n from a multivariate time series, an m-truncated estimate $\widehat{f}^m_{k\ell:\boldsymbol{v}|p}(\omega)$ of $f_{k\ell:\boldsymbol{v}|p}(\omega)$ is constructed by means of the following procedure.*

(a) Use the univariate marginals $\{y_{k,t}\}_{t=1}^n$ and $\{y_{\ell,t}\}_{t=1}^n$ to find estimates $\widehat{G}_{k:n}$ and $\widehat{G}_{\ell:n}$ of the corresponding marginal cumulative distribution functions, and compute from this the pseudo-normalised observations $\left\{\widehat{z}_{k,t} \coloneqq \Phi^{-1}\left(\widehat{G}_{k:n}(y_{k,t})\right)\right\}_{t=1}^n$ and $\left\{\widehat{z}_{\ell,t} \coloneqq \Phi^{-1}\left(\widehat{G}_{\ell:n}(y_{\ell,t})\right)\right\}_{t=1}^n$.

(b) Create the lag h pseudo-normalised pairs $\left\{(\widehat{z}_{k,t+h}, \widehat{z}_{\ell,t})\right\}_{t=1}^{n-h}$ for $h = 0, \ldots, m$, and estimate for the point $\boldsymbol{v} = (v_1, v_2)$ the local Gaussian cross-correlations $\left\{\widehat{\rho}_{k\ell:\boldsymbol{v}|p}(h|\boldsymbol{b}_h)\right\}_{h=0}^m$, where the $\{\boldsymbol{b}_h\}_{h=0}^m$ is the bandwidths that are used for the different lags.

(c) Create the lag h pseudo-normalised pairs $\left\{(\widehat{z}_{\ell,t+h}, \widehat{z}_{k,t})\right\}_{t=1}^{n-h}$ for $h = 1, \ldots, m$, and estimate for the diagonally reflected point $\breve{\boldsymbol{v}} = (v_2, v_1)$ the local Gaussian cross-correlations $\left\{\widehat{\rho}_{\ell k:\breve{\boldsymbol{v}}|p}(h|\boldsymbol{b}_h)\right\}_{h=0}^m$.

(d) Adjust eq. (2.4b) from lemma 2.2(b) with some lag-window function $\lambda_m(h)$ to get the estimate

$$\widehat{f}^m_{k\ell:\boldsymbol{v}|p}(\omega) \coloneqq \widehat{\rho}_{k\ell:\boldsymbol{v}|p}(0) + \sum_{h=1}^m \lambda_m(h) \cdot \widehat{\rho}_{\ell k:\breve{\boldsymbol{v}}|p}(h) \cdot e^{+2\pi i\omega h} + \sum_{h=1}^m \lambda_m(h) \cdot \widehat{\rho}_{k\ell:\boldsymbol{v}|p}(h) \cdot e^{-2\pi i\omega h}, \tag{2.6}$$

where the $\{\boldsymbol{b}_h\}_{h=0}^m$ has been suppressed from the notation in order to get a more compact formula.

Definition 2.5. *For a multivariate sample $\{\boldsymbol{y}_t\}_{t=1}^n$ of size n, as described in definition 2.4, the m-truncated estimates of the local Gaussian versions of the cospectrum, quadrature spectrum, amplitude spectrum and phase spectrum is given by*

$$c^m_{k\ell:\boldsymbol{v}|p}(\omega) \coloneqq \operatorname{Re}\left(f^m_{k\ell:\boldsymbol{v}|p}(\omega)\right) = \widehat{\rho}_{k\ell:\boldsymbol{v}|p}(0) + \sum_{h=1}^\infty \cos(2\pi\omega h)\left[\widehat{\rho}_{k\ell:\boldsymbol{v}|p}(h) + \widehat{\rho}_{k\ell:\breve{\boldsymbol{v}}|p}(h)\right], \tag{2.7a}$$

$$q^m_{k\ell:\boldsymbol{v}|p}(\omega) \coloneqq -\operatorname{Im}\left(f^m_{k\ell:\boldsymbol{v}|p}(\omega)\right) = \sum_{h=1}^\infty \sin(2\pi\omega h)\left[\widehat{\rho}_{k\ell:\boldsymbol{v}|p}(h) - \widehat{\rho}_{k\ell:\breve{\boldsymbol{v}}|p}(h)\right], \tag{2.7b}$$

$$\alpha^m_{k\ell:\boldsymbol{v}|p}(\omega) \coloneqq \operatorname{Mod}\left(f^m_{k\ell:\boldsymbol{v}|p}(\omega)\right) = \sqrt{\left(c^m_{k\ell:\boldsymbol{v}|p}(\omega)\right)^2 + \left(q^m_{k\ell:\boldsymbol{v}|p}(\omega)\right)^2}, \tag{2.7c}$$

$$\phi^m_{k\ell:\boldsymbol{v}|p}(\omega) \coloneqq \operatorname{Arg}\left(f^m_{k\ell:\boldsymbol{v}|p}(\omega)\right) \in (-\pi, \pi]. \tag{2.7d}$$

Remark 2.6. The comments in (Jordanger and Tjøstheim, 2017, remarks 2.5 to 2.8) holds for the present case too. In particular, the estimated marginal cumulative distributions $\widehat{G}_{k:n}$ and $\widehat{G}_{\ell:n}$ from definition 2.4(a) can either be based on the (rescaled) empirical cumulative distribution

functions or they could be built upon a logspline technique like the one implemented in Otneim and Tjøstheim (2016). Furthermore, for the asymptotic investigation, the arguments in (Otneim and Tjøstheim, 2016, Section 3) reveals that the pseudo-normalisation of the marginals does not affect the final convergence rates, which (as was done in (Jordanger and Tjøstheim, 2017)) implies that the present theoretical analysis can ignore the distinction between the original observations and the pseudo-normalised observations.

2.5 Asymptotic theory for $\widehat{f}^{m}_{k\ell:\boldsymbol{v}|p}(\omega)$

The asymptotic theory for the local Gaussian cross-spectrum $\widehat{f}^{m}_{k\ell:\boldsymbol{v}|p}(\omega)$ follows from a few minor adjustments of the asymptotic theory that was developed for the local Gaussian auto-spectra. As in (Jordanger and Tjøstheim, 2017, section 2.4), the assumptions and results will be stated for the original observations instead of the pseudo-normalised observations, since this makes the analysis easier and since the final convergence rates are unaffected by this distinction, see remark 2.6 for details.

2.5.1 Some definitions and an assumption for $\boldsymbol{Y}_t$

As for the univariate case in (Jordanger and Tjøstheim, 2017), the assumptions to be imposed on the k and ℓ components of the multivariate times series $\{\boldsymbol{Y}_t\}_{t\in\mathbb{Z}}$ need to be phrased relative to the bivariate pairs that can be created as different combinations of elements from the univariate marginals $\{Y_{k,t}\}_{t\in\mathbb{Z}}$ and $\{Y_{\ell,t}\}_{t\in\mathbb{Z}}$. Note that the *folding property* from item (d) of definition 2.4 implies that it is sufficient to formulate the assumption based on non-negative values of the lag h.

Definition 2.6. *For a strictly stationary multivariate time series $\{\boldsymbol{Y}_t\}_{t\in\mathbb{Z}}$, with $\boldsymbol{Y}_t = \big(Y_{1,t},\ldots,Y_{d,t}\big)$, and for a selected pair of indices k and ℓ, define the following bivariate pairs from the univariate marginals $\{Y_{k,t}\}_{t\in\mathbb{Z}}$ and $\{Y_{\ell,t}\}_{t\in\mathbb{Z}}$.*

$$\boldsymbol{Y}_{k\ell:h:t} \coloneqq \big[Y_{k,t+h}, Y_{\ell,t}\big]', \qquad h \geq 0, \tag{2.8a}$$

$$\boldsymbol{Y}_{\ell k:h:t} \coloneqq \big[Y_{\ell,t+h}, Y_{k,t}\big]', \qquad h \geq 1, \tag{2.8b}$$

and let $g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})$ and $g_{\ell k:h}(\boldsymbol{y}_{\ell k:h})$ denote the respective probability density functions.

The basic idea for the construction of $f_{k\ell:\boldsymbol{v}|p}(\omega)$ is that a point $\boldsymbol{v} = (v_1, v_2)$ should be selected at which for all h the density functions $g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})$ of $\boldsymbol{Y}_{k\ell:h:t}$ will be approximated by $\psi_p\big(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\big)$, where ψ_p is one of the bivariate Gaussian density functions from eq. (2.1). The correlation-parameter from the approximating Gaussian density function will be denoted $\rho_{k\ell:\boldsymbol{v}|p}(h)$, and it will be referred to as the local Gaussian lag h cross-correlation of $Y_{k,t+h}$ and $Y_{\ell,t}$ (in that order) at the point $\boldsymbol{v}$.

This local investigation requires a bandwidth vector $\boldsymbol{b} = (b_1, b_2)$ and a kernel function $K(\boldsymbol{w})$, which is used to define $K_{k\ell:h:\boldsymbol{b}}(\boldsymbol{y}_{k\ell:h} - \boldsymbol{v}) \coloneqq \frac{1}{b_1 b_2} K\Big(\frac{y_{k,h}-v_1}{b_1}, \frac{y_{\ell,0}-v_2}{b_2}\Big)$, which in turn is used in

$$q_{\boldsymbol{v}|k\ell:h:\boldsymbol{b}|p} \coloneqq \int_{\mathbb{R}^2} K_{k\ell:h:\boldsymbol{b}}(\boldsymbol{y}_{k\ell:h} - \boldsymbol{v}) \left[\psi_p\big(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\big) - \log \psi_p\big(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\big)\, g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})\right] \mathrm{d}\boldsymbol{y}_{k\ell:h}, \tag{2.9}$$

a minimiser of which should satisfy the vector equation

$$\int_{\mathbb{R}^2} K_{k\ell:h:\boldsymbol{b}}(\boldsymbol{y}_{k\ell:h} - \boldsymbol{v})\, \boldsymbol{u}_{k\ell:h|p}\big(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\big) \left[g_{k\ell:h}(\boldsymbol{y}_{k\ell:h}) - \psi_p\big(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\big)\right] \mathrm{d}\boldsymbol{y}_{k\ell:h} = \mathbf{0}, \tag{2.10}$$

where $\boldsymbol{u}_{k\ell:h|p}\big(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\big) \coloneqq \boldsymbol{\nabla}_{k\ell:h|p} \log \psi_p\big(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\big)$ is the score function of $\psi_p\big(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\big)$ (with $\boldsymbol{\nabla}_{k\ell:h|p} \coloneqq \partial/\partial\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}$). Under the assumption that there is a bandwidth $\boldsymbol{b}_{k\ell:h:0}$ such that there exists a minimiser $\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h:\boldsymbol{b}|p}$ of eq. (2.9) which satisfies eq. (2.10) for any $\boldsymbol{b}$ with $\mathbf{0} < \boldsymbol{b} < \boldsymbol{b}_{k\ell:h:0}$, this $\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h:\boldsymbol{b}|p}$ will be referred to as the population value for the given bandwidth $\boldsymbol{b}$.

Remark 2.7. This approach was introduced in a more general context in Hjort and Jones (1996), where it was used to define a local approach to density estimation, and the new idea in (Tjøstheim and Hufthammer, 2013) was to focus upon the estimated local Gaussian parameters $\widehat{\boldsymbol{\theta}}_{\boldsymbol{v}|k\ell:h|p}$ (instead of the estimated densities). The asymptotic properties of the estimated parameters was investigated in (Tjøstheim and Hufthammer, 2013) by the help of the Klimko-Nelson approach[3] and a suitably defined local penalty function $Q_{\boldsymbol{v}|k\ell:h:n|p}\big(\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\big)$ (see eq. (B.1) in appendix B.1).

The assumptions to be imposed on Y_t is related to the estimation of eq. (2.10), and thus requires a few additional definitions.

Definition 2.7. *For $\psi_p\big(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_p\big)$ the local Gaussian density used when approximating $g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})$ at the point $\boldsymbol{v} = (v_1, v_2)$, and for $\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h:\boldsymbol{b}|p}$ the population value that minimises the penalty function $q_{\boldsymbol{v}|k\ell:h:\boldsymbol{b}|p}$ from eq. (2.9), define for all $h \in \mathbb{N}$ and all $q \in \{1, \ldots, p\}$*

$$u_{k\ell:h:q:\boldsymbol{b}}(\boldsymbol{w}) \coloneqq \frac{\partial}{\partial\theta_{p:q}} \log\left(\psi_p\big(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_p\big)\right)\Big|_{\left(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_p\right) = \left(\boldsymbol{w}; \boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h:\boldsymbol{b}|p}\right)}, \tag{2.11}$$

where $\partial/\partial\theta_{p:q}$ is the q^{th} partial derivative (with respect to $\boldsymbol{\theta}_p$).

The following requirements on the kernel function are identical to those given in (Jordanger and Tjøstheim, 2017, definition B.9).

[3] The Klimko-Nelson approach (see Klimko and Nelson (1978)) shows how the asymptotic properties of *an estimate of the parameters of a penalty function Q* can be expressed relative to the asymptotic properties of (entities related to) the penalty function itself. The interested reader can consult (Jordanger and Tjøstheim, 2017, appendix B.1) for a more detailed presentation of the Klimko-Nelson approach when a local penalty-function is used.

Definition 2.8. *From a bivariate, non-negative, and bounded kernel function $K(\boldsymbol{w})$, that satisfies*

$$\int_{\mathbb{R}^2} K(w_1, w_2)\, \mathrm{d}w_1 \mathrm{d}w_2 = 1, \tag{2.12a}$$

$$\mathcal{K}_{1:k}(w_2) \coloneqq \int_{\mathbb{R}^1} K(w_1, w_2)\, w_1^k\, \mathrm{d}w_1 \qquad \textit{is bounded for } k \in \{0,1,2\}, \tag{2.12b}$$

$$\mathcal{K}_{2:\ell}(w_1) \coloneqq \int_{\mathbb{R}^1} K(w_1, w_2)\, w_2^\ell\, \mathrm{d}w_2 \qquad \textit{is bounded for } \ell \in \{0,1,2\}, \tag{2.12c}$$

$$\int_{\mathbb{R}^2} K(w_1, w_2)\, |w_1^k w_2^\ell|\, \mathrm{d}w_1 \mathrm{d}w_2 < \infty, \qquad k, \ell \geq 0 \textit{ and } k + \ell \leq 2 \cdot \lceil \nu \rceil\,, \tag{2.12d}$$

where $\nu > 2$ is from assumption 2.1(b) (and $\lceil \cdot \rceil$ is the ceiling function), define

$$K_{h:\boldsymbol{b}}(\boldsymbol{y}_{k\ell:h} - \boldsymbol{v}) \coloneqq \frac{1}{b_1 b_2} K\left(\frac{y_h - v_1}{b_1}, \frac{y_0 - v_2}{b_2}\right). \tag{2.13}$$

Definition 2.9. *Based on $\boldsymbol{Y}_{k\ell:h:t}$, $u_{k\ell:h:q:\boldsymbol{b}}(\boldsymbol{w})$ and $K_{h:\boldsymbol{b}}(\boldsymbol{y}_{k\ell:h} - \boldsymbol{v})$, define the new bivariate random variables $X^{n|\boldsymbol{v}}_{k\ell:h:q:t}$ as follows,*

$$X^{n|\boldsymbol{v}}_{k\ell:h:q:t} \coloneqq \sqrt{b_1 b_2} K_{h:\boldsymbol{b}}(\boldsymbol{Y}_{k\ell:h:t} - \boldsymbol{v})\, u_{k\ell:h:q:\boldsymbol{b}}(\boldsymbol{Y}_{k\ell:h:t})\,. \tag{2.14}$$

Note that a product of the random variables $X^{n|\boldsymbol{v}}_{k\ell:h:q:t}$ and $X^{n|\boldsymbol{v}}_{k\ell:i:r:s}$ will be a function of $Y_{k,t+h}$, $Y_{\ell,t}$, $Y_{k,s+i}$ and $Y_{\ell,s}$, which depending on the configuration of the indices h, i, s, t will be either a bivariate, trivariate or tetravariate function. The expectation of the product $X^{n|\boldsymbol{v}}_{k\ell:h:q:t} \cdot X^{n|\boldsymbol{v}}_{k\ell:i:r:s}$ will thus (depending on these indices) either require a bivariate, trivariate or tetravariate density function.

Assumption 2.1. The multivariate process $\{\boldsymbol{Y}_t\}_{t\in\mathbb{Z}}$ will be assumed to satisfy the following properties, with $\boldsymbol{v} = (v_1, v_2)$ (in item (d)) the point at which $\widehat{f}^m_{k\ell:\boldsymbol{v}|p}(\omega)$, the estimate of $f_{k\ell:\boldsymbol{v}|p}(\omega)$, is to be computed.

(a) $\{\boldsymbol{Y}_t\}_{t\in\mathbb{Z}}$ is strictly stationary.
(b) $\{\boldsymbol{Y}_t\}_{t\in\mathbb{Z}}$ is strongly mixing, with mixing coefficient $\alpha(j)$ satisfying

$$\sum_{j=1}^{\infty} j^a \left[\alpha(j)\right]^{1-2/\nu} < \infty \qquad \text{for some } \nu > 2 \text{ and } a > 1 - 2/\nu. \tag{2.15}$$

(c) $\mathrm{Var}(||\boldsymbol{Y}_t||^2) < \infty$, where $||\cdot||$ is the Euclidean norm.

The bivariate density functions $g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})$ and $g_{\ell k:h}(\boldsymbol{y}_{\ell k:h})$, corresponding to the lag h pairs introduced in eq. (2.8), must satisfy the following requirements for a given point $\boldsymbol{v} = (v_1, v_2)$.

(d) $g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})$ is differentiable at $\boldsymbol{v}$, such that Taylor's theorem can be used to write $g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})$ as

$$g_{k\ell:h}(\boldsymbol{y}_{k\ell:h}) = g_h(\boldsymbol{v}) + \boldsymbol{\mathfrak{g}}_h(\boldsymbol{v})' \left[\boldsymbol{y}_{k\ell:h} - \boldsymbol{v}\right] + \boldsymbol{\mathfrak{R}}_h(\boldsymbol{y}_{k\ell:h})' \left[\boldsymbol{y}_{k\ell:h} - \boldsymbol{v}\right], \tag{2.16}$$
$$\text{where } \boldsymbol{\mathfrak{g}}_h(\boldsymbol{v}) = \left[\frac{\partial}{\partial y_h} g_{k\ell:h}(\boldsymbol{y}_{k\ell:h}) \Big|_{\boldsymbol{y}_{k\ell:h}=\boldsymbol{v}} , \frac{\partial}{\partial y_0} g_{k\ell:h}(\boldsymbol{y}_{k\ell:h}) \Big|_{\boldsymbol{y}_{k\ell:h}=\boldsymbol{v}} \right]'$$
$$\text{and } \lim_{\boldsymbol{y}_{k\ell:h} \longrightarrow \boldsymbol{v}} \boldsymbol{\mathfrak{R}}_h(\boldsymbol{y}_{k\ell:h}),$$

with the same requirement for $g_{\ell k:h}(\boldsymbol{y}_{\ell k:h})$ at the diagonally reflected point $\breve{\boldsymbol{v}} = (v_2, v_1)$.

(e) There exists a bandwidth $\boldsymbol{b}_{k\ell:h:0}$ such that there for every $\mathbf{0} < \boldsymbol{b} < \boldsymbol{b}_{k\ell:h:0}$ is a unique minimiser $\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h:\boldsymbol{b}|p}$ of the penalty function $q_{\boldsymbol{v}|k\ell:h:\boldsymbol{b}|p}$ from eq. (2.9).

(f) The collection of bandwidths $\{\boldsymbol{b}_{k\ell:h:0}\}_{h\in\mathbb{Z}}$ has a positive infimum, i.e. there exists a $\boldsymbol{b}_{k\ell:0}$ such that[4]

$$\mathbf{0} < \boldsymbol{b}_{k\ell:0} \coloneqq \inf_{h\in\mathbb{Z}} \boldsymbol{b}_{k\ell:h:0}, \tag{2.17}$$

which implies that this $\boldsymbol{b}_{k\ell:0}$ can be used simultaneously for all the lags.

(g) For $X^{n|\boldsymbol{v}}_{k\ell:h:q:t}$ from definition 2.9, the related bivariate, trivariate and tetravariate density functions must be such that the expectations $\mathrm{E}\left[X^{n|\boldsymbol{v}}_{k\ell:h:q:t}\right]$, $\mathrm{E}\left[\left|X^{n|\boldsymbol{v}}_{k\ell:h:q:t}\right|^{\nu}\right]$ and $\mathrm{E}\left[X^{n|\boldsymbol{v}}_{k\ell:h:q:t} \cdot X^{n|\boldsymbol{v}}_{k\ell:i:r:s}\right]$ all are finite.

Remark 2.8. The present assumption 2.1 is in essence identical to (Jordanger and Tjøstheim, 2017, assumption 2.1) with some extra indices, so the remarks from (Jordanger and Tjøstheim, 2017) is of intereset here too. In particular, the α-mixing requirement in item (b) implies that $Y_{k,t+h}$ and $Y_{\ell,t}$ will be asymptotically independent as $h \to \infty$, i.e. the bivariate density functions $g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})$ will for large lags h approach the product of the marginal densities, and the situation will thus stabilise when h is large enough. This is in particular of importance for item (f), since it implies that it will be possible to find a nonzero $\boldsymbol{b}_{k\ell:0}$ that works for all h. Moreover, the finiteness assumptions in item (g) are always trivially satisfied if the required density-functions are finite.

2.5.2 An assumption for $\left(Y_{k,t}, Y_{\ell,t}\right)$ and the score function $\boldsymbol{u}_p(\boldsymbol{w};\boldsymbol{\theta}_p)$ of $\psi_p(\boldsymbol{w};\boldsymbol{\theta}_p)$

The following assumption is in essence identical to (Jordanger and Tjøstheim, 2017, assumption 2.2), which was included due to the need for the asymptotic results from Tjøstheim and Hufthammer (2013) to be applied for all the different lags h.

Assumption 2.2. The collection of local Gaussian parameters $\left\{\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\right\}$ at the point $\boldsymbol{v}$ for the bivariate probability density functions $g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})$, must all be such that

(a) $\boldsymbol{u}_p(\boldsymbol{v};\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}) \neq \mathbf{0}$ for all finite h.

(b) $\lim \boldsymbol{u}_p(\boldsymbol{v};\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}) \neq \mathbf{0}$.

Note that an inspection of the p equations in $\boldsymbol{u}_p(\boldsymbol{w};\boldsymbol{\theta}_p) = \mathbf{0}$ can be used to identify when items (a) and (b) of assumption 2.2 might fail, cf. the discussion in (Jordanger and Tjøstheim, 2017, section 2.4.2) for further details.

[4] Inequalities involving vectors are to be interpreted in a component-wise manner.

2.5.3 Assumptions for n, m and $\boldsymbol{b} = (b_1, b_2)$

The following assumption is identical to (Jordanger and Tjøstheim, 2017, assumption 2.3). The internal consistency of it was verified in (Jordanger and Tjøstheim, 2017, lemma C.3).

Assumption 2.3. Let $m \coloneqq m_n \to \infty$ be a sequence of integers denoting the number of lags to include, and let $\boldsymbol{b} \coloneqq \boldsymbol{b}_n \to \mathbf{0}^+$ be the bandwidths used when estimating the local Gaussian correlations for the lags $h = 1, \dots, m$ (based on n observations). Let b_1 and b_2 refer to the two components of $\boldsymbol{b}$, and let α, ν and a be as introduced in assumption 2.1(b). Let $s \coloneqq s_n \to \infty$ be a sequence of integers such that $s = o\Big(\sqrt{nb_1b_2/m}\Big)$, and let τ be a positive constant. The following requirements must be satisfied for these entities.[5]

(a) $\log n/n(b_1b_2)^5 \longrightarrow 0$, (only required for the case $p = 5$).
(b) $nb_1b_2/m \longrightarrow \infty$.
(c) $m^\delta(b_1 \vee b_2) \longrightarrow 0$, where $\delta = 2 \vee \frac{\nu(a+1)}{\nu(a-1)-2}$.
(d) $\sqrt{nm/b_1b_2} \cdot s^\tau \cdot \alpha(s-m+1) \longrightarrow \infty$.
(e) $m = o\big((nb_1b_2)^{\tau/(2+5\tau)-\lambda}\big)$, for some $\lambda \in (0, \tau/(2+5\tau))$.
(f) $m = o(s)$.

2.5.4 Convergence theorems for $\widehat{f}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)$, $\widehat{\alpha}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)$ and $\widehat{\phi}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)$

See appendix A for the proofs of the theorems stated below.

Theorem 2.10. *The estimate $\widehat{f}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) = \widehat{c}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) - i \cdot \widehat{q}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)$ of the local Gaussian cross-spectrum $f_{k\ell:\boldsymbol{v}|p}(\omega) = c_{k\ell:\boldsymbol{v}|p}(\omega) - i \cdot q_{k\ell:\boldsymbol{v}|p}(\omega)$, will under assumptions 2.1 to 2.3 satisfy*

$$\sqrt{n(b_1b_2)^{(p+1)/2}/m} \cdot \left(\begin{bmatrix} \widehat{c}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) \\ \widehat{q}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) \end{bmatrix} - \begin{bmatrix} c_{k\ell:\boldsymbol{v}|p}(\omega) \\ q_{k\ell:\boldsymbol{v}|p}(\omega) \end{bmatrix} \right) \xrightarrow{d} \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{c|k\ell:\boldsymbol{v}|p}(\omega) & 0 \\ 0 & \sigma^2_{q|k\ell:\boldsymbol{v}|p}(\omega) \end{bmatrix} \right), \tag{2.18}$$

when $\omega \notin \frac{1}{2} \cdot \mathbb{Z} = \big\{\dots, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, \dots\big\}$, where the variances $\sigma^2_{c|k\ell:\boldsymbol{v}|p}(\omega)$ and $\sigma^2_{q|k\ell:\boldsymbol{v}|p}(\omega)$ are given by

$$\sigma^2_{c|k\ell:\boldsymbol{v}|k\ell|p}(\omega) = \lim_{m\to\infty} \frac{1}{m} \left(\tilde{\sigma}^2_{\boldsymbol{v}|k\ell|p}(0) + \sum_{h=1}^{m} \lambda^2_m(h) \cdot \cos^2(2\pi\omega h) \cdot \big\{ \tilde{\sigma}^2_{\boldsymbol{v}|k\ell|p}(h) + \tilde{\sigma}^2_{\check{\boldsymbol{v}}|\ell k|p}(h) \big\} \right) \tag{2.19a}$$

$$\sigma^2_{q|k\ell:\boldsymbol{v}|k\ell|p}(\omega) = \lim_{m\to\infty} \frac{1}{m} \left(\sum_{h=1}^{m} \lambda^2_m(h) \cdot \sin^2(2\pi\omega h) \cdot \big\{ \tilde{\sigma}^2_{\boldsymbol{v}|k\ell|p}(h) + \tilde{\sigma}^2_{\check{\boldsymbol{v}}|\ell k|p}(h) \big\} \right), \tag{2.19b}$$

with $\tilde{\sigma}^2_{\boldsymbol{v}|k\ell|p}(h)$ and $\tilde{\sigma}^2_{\check{\boldsymbol{v}}|\ell k|p}(h)$ the asymptotic variances related to the estimates $\widehat{\rho}_{k\ell:\boldsymbol{v}|p}(h)$ and $\widehat{\rho}_{\ell k:\check{\boldsymbol{v}}|p}(h)$, see theorem B.1 for the details.

The local Gaussian quadrature spectrum is identical to zero when $\omega \in \frac{1}{2} \cdot \mathbb{Z}$, and for those frequencies the following asymptotic result holds under the given assumptions

$$\sqrt{n(b_1b_2)^{(p+1)/2}/m} \cdot \Big(\widehat{f}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) - f_{k\ell:\boldsymbol{v}|p}(\omega) \Big) \xrightarrow{d} \mathrm{N}\big(0, \sigma^2_{c|k\ell:\boldsymbol{v}|p}(\omega)\big), \qquad \omega \in \tfrac{1}{2} \cdot \mathbb{Z}. \tag{2.20}$$

[5] Notational convention: '$\vee$' denotes the maximum of two numbers, whereas '$\wedge$' denotes the minimum.

The asymptotic results for the local Gaussian amplitude- and phase-spectra is a direct consequence of theorem 2.10 and (Brockwell and Davis, 1986, proposition 6.4.3, p. 211).

Theorem 2.11. *Under assumptions 2.1 to 2.3, when* $\alpha_{k\ell:\boldsymbol{v}|p}(\omega) > 0$ *and* $\omega \notin \frac{1}{2} \cdot \mathbb{Z}$*, the estimate* $\widehat{\alpha}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) = \sqrt{\left(\widehat{c}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)\right)^2 + \left(\widehat{q}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)\right)^2}$ *satisfies*

$$\sqrt{n(b_1 b_2)^{(p+1)/2}/m} \cdot \left(\widehat{\alpha}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) - \alpha_{k\ell:\boldsymbol{v}|p}(\omega)\right) \xrightarrow{d} \mathrm{N}(0, \sigma^2_\alpha(\omega)), \tag{2.21}$$

where $\sigma^2_\alpha(\omega)$ *is given relative to* $\sigma^2_{c|k\ell:\boldsymbol{v}|p}(\omega)$ *and* $\sigma^2_{q|k\ell:\boldsymbol{v}|p}(\omega)$ *(from eq.* (A.9)*in theorem 2.10) as*

$$\sigma^2_\alpha = \left(c^2_{k\ell:\boldsymbol{v}|p}(\omega) \cdot \sigma^2_{c|k\ell:\boldsymbol{v}|p}(\omega) + q^2_{k\ell:\boldsymbol{v}|p}(\omega) \cdot \sigma^2_{q|k\ell:\boldsymbol{v}|p}(\omega)\right) / \alpha^2_{k\ell:\boldsymbol{v}|p}(\omega). \tag{2.22}$$

Theorem 2.12. *Under assumptions 2.1 to 2.3, when* $\alpha_{k\ell:\boldsymbol{v}|p}(\omega) > 0$ *and* $\omega \notin \frac{1}{2} \cdot \mathbb{Z}$*, the estimate* $\widehat{\phi}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) = \operatorname{args}\left(\widehat{c}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) - i \cdot \widehat{q}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)\right)$ *satisfies*

$$\sqrt{n(b_1 b_2)^{(p+1)/2}/m} \cdot \left(\widehat{\phi}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) - \phi_{k\ell:\boldsymbol{v}|p}(\omega)\right) \xrightarrow{d} \mathrm{N}\left(0, \sigma^2_\phi(\omega)\right), \tag{2.23}$$

where $\sigma^2_\phi(\omega)$ *is given relative to* $\sigma^2_{c|k\ell:\boldsymbol{v}|p}(\omega)$ *and* $\sigma^2_{q|k\ell:\boldsymbol{v}|p}(\omega)$ *(from eq.* (A.9)*in theorem 2.10) as*

$$\sigma^2_\phi(\omega) = \left(q^2_{k\ell:\boldsymbol{v}|p}(\omega) \cdot \sigma^2_{c|k\ell:\boldsymbol{v}|p}(\omega) + c^2_{k\ell:\boldsymbol{v}|p}(\omega) \cdot \sigma^2_{q|k\ell:\boldsymbol{v}|p}(\omega)\right) / \alpha^4_{k\ell:\boldsymbol{v}|p}(\omega). \tag{2.24}$$

Remark 2.9. The asymptotic normality results in theorems 2.10 to 2.12 do not necessarily help much if computations of pointwise confidence intervals for the estimated local Gaussian estimates are of interest, since it in practice may be unfeasible to find decent estimates of the variances $\sigma^2_{c|k\ell:\boldsymbol{v}|p}(\omega)$ and $\sigma^2_{q|k\ell:\boldsymbol{v}|p}(\omega)$ that occurs in theorem 2.10. The pointwise confidence intervals will thus later on either be estimated based on suitable quantiles obtained by repeated sampling from a known distribution, or they will be based on bootstrapping techniques for those cases where real data has been investigated. Confer Teräsvirta et al. (2010, ch. 7.2.5 and 7.2.6) for further details with regard to the need for bootstrapping in such situations.

3 Examples

This section will investigate if the m-truncated estimates of the local Gaussian cross-spectrum $f_{k\ell:\boldsymbol{v}|p}(\omega)$ might be of interest to consider when multiviarate time series are encountered. Since $\widehat{f}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)$ is complex-valued, the actual investigation will be based on plots of the corresponding local Gaussian versions of the cospectrum, quadrature spectrum, phase spectrum and the amplitude spectrum.

The present setup is similar to the one used for the investigation of the local Gaussian auto-spectrum, see (Jordanger and Tjøstheim, 2017, section 3), i.e. it will first be checked that the expected result is obtained when the origin of the data is a bivariate Gaussian time series. After this a simulation from a bivariate extension of the *local trigonometric* time series from (Jordanger and Tjøstheim, 2017, section 3.3.2) will be considered, and there will moreover be an example based on real multivariate data together with a simulation from a model based on these data. As in (Jordanger and Tjøstheim, 2017), it will be seen that plots related to the estimated $\widehat{f}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)$ might be useful as an exploratory tool, i.e. that this approach can detect nonlinear dependencies and periodicities between the variables.

Several parameters must be specified in order for an m-truncated estimate of $f_{k\ell:\boldsymbol{v}|p}(\omega)$ to be computed, and for the examples in the present paper the following values will be used:

1. p, the number of parameters in the local Gaussian approximation. Only the value $p = 5$ will be used in the present examples, since the results based on $p = 1$ in general fails to capture the local structure in a proper manner (see the discussion in (Jordanger and Tjøstheim, 2017, section 3.7) for further details).
2. $\boldsymbol{v} = (v_1, v_2)$, the points to investigate. The present investigation will consider points whose first and second coordinates correspond to the 10%, 50% and 90% percentiles of the standard normal distribution, i.e. the values are -1.28, 0 and 1.28. Information about the point of investigation is contained in the upper right corner of the relevant plots, where it will be marked as `10%::50%` and so on.
3. ω, the frequencies to investigate. Values between 0 and $\frac{1}{2}$.
4. $\boldsymbol{b} = (b_1, b_2)$, the bandwidth-vector to be used when computing the local Gaussian autocorrelations. The value $\boldsymbol{b} = (.6, .6)$ has been used for all of the cases in this paper.
5. m, the truncation level, i.e. the number of lags to include in the estimate $\widehat{f}^{\,m}_{\boldsymbol{v}|p}(\omega)$. The value $m = 10$ has been used in this investigation, and this number is by default given in the upper left corner of the relevant plots.
6. $\lambda_m(h)$, the weighting function to be used for the smoothing of the different lags. The Tukey-Hanning lag-window kernel has been used for all the present examples, i.e.

$$\lambda_m(h) = \begin{cases} \frac{1}{2} \cdot \left(1 + \cos\left(\pi \cdot \frac{h}{m}\right)\right) & |h| \leq m, \\ 0 & |h| > m. \end{cases}$$

Remark 3.1. This list of parameters is similar to the one used when estimating the local Gaussian auto-spectra in (Jordanger and Tjøstheim, 2017, section 3.1), with the main exception that the value 0.6 is used instead of 0.5 in the bandwidth vector $\boldsymbol{b}$. This adjustment is partially due to the fact that the time series in this paper are sligthly shorter than those used in (Jordanger and Tjøstheim, 2017), i.e. length 1859 versus length 1974. (All the time series have the same length as the one encountered for the real sample.)

Remark 3.2. It was noted in (Jordanger and Tjøstheim, 2017) that it was natural to require that the bandwidth $\boldsymbol{b} = (b_1, b_2)$ should satisfy $b_1 = b_2$ when the local Gaussian autocorrelations $\rho_{kk:\boldsymbol{v}|p}(h)$ should be estimated, since both of the components in the lag h pseudo-normalised pairs originated from the same univariate time series. For the estimation of local Gaussian cross-correlations $\rho_{k\ell:\boldsymbol{v}|p}(h)$, it is the pseudo-normalisation of the marginals that justifies the assumed equality of b_1 and b_2.

Remark 3.3. The *pointwise confidence bands*[6] shown in the plots later on are all based upon $R = 100$ replicates. Repeated independent samples from the known model was used to construct the confidence bands in section 3.3, whereas block-bootstrap was used for the real data example in section 3.4. The lower and upper limits of the pointwise confidence bands are based on the 0.05 and 0.95 quantiles of the resulting collection of estimated local Gaussian spectral densities (truncated at lag m), and thus gives estimated 90% pointwise confidence bands for $c^m_{k\ell:\boldsymbol{v}|p}(\omega)$, $q^m_{k\ell:\boldsymbol{v}|p}(\omega)$, $\alpha^m_{k\ell:\boldsymbol{v}|p}(\omega)$, and $\phi^m_{k\ell:\boldsymbol{v}|p}(\omega)$.

[6] The pointwise confidence band gives for each frequency ω a confidence interval for the values of $c^m_{k\ell:\boldsymbol{v}|p}(\omega)$, $q^m_{k\ell:\boldsymbol{v}|p}(\omega)$, $\alpha^m_{k\ell:\boldsymbol{v}|p}(\omega)$, and $\phi^m_{k\ell:\boldsymbol{v}|p}(\omega)$.

Remark 3.4. It will in general be hard to know in advance which input parameters it would be natural to employ for a given time series, and it is thus necessary and recommended to find estimates $\widehat{f}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)$ for several different parameter combinations and then use an interactive tool to inspect the corresponding plots to see if some interesting features might be present. The `R`-package `localgaussSpec`[7] has been created to take care of both of these parts, i.e. it allows an integrated interactive investigation of the results by means of a `shiny`-application.[8]

Remark 3.5. The scripts that generated the plots presented in this paper is included as a part of the `R`-package `localgaussSpec`, so the interested reader can run these and check out how the plots change when the input parameters in the estimation algorithm are modified. Note that any conclusions based on individual static plots, like those presented in this paper, in general should be considered with some caution, in particular when there is a lack of data-driven methods that can justify the parameter-configuration that was used in the estimation-algorithm. Moreover, it is important to be aware of the fact that small-sample variation might occur not only if an estimate is computed for a time series sample that is too short, but that it also can be small-sample variation that dominates if the point $\boldsymbol{v}$ lies too far out in the periphery, or if the truncation level m and the bandwidth $\boldsymbol{b}$ are not keept within reasonable limits. See the discussion in section 4 for further details.

Remark 3.6. The `R`-package `localgauss`, see Berentsen et al. (2014), was used for the estimation of the local Gaussian auto- and cross-correlations for the $p = 5$ case. These estimates are returned with an indicator (named `eflag`) that reveals whether or not the estimation algorithm converged numerically to the estimate, and this numerical convergence-information has then been added to the relevant plots in their lower left corner. In particular, '`NC = OK`' will be used to show that all the required estimates had a successful numerical convergence. Contrary, '`NC = FAIL`' will represent that problems did occur during the estimation algorithm. It should be noted that convergence-problems hardly occurs when the computations are based on pseudo-normalised observations.

3.1 Some basic simulations

This section will check that the estimates of $f^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)$ behaves as expected for a few simple simulated bivariate examples, where the underlying models in essence are bivariate extensions of the models encountered in (Jordanger and Tjøstheim, 2017, section 3.3), and it will be seen that a joint inspection of the `Co`-, `Quad`- and `Phase`-plots might be useful as an exploratory tool for nonlinear dependencies in multivariate time series. The `Amplitude`-plots have not been included here since the interesting details (in most cases) already would have been detected by the other plots.

3.1.1 Bivariate Gaussian white noise

It is known from lemma 2.2(a) that the local Gaussian cross-spectrum coincide with the ordinary cross-spectrum when the time series under investigation is Gaussian. The plots in fig. 1 shows the `Co`-, `Quad`- and `Phase`-plots based on 100 independent samples of length 1859 from a bivariate Gaussian distribution with standard normal marginals and correlation 0.35. The left column of fig. 1 shows the situation for a point off the diagonal, whereas the right column

[7] See footnote 1 (page 3) for details about the installation of the `localgaussSpec`-package.
[8] See Chang et al. (2017) for details about `shiny`.

shows the situation for a point at the center of the diagonal, i.e. $v_1 = v_2 = 0$. Note that the global spectra are identical for all the points, i.e. the red components are the same for each row of fig. 1.

In this simple case, where the true values of the local Gaussian versions of the spectra coincides with the ordinary global spectra, it follows that the `Co`-, `Quad`- and `Phase`-spectra (for any truncation level m) respectively should be the constants 0.35, 0 and 0. Figure 1 shows that the red and blue dotted lines, that respectively represents the estimates of the global and local m-truncated spectra,[9] seems to match these true values quite reasonably – and this provides a sanity check of the code that generated these plots. Note that the 90% pointwise confidence interval for the local Gaussian versions (blue ribbons) are wider than those for the ordinary spectra (red ribbons) since the bandwidth used for the estimation of the local Gaussian cross-correlations, in this case $\boldsymbol{b} = (0.6, 0.6)$, reduces the number of observations that effectively contributes to the computation of the local Gaussian spectra.

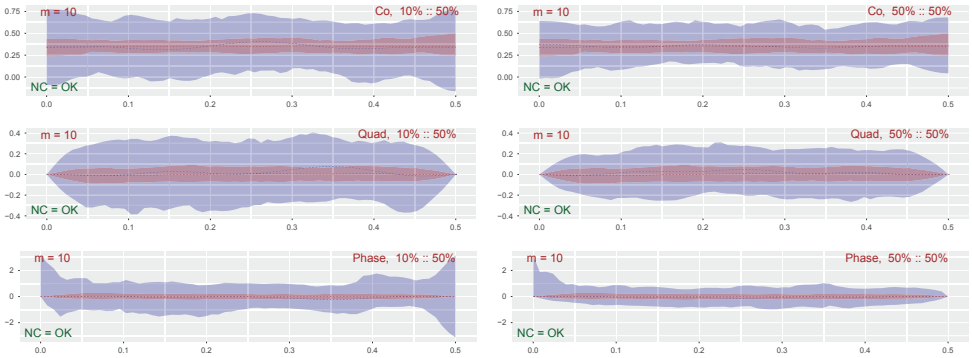

Figure 1: i.i.d. bivariate Gaussian white noise.

3.1.2 Bivariate local trigonometric examples

With the exception of the Gaussian time series, it is not known what the true local Gaussian spectral densities should look like – which makes it hard to know whether or not the `Co`-, `Quad`- and `Phase`-plots looks like they are expected to do. However, it is possible to extend to the bivariate case the strategy that was outlined in (Jordanger and Tjøstheim, 2017, section 3.3.2), i.e. that *bivariate local trigonometric time series* can be constructed for which it, at some designated points $\boldsymbol{v}$, can be given a heuristic argument for the expected shape of the local Gaussian spectra.

The heuristic argument needed for the bivariate case is identical in structure to the one used in the univariate case, and for the present case the reference for the plots later on is based on the following simple bivariate model,

$$Y_{1,t} = \cos(2\pi\alpha t + \phi) + w_t \text{ and } Y_{2,t} = \cos(2\pi\alpha t + \phi + \theta) + w_t, \tag{3.1}$$

[9] The dotted lines represents the medians of the estimated values, whereas the 90% pointwise confidence intervals are based on the 5% and 95% quantiles of these samples.

where w_t is Gaussian white noise with mean zero and standard deviation σ, and where it in addition is such that α and θ are fixed for all the replicates whereas ϕ is drawn uniformly from $[0, 2\pi)$ for each individual replicate. A realisation with $\sigma = 0.75$, $\alpha = 0.302$ and $\theta = \pi/3$ has been used for the `Co-`, `Quad-`, and `Phase`-plots shown in fig. 2, where 100 independent samples of length 1859 were used to get the estimates of the m-truncated spectra and their corresponding 90 pointwise confidence intervals (based on the bandwidth $\boldsymbol{b} = (0.6, 0.6)$). Some useful remarks can be based on this plot, before *the bivariate local trigonometric case* is defined and investigated.

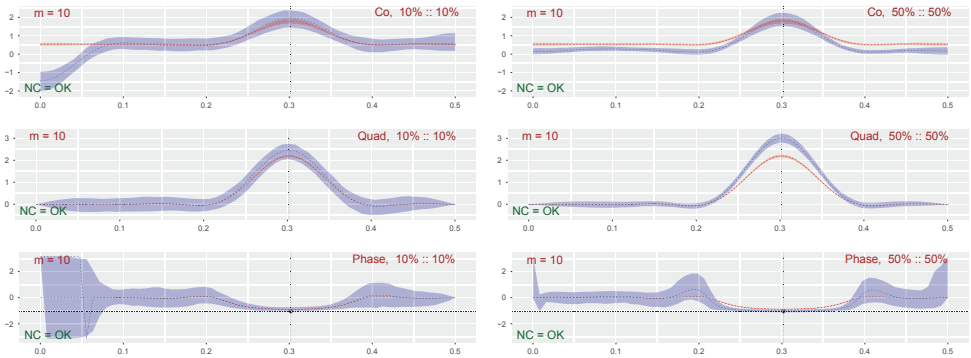

Figure 2: Realisation of eq. (3.1), with $\alpha = 0.302$ and $\theta = \pi/3$.

Remark 3.7. In this particular case, the local Gaussian spectra in fig. 2 shares many similarities with the corresponding global spectra. In particular, the peak of the `Co-` and `Quad`-plots lies both for the local and global spectra at the frequency $\omega = \alpha$ (shown in the plots as a vertical line), and the corresponding `Phase`-plots at this frequency lies quite close to the phase-adjustment $\theta = \pi/3$ (shown as a horizontal line, positioned with an appropriate sign adjustment). This phenomenon is present both for the point at `10%::10%` and the point at `50%::50%`, but it should be noted that this nice match does not hold for all values of σ. In fact, experiments with different values for σ (plots not included in this paper) indicates that the difference between the local and global spectra becomes larger (in particular for the point `10%::10%`) when σ becomes smaller. See (Jordanger and Tjøstheim, 2017, fig. 9) for an example of the results that can occur.

Remark 3.8. The values of the `Co-` and `Quad`-plots (for a given frequency ω) are (for each frequency) related to the corresponding values of the `Amplitude-` and `Phase`-plots by the following simple relations,

$$-\texttt{Quad-plot}/\texttt{Co-plot} = \tan(\texttt{Phase-plot}), \tag{3.2a}$$

$$\texttt{Co-plot} = \texttt{Amplitude-plot} \cdot \cos(\texttt{Phase-plot}), \tag{3.2b}$$

$$-\texttt{Quad-plot} = \texttt{Amplitude-plot} \cdot \sin(\texttt{Phase-plot}), \tag{3.2c}$$

which follows trivially from the way these spectra are defined relative to Cartesian or polar representations of the complex-valued cross-spectra, cf. eq. (1.2) and definition 2.3. For the example investigated in fig. 2, where the `Phase` plot is close to $-\pi/3$ at $\alpha = 0.302$, it thus

follows that the peak for the Quad-spectrum should be approximately $\sqrt{3}$ times larger than the peak of the Co-spectrum.

Remark 3.9. It can be enlightening to compare the Co-, Quad- and Phase-plots in fig. 2 with a plot that shows the underlying estimates upon which the pointwise confidence intervals were based. Such a plot is shown in fig. 3, where the left panel presents the complex-valued estimates of the local Gaussian cross-spectrum at the frequency $\omega = \alpha$, and where medians and quantiles relative to a polar representation, i.e. $z = re^{i\theta}$, have been added to the plot. The center panel shows the same estimated values, but this time the median/quantiles are based on a Cartesian representation $z = x + iy$. These two panels gives a geometrical view to the observations presented in remark 3.8. The third panel of fig. 3 presents a zoomed in version of the estimated values of the local Gaussian cross-spectrum, and it gives a reminder that it in principle is possible to extract more information from these estimates than what has been done so far. A closer inspection of these estimates could e.g. be used to see how much they (for the given m-truncation) deviate from the expected asymptotic distributions that was given in section 2.5.4.

Remark 3.10. The wide pointwise confidence band observed for ω near 0 in the 10%::10%-Phase-plot, is related to the branch-cut that occurs at $-\pi$ in the definition of the phase-spectrum, cf. definition 2.3. The simple algorithm used for the creation of the pointwise confidence intervals has not been tweaked to properly cover the case where the majority of the estimates lies in the second and third quadrants of the complex plane, which implies that the Co- and Quad-plots should be consulted instead when the Phase-plot misbehaves in this manner.

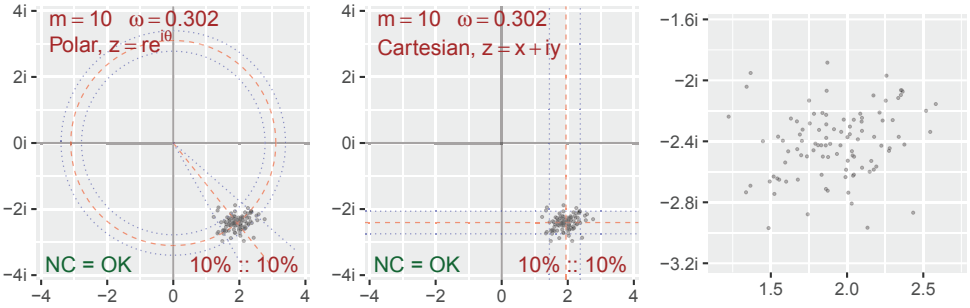

Figure 3: Complex-valued representation of 100 samples of $f_{12;v}(\omega)$ from eq. (3.1), at the peak frequency $\omega = 0.302$. Left panel: Pointwise 90% confidence bands based on polar representation. Center panel: Pointwise 90% confidence bands based on Cartesian representation. Right panel: Zoomed in plot.

The bivariate local trigonometric case: Two bivariate extensions of the artificial *local trigonometric* time series from (Jordanger and Tjøstheim, 2017, section 3.3.2) will now be considered. The parameters in these time series can be selected in such a manner that the global spectra are similar to those encountered for white noise, whereas the local Gaussian spectra contains different peaks at different points. The same heuristic argument that was used in (Jordanger and Tjøstheim, 2017) implies that it is *expected* that the local periodicity at some designated points should resemble those in fig. 2, and samples from these models can thus be

used to check if the `Co-`, `Quad-` and `Phase`-plots can detect these local properties. It will be seen that the answer indeed is affirmative for the designated points for both of the models considered here, see figs. 4 and 6, but it will also be noted (see fig. 5) that the plots can be quite different when other points are investigated.

The algorithm for the bivariate samples $\left(Y_{1,t}, Y_{2,t}\right)$ (defined below) is based on the univariate algorithm from (Jordanger and Tjøstheim, 2017), i.e. the marginals $Y_{1,t}$ and $Y_{2,t}$ will both be generated by the same *local trigonometric formula* that was used in (Jordanger and Tjøstheim, 2017, section 3.3.2), but with the distinction that an additional phase-adjustment is used for the second component. To clarify this, here follows the required modification of the univariate setup from (Jordanger and Tjøstheim, 2017), with the specification of the parameters used in the present computations.

The *bivariate local trigonometric* time series are constructed by the following principle: For a given $r \geq 2$, first select a collection of different base levels $(L_1, \ldots, L_r)$ at the y-axis, and a collection of amplitudes $(A_1, \ldots, A_r)$ and amplitude adjustments $(A'_1, \ldots, A'_r)$. Create a stochastic amplitude $A_i(t)$, for each t, by selecting uniformly a value from the interval spanned by A_i and A'_i. Then select a collection of frequencies $(\alpha_1, \ldots, \alpha_r)$, and two collections of phase-adjustments $(\phi_1, \ldots, \phi_r)$ and $(\theta_1, \ldots, \theta_r)$. Finally, assign a probability p_i to each $i = 1, \ldots, r$, such that $\sum_{i=1}^{r} p_i = 1$. These ingredients enables the definition of the following functions for $i = 1, \ldots, r$,

$$C_{1,i}(t) = L_i + A_i(t) \cdot \cos\left(2\pi\alpha_i t + \phi_i\right), \tag{3.3a}$$

$$C_{2,i}(t) = L_i + A_i(t) \cdot \cos\left(2\pi\alpha_i t + \phi_i + \theta_i\right), \tag{3.3b}$$

from which the stochastic variables $Y_{1,t}$ and $Y_{2,t}$ can be created by means of the probabilities $(p_1, \ldots, p_r)$, i.e. let N_t be a random variable that with probability p_i takes the value i, and define

$$Y_{1,t} \coloneqq \sum_{i=1}^{r} C_{1,i}(t) \cdot \mathbb{1}\{N_t = i\}, \tag{3.4a}$$

$$Y_{2,t} \coloneqq \sum_{i=1}^{r} C_{2,i}(t) \cdot \mathbb{1}\{N_t = i\}. \tag{3.4b}$$

The indicator function $\mathbb{1}\{\}$ ensures that only one of the r functions in the sum contributes for a given value t. Note that it is assumed that the phases ϕ_i are uniformly drawn (one time for each realisation) from the interval between 0 and 2π, and that it moreover also is assumed that the stochastic processes ϕ_i, $A_i(t)$ and N_t are independent of each other. Based on this, the auto- and cross-covariances can can be given as functions of L_i and p_i, from which it then is fairly easy to select a combination of input parameters that returns a $\left(Y_{1,t}, Y_{2,t}\right)$-process that looks like white noise.

The generating models for the two time series presented in this section both have $r = 4$ components with base levels L_i in $(-2, -1, 0, 1)$, amplitude-functions $A_i(t)$ defined by A_i in $(1.0, 0.5, 0.3, 0.5)$ and A'_i in $(0.5, 0.2, 0.2, 0.6)$, and frequencies α_i in $(0.267, 0.091, 0.431, 0.270)$.

The probabilities p_i in $(0.05, 0.28, 0.33, 0.33)$ was used to sample[10] which component to include in $Y_{1,t}$ and $Y_{2,t}$.

The distinction between the two models are due to the selection of the additional phase-adjustments θ_i. The model investigated in figs. 4 and 5 have a constant phase adjustment of $\theta = \pi/3$, whereas the model investigated in fig. 6 have individual phase-adjustments given as $(\theta_1, \theta_2, \theta_3, \theta_4) = (\pi/3, \pi/4, 0, \pi/2)$.

To complete the specification of the setup, note that the 90% pointwise confidence intervals in figs. 4 to 6 all are based on 100 independent samples of length 1859 from the above described models, and that the bandwidth $\boldsymbol{b} = (0.6, 0.6)$ was used in the computation of the local Gaussian cross-correlations.

Constant phase adjustment: The case where the phase difference $\theta = \pi/3$ was used for all the θ_i is investigated in figs. 4 and 5. Figure 4 (see page 22) shows the `Co-`, `Quad-` and `Phase`-plots for the three designated points `10%::10%`, `50%::50%` and `90%::90%`, for which the heuristic argumentation implies that the results should look a bit like the situation encountered in fig. 2.

Remark 3.11. The three points investigated in fig. 4 corresponds to the function-components in eq. (3.3) with indices $i = 2, 3, 4$. The point that corresponds to the $i = 1$ component would, due to the combination of the low probability p_1 and the placement of the level L_1, lie too far out in the tail to be properly investigated by the present sample size and truncation point. The $i = 1$ component was included in the example in (Jordanger and Tjøstheim, 2017) in order to show that an exploratory approach based on local Gaussian spectra can fail to detect local signals that are much weaker than the dominating ones.

For the points investigated in fig. 4, it seems to be the case that the local Gaussian part of the `Co-`, `Quad-` and `Phase`-plots together reveal local properties in accordance with the outcome expected from the knowledge of the generating model – and these local structures are not detected by the ordinary global spectra, which in this case (due to the values used for L_i and p_i) are flat. The left column investigates a point at the lower tail of the diagonal, and it can there be observed that both the `Co-` and `Quad`-plots have a peak close to the leftmost α-value – and the value of the corresponding `Phase`-plot for frequencies close to this α-value lies quite close to the phase difference between the first and second component. A similar situation is present for the three plots shown in the right column, where a point at the upper tail of the diagonal are investigated. Moreover, in accordance with the general observation in remark 3.8, the peaks of the `Quad`-plots are higher than those of the `Co`-plots in this case due to the phase-difference θ that was used in the input parameters.

For the center column of fig. 4, which investigates the point at the center of the diagonal, it can be seen that the `Quad`- and `Phase`-plots in addition to the expected α-value also detects the presence of the other α-values. The `Phase`-plot is for this point not that close to the expected value, but that situation changes if the truncation is performed at a higher lag than $m = 10$. The center column thus shows the importance of considering a range of values for the truncation point when such plots are investigated. The additional peaks that are detected in the center column are due to *contamination* from the neighbouring regions.

[10] The printed probabilities might not add to one! This is due to the fact that these values was rounded in `R` before they were included in this document by the means of the `R`-package `knitr`, see (Xie, 2015, 2016) for details about dynamic documents.

Figure 5 present plots based on the same sample that was used for fig. 4, but now the investigated points are no longer among the designated ones on the diagonal. The plots in fig. 5 shows that the Co-, Quad- and Phase-plots at the point 10%::90% looks more like the i.i.d. white noise that was encountered in fig. 1, whereas the plots for the two points 10%::50% and 50%::90% does detect the presence of local phenomena. It might not be any obvious interpretation of these plots when seen isolated, but it should at least be noted that the plots for the two points 10%::50% and 50%::90% have troughs for the α-values that corresponds to the first and second coordinates of these points – and this seen in conjunction with the previously investigated points in fig. 4 does support the idea that there are local features in the data that depends on these α-values.

Individual phase adjustment: Figure 6 investigates the same designated points as fig. 4 did, but now the samples have been generated from the model where the second variate used the individual phase-adjustments specified in $(\theta_1, \theta_2, \theta_3, \theta_4) = (\pi/3, \pi/4, 0, \pi/2)$. Horizontal lines have been added to the Phase-plots to show all of these θ_i-values (adjusted to have the correct sign), and for each of the designated points the intersection with the relevant vertical α_i-line has been highlighted to show the expected outcome based on the knowledge of the model.

The Co-, Quad- and Phase-plots in fig. 6 does behave in accordance with what was observed in fig. 4, i.e. the Phase-plots lies close to the expected θ_i-value when the frequency ω is near the corresponding α_i-value, and the height of the corresponding Co- and Quad-peaks are in accordance with the values of the Phase-plots. In particular, the phase-adjustment is $\theta_2 = \pi/4$ for the point 10%::10%, which implies that the Co- and Quad-peaks should rise approximately to the same height above their respective baselines, which seems to be fairly close to the observed result. For the points 50%::50% and 90%::90% the situation is clearer since the respective local frequencies $\theta_3 = 0$ and $\theta_4 = \pi/2$ then implies that only the Co-plot should have a peak for the point 50%::50% and only the Quad-plot should have a peak for the point 90%::90%, again in agreement with the impression based on fig. 6.

Remark 3.12. The examples investigated in figs. 2 to 6 do not satisfy the requirements needed for the asymptotic results (both for the global and local cases) to hold true, in particular the local Gaussian cross-correlations will in these cases not be absolutely summable.[11] Despite this, the examples are still of interest since they shows that an exploratory tool based on the local Gaussian spectra in this case can detect information that is in agreement with what could be expected based on the parameters used in the models for $\left(Y_{1,t}, Y_{2,t}\right)$.

Remark 3.13. The generating model will of course not be known when a real multivariate time series is encountered, so it is important to estimate the local Gaussian cross-spectrum at a wide range of points $\boldsymbol{v}$ in the plane and a wide range of truncation levels m. This is necessary since it can happen, like seen in fig. 5, that there are some points where the result might look like it has been computed based on i.i.d. white noise, whereas other points shows that some local phenomenon could be present.

Remark 3.14. Even when it might not be obvious how to interpret the results shown in the Co-, Quad- and Phase-plots, it should be noted that they can be used as an exploratory tool that can detect nonlinear traits in the observations. Moreover, these plots can also be used to investigate if a model fitted to the data contains elements that can mimic the observed features. The recipe

[11] In this respect the situation is similar to the detection of a pure sinusoidal for the global spectrum.

for this approach would then be to first select a model, then estimate parameters based on the available sample, and finally use the resulting fitted model to generate independent samples of the same length as the sample. Section 3.2 will show an example of this approach, cf. figs. 9 and 12.

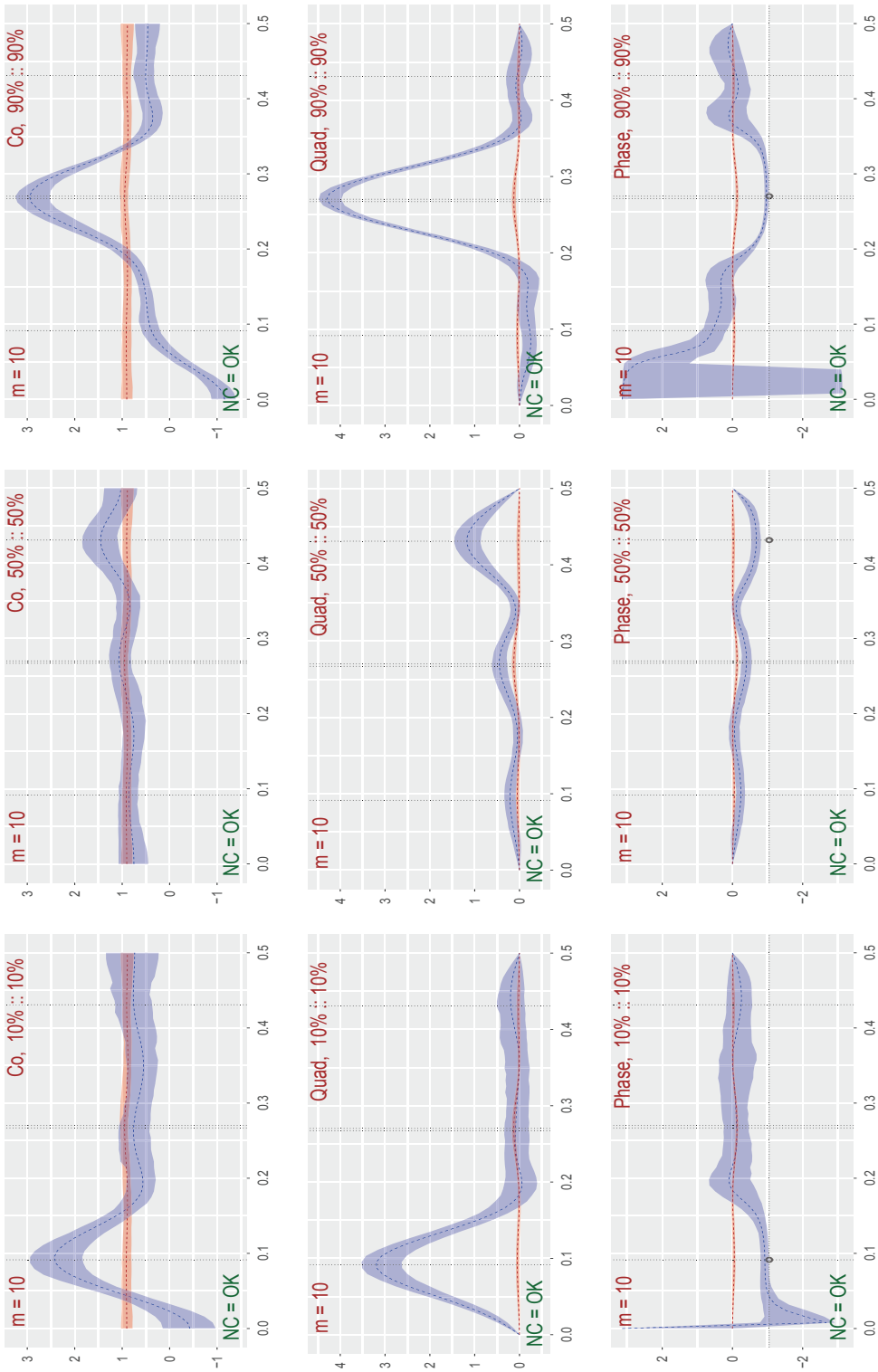

Figure 4: Plots related to eq. (3.4), common phase adjustment, designated points along the diagonal.

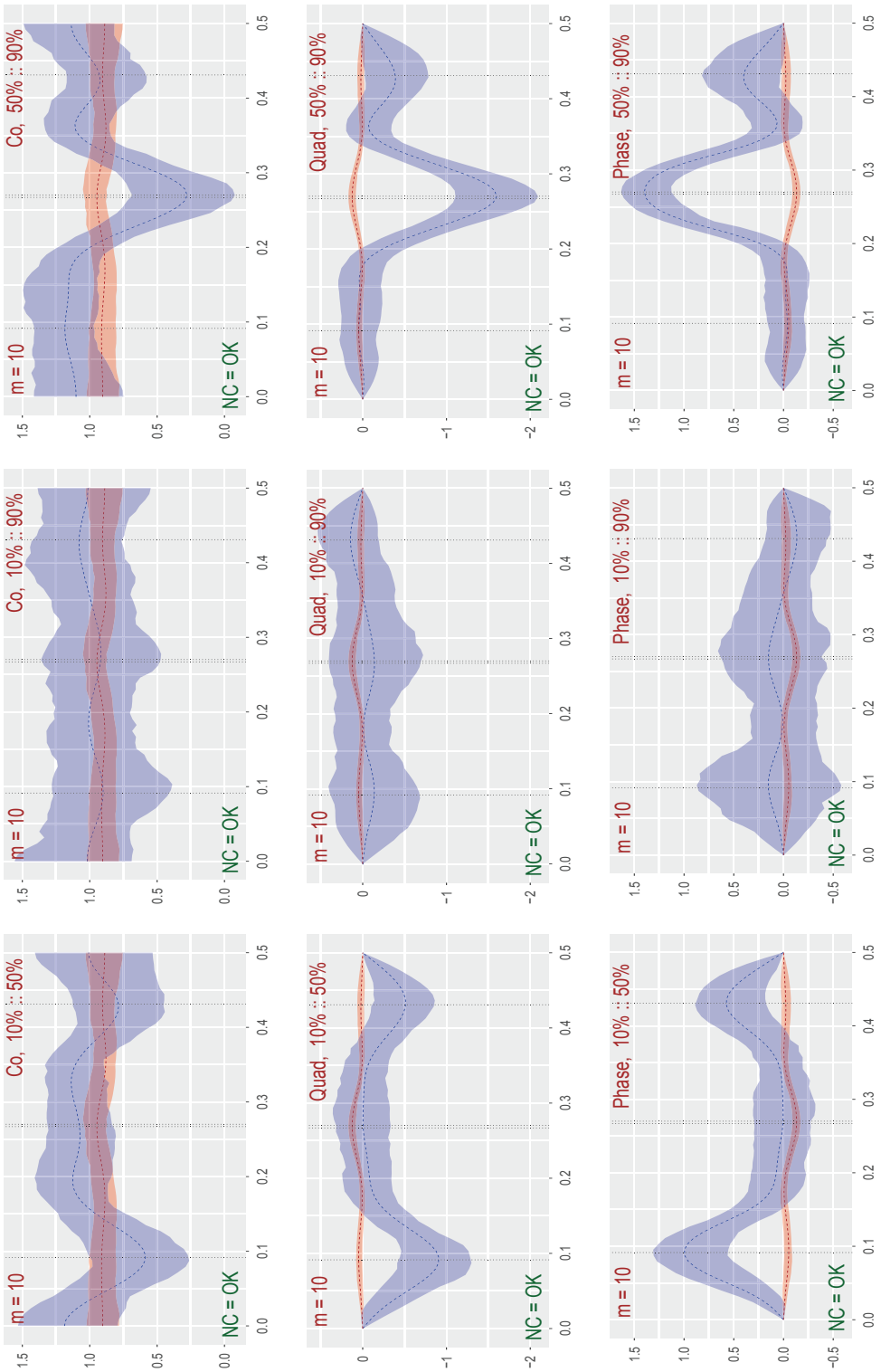

Figure 5: Plots related to eq. (3.4), common phase adjustment, behaviour away from the designated points.

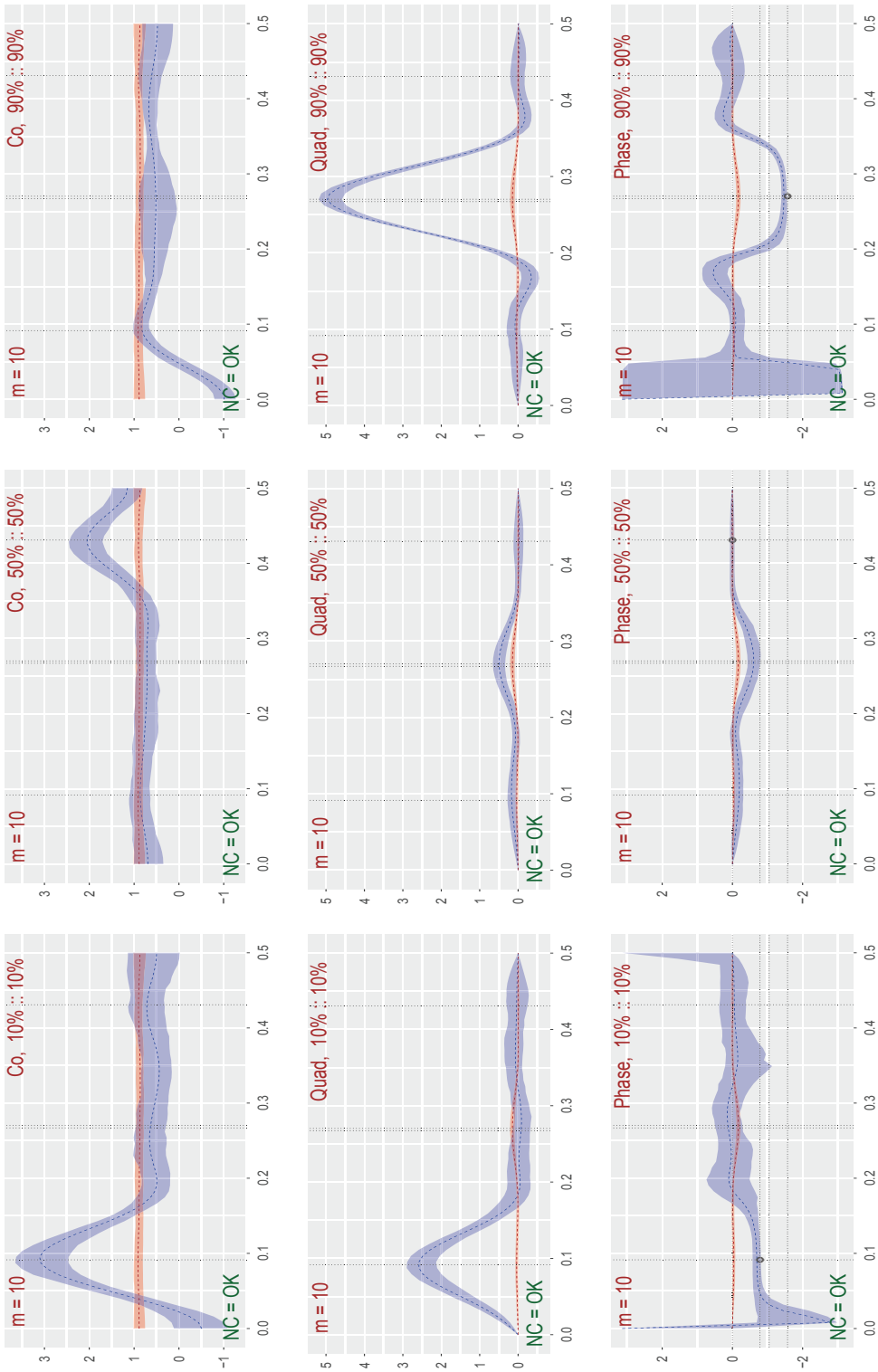

Figure 6: Plots related to eq. (3.4), individual phase adjustments, designated points along the diagonal.

3.2 A real multivariate time series and a (badly) fitted GARCH-type model

This section will show how the Co-, Quad- and Phase-plots can be used as an exploratory tool on some financial data, and then it will be seen how this approach can be used to get a visual impression of the quality of a multivariate GARCH-type model fitted to these data.

The multivariate time series sample to be considered in this section will be a bivariate subset of the (log-returns of the) tetravariate EuStockMarkets-sample from the datasets-package of R, R Core Team (2017). This data-set has been selected since it has a length that should be large enough to justify the assumption that the observed features in the Co-, Quad- and Phase-plots are not solely there due to small sample variation.

The EuStockMarkets contains 1860 daily closing prices collected in the period 1991-1998, from the following four major European stock indices: Germany DAX (Ibis), Switzerland SMI, France CAC, and UK FTSE. The data was sampled in business time, i.e., weekends and holidays was omitted.

The log-returns of the EuStockMarkets values gives a tetravariate data-set that it seems natural to model with some multivariate GARCH-type model, and the R-package rmgarch, Ghalanos (2015b), was thus used for that purpose. Note that the present paper only aims at showing how this kind of analysis can be performed, so only one very simple model was investigated – which thus gave a rather badly fitted model for the data at hand.

3.2.1 The DAX-CAC subset of the EuStockMarkets-log-returns

The log-returns of the bivariate EuStockMarkets-subset $(Y_1, Y_3) = (\text{DAX}, \text{CAC})$, of length 1859, will now be investigated. The individual pseudo-normalised traces of these observations are shown in fig. 7, and it will be from these pseudo-normalised observations that the local Gaussian cross-correlations will be computed.

The plots of the resulting m-truncated global and local spectra are shown as the red and blue lines in fig. 9, and the 90% pointwise confidence intervals has been created based on 100 block-bootstrap replicates using a block-length of 100.

No data-driven algorithm is present for the selection of the blocklength in this case, cf. section 4.1, and the rationale for the use of a given blocklength is thus based on a visual inspection of fig. 8, where the plot of the local Gaussian cross-correlations (based on the bandwidth $\boldsymbol{b} = (0.6, 0.6)$ for all the lags) are presented for the three points that will be investigated in fig. 9. It seems plausible, based on an inspection of fig. 8, that a blocklength of 100 might be needed in this case – but note that a more rigorous investigation should have been employed if this was to be an actual in depth analysis of the dependence on the blocklength.

The three points considered in fig. 9 all lie on the diagonal, since it seems easier to give an interpretation for those points. In particular, the point 10%::10% represent a situation where the market goes down both in Germany and France, whereas the points 50%::50% and 90%::90% similarly represent cases where the market either is stable or goes up in both countries.

Thus, for the purpose of the present paper, it suffices to point out that the Co-, Quad- and Phase-plots of fig. 9 indicates that the data contains nonlinear traits, which in particular is visible in the Co-plot for the point 10%::10% and for all the plots related to the point 90%::90%. It should be noted that the Co-plots at the frequency $\omega = 0$ simply gives a weighted sum of the local Gaussian cross-correlations (between $\left(Y_{1,t+h}, Y_{3,t}\right)$) seen in fig. 8, so the Co-plot peaks at

$\omega = 0$ for the points `10%::10%` and `90%::90%` are thus as expected, and the lack of a `Co`-plot peak at $\omega = 0$ for the point `50%::50%` also seems natural in view of fig. 8. It should also be noted that the $\omega = 0$ peak for the `90%::90%` `Co`-plot is lower than the corresponding peak for `10%::10%`, but this seems in this case to be due to the low truncation level used for the plots, i.e. these two peaks attain approximately the same height when a higher truncation level is applied.

It seems, for the particular parameter-configuration that generated the plots in fig. 9, to be the case that the point `90%::90%` has the most interesting `Quad`- and `Phase`-plots, but again, as noted above, it may be premature to put too much emphasis on this particular plot given the uncertainties involved in the selection of the bandwidth $\boldsymbol{b}$, the truncation level m, and the block-length to be used in the bootstrap. In particular the block-length is critical with regard to this, since a shorter block-length tends to give wider confidence-intervals, and that might result in plots where a seemingly significant difference between the local and global spectra disappears.

It should also be noted that a low number of bootstrapped replicates can be a source of small sample variation for the width of the estimated pointwise confidence intervals, and this is important to keep in mind if a minor gap is observed between the pointwise confidence intervals for the local and global spectra. Such gaps could appear or disappear when the algorithm is used to generate new computations based on the same number of bootstrapped replicates, a behaviour that in particular has been observed for the rightmost peak/trough of the `Quad`- and `Phase`-plots at the point `90%::90%` in fig. 9.

This kind of ambiguity can be countered by increasing the number of bootstrapped replicates, but that has not been done for the present example due to the uncertainty with regard to the size of the selected blocklength. In particular, if fig. 9 is recreated with the blocklength 50 instead of 100,[12] then the pointwise confidence interval for the `Phase`-plot at the point `90%::90%` widens enough to remove those occurrences where it seems that it could be a significant trough at the rightmost side. The peak at the center of the `90%::90%` `Phase`-plot does however remain significantly different from the global spectrum even with a smaller blocklength, which strengthens the impression that something of interest might be present at that frequency. However, it is important to keep in mind that this impression is based on the present combination of bandwidth and truncation level – and there are at the moment no data-driven method for the selection of these parameters. (A positive phase difference is consistent with the `90%::90%` cross-correlation plot in fig. 8, which might indicate that the DAX is leading over CAC when the market is going up.)

Remark 3.15. Note that the `shiny`-interface in the `R`-package `localgaussSpec` should be used if it is of interest to pursue a further analysis of the local Gaussian spectra of the (log-returns of the) `EuStockMarkets`-data, since that enables an interactive investigation that shows how the estimates vary based on different bandwidths $\boldsymbol{b}$ and truncation levels m, and moreover, it would also allow an investigation of how much the selection of the block-length for the bootstrap-procedure influences the widths of the pointwise confidence intervals in the `Co`-, `Quad`- and `Phase`-plots.

[12] Based on the lower panel of fig. 8, it might be ample reason to consider a block-length of 50 to be too short.

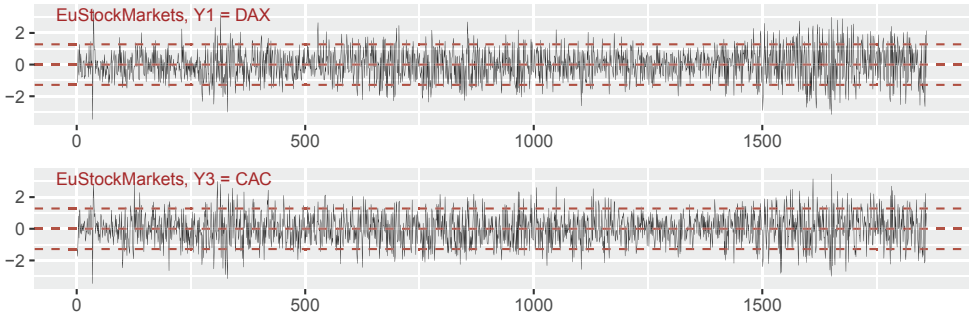

Figure 7: EuStockMarkets, pseduo-normalised components under investigation.

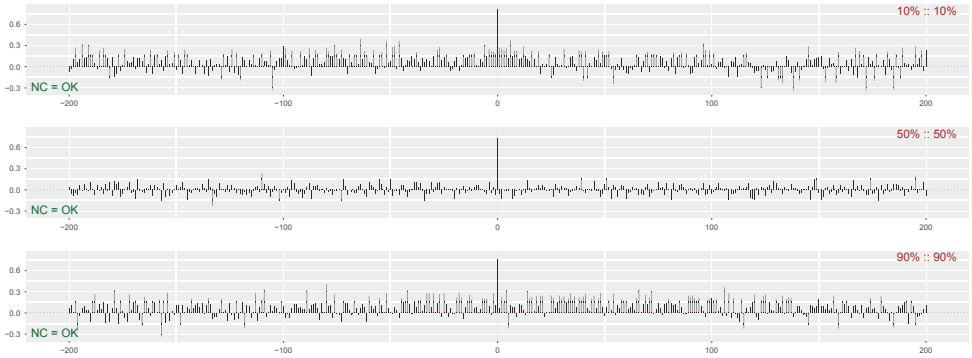

Figure 8: EuStockMarkets, local Gaussian cross-correlations between $\left(Y_{1,t+h}, Y_{3,t}\right)$, i.e. DAX is leading CAC.

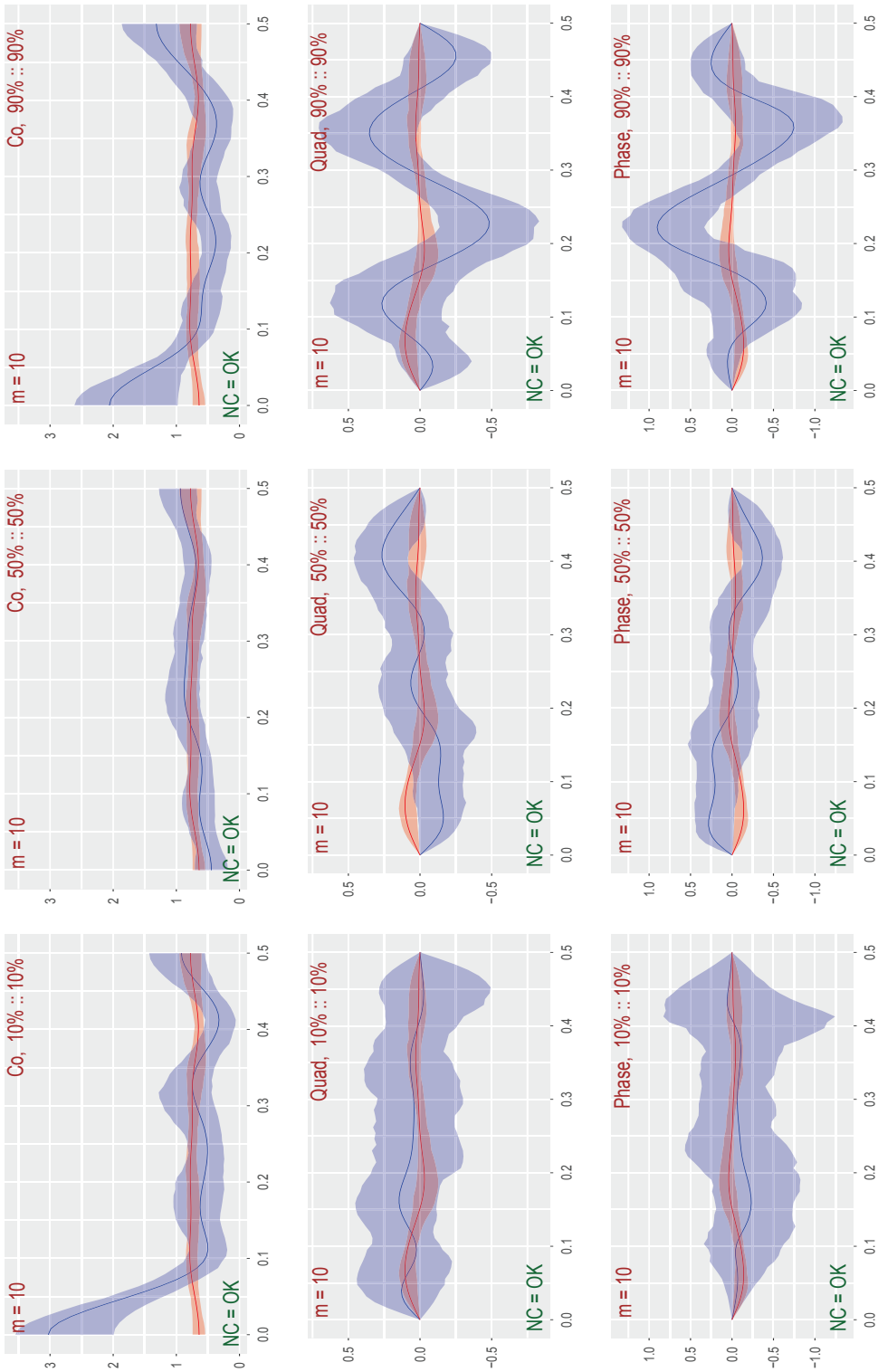

Figure 9: EuStockMarkets, local Gaussian spectra based on $\left(Y_{1,t+h}, Y_{3,t}\right)$, i.e. DAX and CAC.

3.2.2 A simple copula GARCH-model fitted to `EuStockMarkets`

It might not be obvious how to interpret the `Co-`, `Quad-` and `Phase` spectra based on the log-returns of the `EuStockMarkets`-data, but they do at least provide an approach where nonlinear dependencies might be detected from a visual inspection of the plots.

Furthermore, it is possible to use this as an exploratory tool in order to investigate whether a model fitted to the original data is capable of reproducing nonlinear traits that match those observed for the data. The procedure is straightforward:

1. Fit the selected model to the data.
2. Perform a local Gaussian spectrum investigation based on simulated samples from the fitted model. The parameters should match those used in the investigation of the original data.
3. Compare the plots based on the original data with corresponding plots based on the simulated data from the model. It can be of interest to not only compare the `Co-`, `Quad-` and `Phase`-plots, but also include plots that shows the traces and the estimated local Gaussian auto- and cross-spectra.

For the present case of interest, item 2 of the list above implies that 100 independent samples of length 1859 will be used as the basis for the construction of the `Co-`, `Quad-` and `Phase`-plots of the fitted model, and the bandwidth $\boldsymbol{b} = (0.6, 0.6)$ will be used for the estimation of the local Gaussian cross-correlations at the three points `10%::10%`, `50%::50%` and `90%::90%`.

The model: The `R`-package `rmgarch` was used to fit a simple multivariate GARCH-type model to the log-returns of the `EuStockMarkets`-data, in order to exemplify the procedure outlined above, i.e. a copula GARCH-model (cGARCH) with the simplest available univariate models for the marginals[13] was fitted to the data, and the resulting model was then used to produce figs. 10 to 12.

The traces: Figure 10 shows the pseudo-normalised trace of the Y_1- and Y_3-variables for one sample from the tetravariate cGARCH-model, and this can be compared with the corresponding pseudo-normalised trace of the DAX and CAC plot for the pseudo-normalised log-returns of the `EuStockMarkets`-data, see fig. 7 on page 27. Obviously, a comparison of one single simulated trace with the trace of the original data might not reveal much, and it should also be noted that it in general might be preferable to compare the traces before they are subjected to the pseudo-normalisation, since that could detect if the model might fail to produce sufficiently extreme outliers.

The local Gaussian correlations: Box-plots, based on the 100 independent estimates of the local Gaussian cross-correlations from the cGARCH-model, are shown in fig. 11. These can be compared with the local Gaussian cross-correlations estimated from the original sample, shown in fig. 8. It should be noted that the computational cost for the production of the box-plots in fig. 11 is substantially larger than the cost for the production of the simpler plots shown in fig. 8, so it is preferable to restrict the attention to a shorter range of lags in fig. 11. Note also that the wide range of lags included in fig. 8 is related to the need for a justification of the selected block-lenght for the bootstrap-algorithm, and it should be possible to judge the suitability of the fitted model from the shorter range of lags included in fig. 11.

[13] See Ghalanos (2015a) for details about the cGARCH-model and other options available in the `rmgarch`-package.

The impression from the lags included in fig. 11 is that the medians of the estimated local Gaussian cross-correlations for the point `50%::50%` lies quite close to zero, whereas the medians for the points `10%::10%` and `90%::90%` mostly lies slightly above zero. Almost none of the boxes for the two latter points seems to be positioned in a manner consistent with the desired outcome for a good match with the corresponding estimated values in fig. 8, and it might thus be ample reason to suspect that this cGARCH-model might better be replaced with another model instead.

The `Co-` `Quad-` and `Phase`-plots: Figure 12 shows the local Gaussian spectra for the same points $\boldsymbol{v}$ and the same configuration of parameters as those used in fig. 9. The `Co-`, `Quad-` and `Phase`-plots of fig. 12 does look like they could originate from i.i.d. white noise – which does not come as a surprise in view of the information about the local Gaussian cross-correlations in fig. 11. The `Co`-plots for the two points `10%::10%` and `90%::90%` does show that the estimates of the m-truncated local Gaussian cospectra, i.e. the blue dashed lines, might have a peak at $\omega = 0$ – but the 90% pointwise confidence intervals are too wide to support a claim that these peaks are significant. A further comparison of these two `Co`-plots with the corresponding `Co`-plots in fig. 9 (beware of different scales for the axes), does moreover show that the confidence intervals from fig. 12 are too narrow (at $\omega = 0$) to encompass the peaks observed in fig. 9 – which indicates that the selected model might be a rather bad approximation to the log-returns of the `EuStockMarkets`-data. It thus seems advisable to look for some other model instead, a natural conclusion given that no effort whatsoever was made with regard to finding reasonable marginal distributions for the copula GARCH-model used in this discussion.

Remark 3.16. It should be noted that a *local Gaussian spectra comparison* of the original data and the fitted model in practise also should include a comparison of the local Gaussian auto-spectra of the marginals, as was done in (Jordanger and Tjøstheim, 2017). These auto-spectra plots (not included in this paper) can provide some additional information useful for the model-selection process. In particular, if a model-selection algorithm for GARCH-type models has been used to pick one marginal model from a given collection of marginal models, then an investigation based on the *local Gaussian auto-spectrum* might reveal if the selected marginal model captures the local traits of the corresponding marginal observations in a reasonable manner.

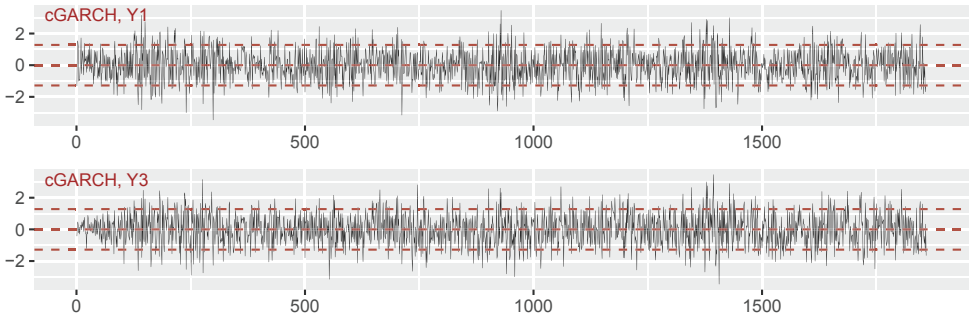

Figure 10: cGARCH, pseduo-normalised components under investigation.

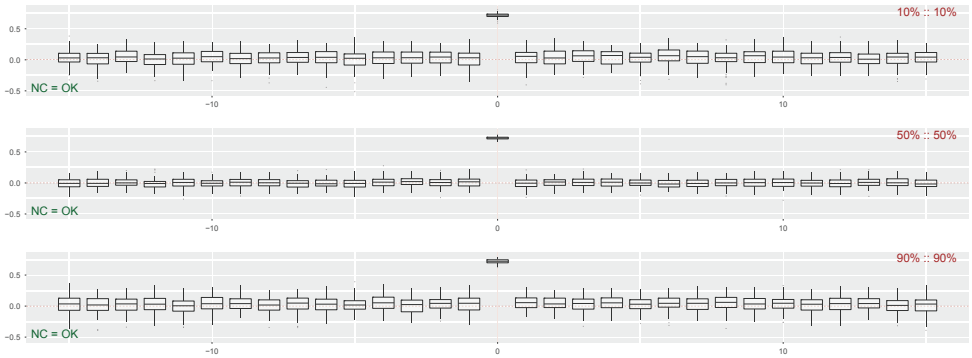

Figure 11: cGARCH, local Gaussian cross-correlations for $\left(Y_{1,t+h}, Y_{3,t}\right)$.

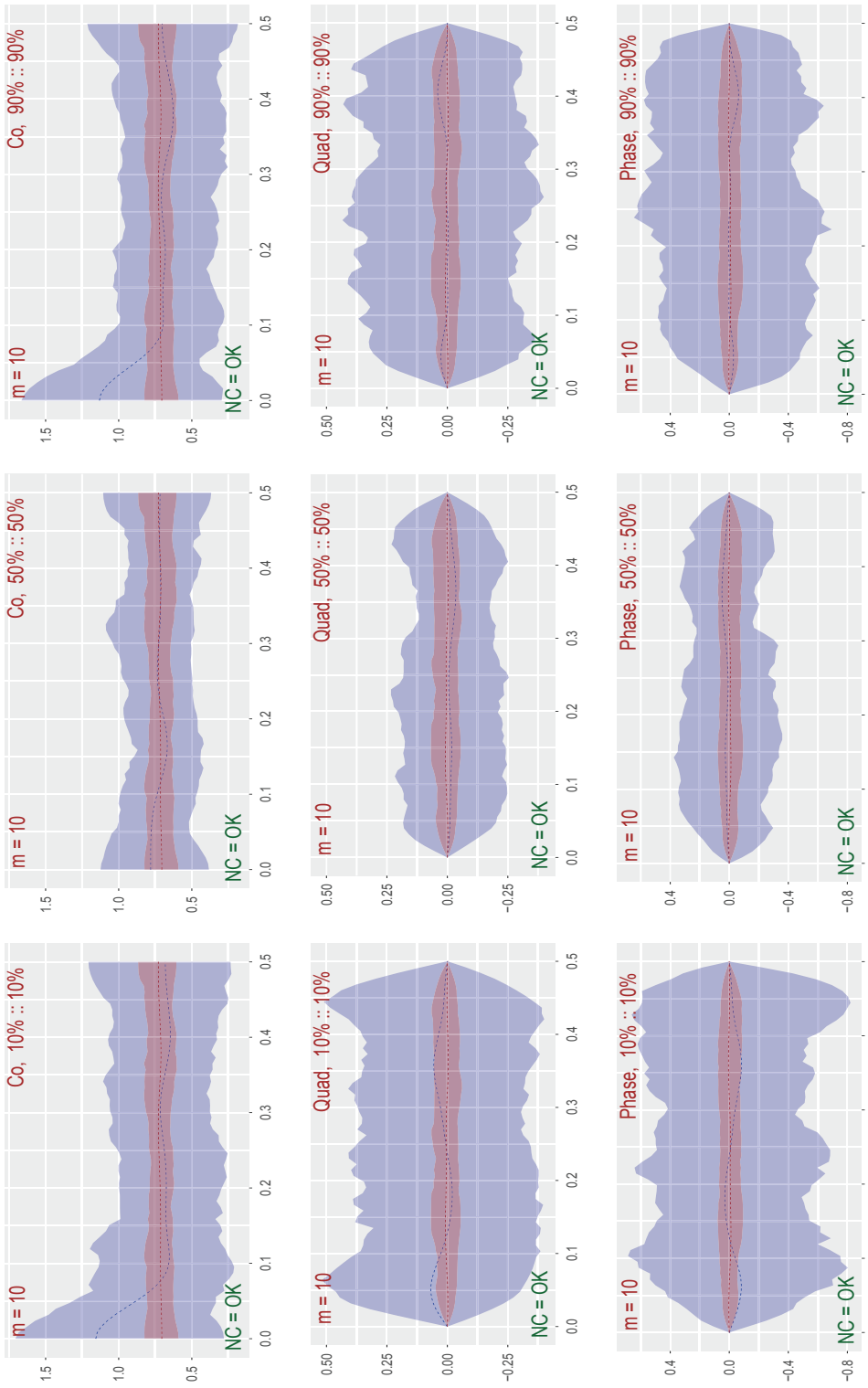

Figure 12: cGARCH, local Gaussian spectra based on $\left(Y_{1,t+h}, Y_{3,t}\right)$, corresponding to DAX and CAC of `EuStockMarkets`.

4 Discussion

This paper extends the *local Gaussian auto-spectrum* from Jordanger and Tjøstheim (2017) to the case of the *local Gaussian cross-spectrum*, and it has been seen that this can be used as an *exploratory tool* to detect non-linear traits in multivariate time series. The simulated *locally trigonometric* examples in section 3.1.2 showed that nonlinear periodicities can be detected with this machinery, but it should be noted that the selection of the point $\boldsymbol{v} = (v_1, v_2)$ was pivotal with regard to what was detected. It was also observed, see the discussion in section 3.2, that this *local Gaussian approach to spectral analysis* might be useful with regard to model selection, i.e. that it can be used to check if a model fitted to the data can reproduce local traits detected in the original sample.

This new local Gaussian approach to spectral analysis is still in its infancy, and many details have not been looked into yet. Some details, i.e. those related to the selection of the input parameters, have already been discussed in the treatment of the univariate case, see (Jordanger and Tjøstheim, 2017, section 4), and these will only be briefly discussed in this section.

4.1 The input parameters

The selection of the point $\boldsymbol{v}$ and the sample length: It is important that the coordinates v_1 and v_2 of the point $\boldsymbol{v} = (v_1, v_2)$ should not correspond to quantiles too far out in the tails, since that could allow small sample variation to dominate even when the sample itself is large,[14] cf. the discussion related to (Jordanger and Tjøstheim, 2017, fig. 2) for further details.

The selection of the bandwidth $\boldsymbol{b}$ and the truncation level m: A problem that can arise when several points $\{\boldsymbol{v}_i\}_{i=1}^r$ are investigated at the same time, is that each point $\boldsymbol{v}_i$ could have its own optimal configuration of bandwidth $\boldsymbol{b}$ and truncation level m – and the question then becomes whether these individual values should be used, or if it would be preferable to instead search for some common values for the bandwidth $\boldsymbol{b}$ and the truncation level m. The interested reader can find a more detailed discussion about this and related issues in (Jordanger and Tjøstheim, 2017, sections 4.1 and 4.2)

The block-length: It was noted in (Jordanger and Tjøstheim, 2017, section 4.3) that the standard algorithms for the selection of the bootstrap block-length depended on the (global) spectral density, and this rendered them rather useless for samples from e.g. GARCH-type models – see (Jordanger and Tjøstheim, 2017) for further details and references.

4.2 The smoothing of the estimates

The estimates of the local Gaussian auto- and cross-spectra are smoothed by the help of weighting functions $\lambda_m(h)$, that work upon the estimated local Gaussian auto- and cross-correlations, cf. definition 2.4(d). For the present investigation, the smoothing function has been inherited from the one used in the global case.

A completely different smoothing strategy can, as explained in (Jordanger and Tjøstheim, 2017, section 4.4), be based on the observation that the estimated m-truncated local Gaussian spectra corresponding to different points $\{\boldsymbol{v}_i\}_{i=1}^r$ will be jointly asymptotically normal and pairwise asymptotically independent (when $m \to \infty$ and $\boldsymbol{b} \to \boldsymbol{0}^+$ as $n \to \infty$). This implies that it should be possible to find a smoothed estimate of the local Gaussian spectral density at a given

[14] It is possible to have a point $\boldsymbol{v}$ far out in the tails if the bandwidth $\boldsymbol{b}$ is large, but the result would then hardly deserve to be referred to as an estimate of the *local* Gaussian spectral density at the point $\boldsymbol{v}$.

point $\boldsymbol{v}$, by the help of the estimated values in a grid of points surrounding $\boldsymbol{v}$. This approach resembles the one used when the ordinary global spectral densities are computed based on the periodogram,[15] but with the important distinction that the smoothing in this case would be over a grid of neighbouring points instead of a range of neighbouring frequencies.

The periodogram-approach for the estimation of the ordinary (global) auto- and cross-spectra provides an efficient estimation regime due to the *Fast Fourier Transform* (FFT), which enables the periodogram to be computed directly from the observations without the need for an explicit computation of all of the estimated auto- and cross-correlations. The effectiveness of FFT is linked to the product structure of the estimated covariance function, i.e. $\widehat{\gamma}(h) = \sum_{t=1}^{n-|h|} y_{t+|h|} \cdot y_t$, and a similar simple connection does not exist for the distribution based approach used for the local Gaussian autocorrelation. The above mentioned grid-based smoothing approach for the local Gaussian spectral density might thus be rather inefficient to use in practice, but it could still be of interest to keep this approach in mind.

4.3 The visualisation of the local Gaussian cross-spectrum

The local Gaussian cross-spectrum is complex valued, which makes a direct graphical investigation of it a bit tricky. It is possible to use a solution like the one in section 3, where a combination of the `Co-`, `Quad-` and `Phase`-plots together provide a graphical setup that enable local traits and periodicities in the data to be discovered, but other alternative might also be of interest to investigate. In particular, plots (like fig. 3) that individually might not be that informative could reveal interesting features if a sequence of them are interactively investigated by the means of a `shiny`-interface, like the one used in the `R`-package `localgaussSpec`.

A completely different approach would be to estimate the local Gaussian cross-spectra for a grid of points $\boldsymbol{v}$ in the plane, and then use three dimensional plots to see how (for a specified frequency ω) the real and imaginary parts, or the amplitude and the phase, varies over the grid. The two dimensional contour-plots corresponding to these three dimensional plots might of course also be considered, or alternative visualisations like the colour-coded phase plots from Wegert and Semmler (2011).

An obvious drawback with the above mentioned approach is that these plots might not be useful unless the number of points in the grid is (sufficiently) large, which could incur large computational costs. It might, despite this, still be of interest to initiate an investigation using such an approach with a coarse grid, since that could detect regions containing promising structures.

4.4 Possible applications

A key property of the local Gaussian spectral densities is that they for *Gaussian time series* coincide with the ordinary spectral densities. This implies that a comparison of ordinary spectra and local Gaussian spectra can detect the presence of non-Gaussian structures, which in particular can be of interest for time series whose ordinary spectra looks like white noise. Any peaks and troughs of the local Gaussian spectra can for such white noise cases naturally be considered as indicators of nonlinear structures in the time series under investigation, and it might then in particular be of interest to investigate if periodicities in the local Gaussian spectra might correspond to real phenomena in the data generating structure.

[15] See e.g. Brockwell and Davis (1986) for details.

Estimates of the local Gaussian spectral densities might also be of interest to consider as an informal tool in addition to existing tools for model selection for time series, i.e. estimates based on *repeated samples from a parametric model fitted to the original observations* can be compared with the corresponding estimates based on *the original observations*. Significant differences between these estimates (taking the pointwise confidence intervals into account) could indicate that the selected model might not be good enough, and that it thus might be prudent to look for a better model.

These applications of the local Gaussian spectra could provide useful insight even when it is unknown whether or not the investigated time series actually satisfies the requirements needed for the asymptotic theory to work. It might however occur problems due to the (at the present state of development) unresolved issues regarding the selection of the input parameters – in particular, cases can be encountered where it is hard to figure out if a peak/trough represents an actual nonlinear phenomenon, or if it is small-sample variation due to an unsavoury combination of input parameters. More work is needed in order to resolve these issues.

5 Conclusion

The *local Gaussian cross-spectrum* $f_{k\ell;\mathfrak{v}|p}(\omega)$ can be used to detect nonlinear dependencies between the marginals $Y_{k,t}$ and $Y_{\ell,t}$ of a multivariate time series $\left\{\boldsymbol{Y}_t = \left(Y_{1,t}, \ldots, Y_{d,t}\right)\right\}_{t\in\mathbb{Z}}$. This provides, together with the *local Gaussian auto-spectrum* from Jordanger and Tjøstheim (2017), a *local Gaussian approach to spectral analysis*, which in particular can be used to detect if the time series under investigation deviates from being Gaussian. For time series whose ordinary auto- and cross-spectra are flat, any peaks and troughs from the *local Gaussian approach* can then be considered indicators of *local nonlinear traits* and *local periodicities*.

Appendix A: The proofs of theorems 2.10 to 2.12

This appendix presents the proofs of the asymptotic results stated in the main part of the paper. The proof of the result for the m-truncated estimate of the local Gaussian cross-spectrum $f_{k\ell:\boldsymbol{v}|p}(\omega)$ is in essence identical to the one encountered in Jordanger and Tjøstheim (2017) for the local Gaussian auto-spectrum $f_{kk:\boldsymbol{v}|p}(\omega)$, whereas the proofs for the estimates of the local Gaussian amplitude- and phase-spectra are identical in structure to those encountered in the ordinary global case. Some technical details needed for the proof of theorem 2.10 are covered in appendix B.

Proof of theorem 2.10, page 11.
The case $\omega \notin \frac{1}{2}\cdot\mathbb{Z}$ will be treated first, since the other case follows from a trivial adjustment of the setup. The key observation for this case is that the sum that defines $\widehat{f}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)$, see eq. (2.6) in definition 2.4(d), implies that $\widehat{c}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)$ and $\widehat{q}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)$ can be realised as the following inner products,

$$\widehat{c}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) = \boldsymbol{\Lambda}'_{c|\overline{\boldsymbol{m}}}(\omega)\cdot\widehat{\boldsymbol{P}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v},\breve{\boldsymbol{v}}) \tag{A.1a}$$

$$\widehat{q}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) = \boldsymbol{\Lambda}'_{q|\overline{\boldsymbol{m}}}(\omega)\cdot\widehat{\boldsymbol{P}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v},\breve{\boldsymbol{v}})\,, \tag{A.1b}$$

where $\boldsymbol{\Lambda}'_{c|\overline{\boldsymbol{m}}}(\omega)$ and $\boldsymbol{\Lambda}'_{q|\overline{\boldsymbol{m}}}(\omega)$ respectively contains the coefficients $\lambda_{m}(h)\cos(2\pi\omega h)$ and $\lambda_{m}(h)\sin(2\pi\omega h)$, and where $\widehat{\boldsymbol{P}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v},\breve{\boldsymbol{v}}) = \big[\widehat{\rho}_{\ell k:\breve{\boldsymbol{v}}|p}(m),\ldots,\widehat{\rho}_{\ell k:\breve{\boldsymbol{v}}|p}(1),\widehat{\rho}_{k\ell:\boldsymbol{v}|p}(0),\ldots,\widehat{\rho}_{k\ell:\boldsymbol{v}|p}(m)\big]'$ contains the $2m+1$ estimates of the local Gaussian cross-correlations (based on the bandwidth $\boldsymbol{b}$). The vector $\widehat{\boldsymbol{P}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v},\breve{\boldsymbol{v}})$ can by the help of a suitable $(2m+1)\times(2m+1)p$ matrix $\boldsymbol{E}'_{\overline{\boldsymbol{m}}|p}$ (based on the vectors $\boldsymbol{e}'_{p}$ that gives $\rho_{k\ell:\boldsymbol{v}|p}(h) = \boldsymbol{e}'_{p}\cdot\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}$) be expressed as

$$\widehat{\boldsymbol{P}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v},\breve{\boldsymbol{v}}) = \boldsymbol{E}'_{\overline{\boldsymbol{m}}|p}\cdot\widehat{\boldsymbol{\Theta}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v},\breve{\boldsymbol{v}})\,, \tag{A.2}$$

where $\widehat{\boldsymbol{\Theta}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v},\breve{\boldsymbol{v}}) = \Big[\widehat{\boldsymbol{\theta}}_{\breve{\boldsymbol{v}}|\ell k:m:n|p},\ldots,\widehat{\boldsymbol{\theta}}_{\breve{\boldsymbol{v}}|\ell k:1:n|p},\widehat{\boldsymbol{\theta}}_{\boldsymbol{v}|k\ell:0:n|p},\ldots,\widehat{\boldsymbol{\theta}}_{\boldsymbol{v}|k\ell:m:n|p}\Big]'$ is the length $(2m+1)p$ vector created by stacking into one vector all the estimated parameters from the local Gaussian approximations. It follows from this that the target of interest can be written as

$$\begin{bmatrix}\widehat{c}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)\\ \widehat{q}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Lambda}'_{c|\overline{\boldsymbol{m}}}(\omega)\\ \boldsymbol{\Lambda}'_{q|\overline{\boldsymbol{m}}}(\omega)\end{bmatrix}\cdot\boldsymbol{E}'_{\overline{\boldsymbol{m}}|p}\cdot\widehat{\boldsymbol{\Theta}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v},\breve{\boldsymbol{v}})\,, \tag{A.3}$$

which together with the asymptotic normality result from theorem B.3 (page 39), i.e.

$$\sqrt{n(b_1b_2)^{(p+1)/2}}\cdot\Big(\widehat{\boldsymbol{\Theta}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v},\breve{\boldsymbol{v}}) - \boldsymbol{\Theta}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v},\breve{\boldsymbol{v}})\Big) \xrightarrow{\;d\;} \mathrm{N}\big(\boldsymbol{0},\Sigma_{\boldsymbol{v}|k\ell:\overline{\boldsymbol{m}}|p}\big)\,, \tag{A.4}$$

and Brockwell and Davis (1986, proposition 6.4.2, p. 211) gives that

$$\sqrt{n(b_1b_2)^{(p+1)/2}/m}\cdot\left(\begin{bmatrix}\widehat{c}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)\\ \widehat{q}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)\end{bmatrix} - \begin{bmatrix}c_{k\ell:\boldsymbol{v}|p}(\omega)\\ q_{k\ell:\boldsymbol{v}|p}(\omega)\end{bmatrix}\right) \tag{A.5}$$

is asymptotically bivariate normally distributed with mean $\mathbf{0}$ and covariance matrix

$$\frac{1}{m} \cdot \left(\begin{bmatrix} \boldsymbol{\Lambda}'_{c|\overline{\underline{m}}}(\omega) \\ \boldsymbol{\Lambda}'_{q|\overline{\underline{m}}}(\omega) \end{bmatrix} \cdot \boldsymbol{E}'_{\overline{\underline{m}}|p} \cdot \Sigma_{\boldsymbol{v}|k\ell:\overline{\underline{m}}|p} \cdot \boldsymbol{E}_{\overline{\underline{m}}|p} \cdot \left[\boldsymbol{\Lambda}_{c|\overline{\underline{m}}}, \boldsymbol{\Lambda}_{q|\overline{\underline{m}}} \right] \right). \tag{A.6}$$

The specified form of the covariance matrix given in theorem 2.10 now follows by the help of some linear algebra, the observation in theorem B.3 that

$$\Sigma_{\boldsymbol{v}|k\ell:\overline{\underline{m}}|p} \coloneqq \left(\bigoplus_{h=m}^{1} \Sigma_{\breve{\boldsymbol{v}}|\ell k:h|p} \right) \bigoplus \left(\bigoplus_{h=0}^{m} \Sigma_{\boldsymbol{v}|k\ell:h|p} \right), \tag{A.7}$$

and the definition $\tilde{\sigma}^2_{\boldsymbol{v}|k\ell|p}(h) \coloneqq \boldsymbol{e}'_p \cdot \Sigma_{\boldsymbol{v}|k\ell:h|p} \cdot \boldsymbol{e}_p$ from theorem B.1.

It is easy to see that both $\sigma^2_{c|k\ell:\boldsymbol{v}|k\ell|p}(\omega)$ and $\sigma^2_{q|k\ell:\boldsymbol{v}|k\ell|p}(\omega)$ from eq. (2.19) are nonzero when $\omega \notin \frac{1}{2} \cdot \mathbb{Z}$, as required for the validity of (Brockwell and Davis, 1986, proposition 6.4.2, p. 211). The proof for the case $\omega \in \frac{1}{2} \cdot \mathbb{Z}$ can be constructed in the same manner, simply ignoring the components having sine-terms. □

The key observation for the proof of theorems 2.11 and 2.12 is that they both follow as a consequence of theorem 2.10 and Brockwell and Davis (1986, proposition 6.4.3, p. 211). Note that these arguments are quite similar to those used for the investigation of the estimates of the ordinary amplitude and phase spectra in (Brockwell and Davis, 1986, p.448-449).

Proof of theorem 2.11, page 12.
First observe that the function $h(x_1, x_2) = \sqrt{x_1^2 + x_2^2}$ implies that

$$\widehat{\alpha}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega) - \alpha_{k\ell:\boldsymbol{v}|p}(\omega) = h\big(\widehat{c}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega), \widehat{q}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)\big) - h\big(c_{k\ell:\boldsymbol{v}|p}(\omega), q_{k\ell:\boldsymbol{v}|p}(\omega)\big), \tag{A.8}$$

and then observe that the asymptotic covariance matrix in theorem 2.10

$$\Sigma_{k\ell:\boldsymbol{v}|p}(\omega) \coloneqq \begin{pmatrix} \sigma^2_{c|k\ell:\boldsymbol{v}|p}(\omega) & 0 \\ 0 & \sigma^2_{q|k\ell:\boldsymbol{v}|p}(\omega) \end{pmatrix}, \tag{A.9}$$

obviously is a symmetric non-negative definite matrix.

It now follows from Brockwell and Davis (1986, proposition 6.4.2, p. 211) that

$$\sqrt{n(b_1 b_2)^{(p+1)/2}/m} \cdot \left\{ h\big(\widehat{c}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega), \widehat{q}^{\,m}_{k\ell:\boldsymbol{v}|p}(\omega)\big) - h\big(c_{k\ell:\boldsymbol{v}|p}(\omega), q_{k\ell:\boldsymbol{v}|p}(\omega)\big) \right\} \xrightarrow{\;d\;} \mathrm{N}(0, \sigma^2_\alpha(\omega)), \tag{A.10}$$

where $\sigma^2_\alpha(\omega) = \boldsymbol{D} \cdot \Sigma_{k\ell:\boldsymbol{v}|p}(\omega) \cdot \boldsymbol{D}'$, with

$$\boldsymbol{D} = \left[\frac{\partial}{\partial x_1} h(x_1, x_2), \frac{\partial}{\partial x_2} h(x_1, x_2) \right] = \left[x_1/\sqrt{x_1^2 + x_2^2}, x_2/\sqrt{x_1^2 + x_2^2} \right] \tag{A.11}$$

evaluated in $(x_1, x_2) = \big(c_{k\ell:\boldsymbol{v}|p}(\omega), q_{k\ell:\boldsymbol{v}|p}(\omega)\big)$.

A simple calculation gives $\boldsymbol{D} = \big[c_{k\ell:\boldsymbol{v}|p}(\omega)/\alpha_{k\ell:\boldsymbol{v}|p}(\omega), q_{k\ell:\boldsymbol{v}|p}(\omega)/\alpha_{k\ell:\boldsymbol{v}|p}(\omega)\big]$, from which it follows that $\sigma^2_\alpha(\omega) = \big(c^2_{k\ell:\boldsymbol{v}|p}(\omega) \cdot \sigma^2_{c|k\ell:\boldsymbol{v}|p}(\omega) + q^2_{k\ell:\boldsymbol{v}|p}(\omega) \cdot \sigma^2_{q|k\ell:\boldsymbol{v}|p}(\omega)\big) / \alpha^2_{k\ell:\boldsymbol{v}|p}(\omega)$. □

Proof of theorem 2.12, page 12.
This argument is quite similar to the proof of theorem 2.11, and only the details that are different will thus be included. In this case the function of interest is $h(x_1, x_2) = \tan^{-1}(x_2/x_1)$, from which it follows that

$$\left[\frac{\partial}{\partial x_1}h(x_1,x_2), \frac{\partial}{\partial x_2}h(x_1,x_2)\right] = \left[-x_2/\left(x_1^2+x_2^2\right), x_1/\left(x_1^2+x_2^2\right)\right]. \tag{A.12}$$

This implies that $\boldsymbol{D} = \left[-q_{k\ell:\boldsymbol{v}|p}(\omega)/\alpha^2_{k\ell:\boldsymbol{v}|p}(\omega), c_{k\ell:\boldsymbol{v}|p}(\omega)/\alpha^2_{k\ell:\boldsymbol{v}|p}(\omega)\right]$ and a simple calculation now gives

$$\sigma^2_\phi(\omega) = \left(q^2_{k\ell:\boldsymbol{v}|p}(\omega)\cdot\sigma^2_{c|k\ell:\boldsymbol{v}|p}(\omega) + c^2_{k\ell:\boldsymbol{v}|p}(\omega)\cdot\sigma^2_{q|k\ell:\boldsymbol{v}|p}(\omega)\right)/\alpha^4_{k\ell:\boldsymbol{v}|p}(\omega), \tag{A.13}$$

which completes the proof. □

Appendix B: The underlying asymptotic results

B.1 The bivariate case, a brief overview and the $\widehat{\rho}_{k\ell:\boldsymbol{v}|p}(h)$-case

The main ingredient for the theoretical setup is a translation of the bivariate results from (Tjøstheim and Hufthammer, 2013) into the multivariate framework, and this is almost identical to the discussion that was given in (Jordanger and Tjøstheim, 2017, appendix B.1.2). The main difference is that two extra indices (k and ℓ) now are needed in order to specify which components from $\boldsymbol{Y}_t = \left(Y_{1,t}, \ldots, Y_{d,t}\right)$ that are investigated.

The basic building-blocks was given in section 2.5.1, see definitions 2.6 to 2.9, and the first target of interest is to define a suitable bivariate penalty function relative to the requirement of eq. (2.10). For a sample of size n from $\{\boldsymbol{Y}_{k\ell:h:t}\}_{t\in\mathbb{Z}}$, and with the present notation, the local penalty function from (Tjøstheim and Hufthammer, 2013) can be described as

$$\begin{aligned} Q_{\boldsymbol{v}|k\ell:h:n|p}\left(\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\right) := &-\sum_{t=1}^{n} K_{k\ell:h:\boldsymbol{b}}(\boldsymbol{Y}_{k\ell:h:t}-\boldsymbol{v})\log\psi_p\left(\boldsymbol{Y}_{k\ell:h:t};\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\right) \\ &+ n\int_{\mathbb{R}^2} K_{k\ell:h:\boldsymbol{b}}(\boldsymbol{y}_{k\ell:h}-\boldsymbol{v})\,\psi_p\left(\boldsymbol{y}_{k\ell:h};\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\right)\mathrm{d}\boldsymbol{y}_{k\ell:h}, \end{aligned} \tag{B.1}$$

and from this, under suitable regularity conditions and by the help of the Klimko-Nelson approach, the following asymptotic normality result can be obtained for the estimated parameters,

$$\sqrt{n(b_1b_2)^{(p+1)/2}}\cdot\left(\widehat{\boldsymbol{\theta}}_{\boldsymbol{v}|k\ell:h:n|p}-\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\right)\xrightarrow{d}\mathrm{N}\left(\boldsymbol{0},\Sigma_{\boldsymbol{v}|k\ell:h|p}\right). \tag{B.2}$$

See Tjøstheim and Hufthammer (2013, Th. 3) for the details for the $p=5$ case, and note that the $p=1$ case follows from a simplification of that result.

Remark B.1. For the $p=1$ case, $\widehat{\boldsymbol{\theta}}_{\boldsymbol{v}|k\ell:h:n|1}$ consists of the single value $\widehat{\rho}_{k\ell:\boldsymbol{v}|1}(h)$, and eq. (B.2) gives the univariate result of interest. For the $p=5$ case, $\widehat{\rho}_{k\ell:\boldsymbol{v}|5}(h)$ can be expressed as $\boldsymbol{e}_5'\cdot\widehat{\boldsymbol{\theta}}_{\boldsymbol{v}|k\ell:h:n|5}$, where $\boldsymbol{e}_5'$ is the unit vector that picks out the correlation from $\widehat{\boldsymbol{\theta}}_{\boldsymbol{v}|k\ell:h:n|5}$, and it follows

from (Brockwell and Davis, 1986, proposition 6.4.2, p. 211) that

$$\sqrt{n(b_1 b_2)^3} \cdot \big(\widehat{\rho}_{k\ell:\boldsymbol{v}|5}(h) - \rho_{k\ell:\boldsymbol{v}|5}(h)\big) \xrightarrow{d} \mathrm{N}\big(0, \boldsymbol{e}_5' \cdot \Sigma_{\boldsymbol{v}|k\ell:h|5} \cdot \boldsymbol{e}_5\big)\,. \tag{B.3}$$

These observations can (by introducing $\boldsymbol{e}_1 \coloneqq 1$) be collected in the following result, which is included in order to give a reference for the statements in theorem 2.10.

Theorem B.1. *Under assumptions 2.1 to 2.3, the following univariate asymptotic result holds for the estimates $\widehat{\rho}_{k\ell:\boldsymbol{v}|p}(h)$ of the local Gaussian cross-correlations $\rho_{k\ell:\boldsymbol{v}|p}(h)$.*

$$\sqrt{n(b_1 b_2)^{(p+1)/2}} \cdot \big(\widehat{\rho}_{k\ell:\boldsymbol{v}|p}(h) - \rho_{k\ell:\boldsymbol{v}|p}(h)\big) \xrightarrow{d} \mathrm{N}\big(0, \tilde{\sigma}^2_{\boldsymbol{v}|k\ell|p}(h)\big)\,, \tag{B.4}$$

where $\tilde{\sigma}^2_{\boldsymbol{v}|k\ell|p}(h) \coloneqq \boldsymbol{e}_p' \cdot \Sigma_{\boldsymbol{v}|k\ell:h|p} \cdot \boldsymbol{e}_p$.

B.2 The asymptotic result for $\widehat{\Theta}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v}, \breve{\boldsymbol{v}})$

The Klimko-Nelson approach from the bivariate case can be extended to the present case of interest in the same manner as it was done for the local Gaussian auto-spectrum in (Jordanger and Tjøstheim, 2017, appendix B).

Definition B.2. *For each bivariate penalty function $Q_{\boldsymbol{v}|k\ell:h:n|p}\big(\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\big)$ (as given in eq. (B.1)), denote by $\widetilde{Q}_{\boldsymbol{v}|k\ell:h:n|p}\big(\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\big)$ the extension of it from a function of $\boldsymbol{Y}_{k\ell:h:t} \coloneqq \big[Y_{k,t+h}, Y_{\ell,t}\big]'$ to a function of $\big[Y_{k,t+m}, \ldots, Y_{k,t}, Y_{\ell,t+m}, \ldots, Y_{\ell,t}\big]'$. Use these extensions do define the new penalty function*

$$Q_{\boldsymbol{v}:\breve{\boldsymbol{v}}|k\ell:\overline{\boldsymbol{m}}:n|p}\big(\boldsymbol{\Theta}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v}, \breve{\boldsymbol{v}})\big) \coloneqq \sum_{h=m}^{1} \widetilde{Q}_{\breve{\boldsymbol{v}}|\ell k:h:n|p}\big(\boldsymbol{\theta}_{\breve{\boldsymbol{v}}|\ell k:h|p}\big) + \sum_{h=0}^{m} \widetilde{Q}_{\boldsymbol{v}|k\ell:h:n|p}\big(\boldsymbol{\theta}_{\boldsymbol{v}|k\ell:h|p}\big)\,. \tag{B.5}$$

The $2m+1$ bivariate components in the sum that defines $Q_{\boldsymbol{v}:\breve{\boldsymbol{v}}|k\ell:\overline{\boldsymbol{m}}:n|p}\Big(\widehat{\boldsymbol{\Theta}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v}, \breve{\boldsymbol{v}})\Big)$ have no common parameters, so the optimisation of the parameters for the different summands can be performed independently. The optimal parameter vector $\widehat{\boldsymbol{\Theta}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v}, \breve{\boldsymbol{v}})$ for the penalty function $Q_{\boldsymbol{v}:\breve{\boldsymbol{v}}|k\ell:\overline{\boldsymbol{m}}:n|p}$ (for a given sample) can thus be constructed by stacking on top of each other the parameter vectors $\widehat{\boldsymbol{\theta}}_{\boldsymbol{v}|k\ell:h|p}$ and $\widehat{\boldsymbol{\theta}}_{\breve{\boldsymbol{v}}|\ell k:h|p}$ that optimise the individual summands in eq. (B.5).

The Klimko-Nelson approach can now be used on the penalty function from eq. (B.5) i.e. four requirements related to the penalty function must be verified before the desired asymptotic result for the paramater-vector $\widehat{\boldsymbol{\Theta}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v}, \breve{\boldsymbol{v}})$ is obtained. The following cross-spectrum analogue of (Jordanger and Tjøstheim, 2017, theorem B.23) can now be stated for the present case of interest.

Theorem B.3. *Under assumptions 2.1 to 2.3, the following asymptotic behaviour holds for the estimated parameters $\widehat{\boldsymbol{\Theta}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v}, \breve{\boldsymbol{v}}) = \Big[\widehat{\boldsymbol{\theta}}_{\breve{\boldsymbol{v}}|\ell k:m:n|p}, \ldots, \widehat{\boldsymbol{\theta}}_{\breve{\boldsymbol{v}}|\ell k:1:n|p}, \widehat{\boldsymbol{\theta}}_{\boldsymbol{v}|k\ell:0:n|p}, \ldots, \widehat{\boldsymbol{\theta}}_{\boldsymbol{v}|k\ell:m:n|p}\Big]'$,*

$$\sqrt{n(b_1 b_2)^{(p+1)/2}} \cdot \Big(\widehat{\boldsymbol{\Theta}}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v}, \breve{\boldsymbol{v}}) - \boldsymbol{\Theta}_{k\ell:\overline{\boldsymbol{m}}|\boldsymbol{b}|p}(\boldsymbol{v}, \breve{\boldsymbol{v}})\Big) \xrightarrow{d} \mathrm{N}\big(\boldsymbol{0}, \Sigma_{\boldsymbol{v}|k\ell:\overline{\boldsymbol{m}}|p}\big)\,, \tag{B.6}$$

where the matrix $\Sigma_{\boldsymbol{v}|k\ell:\overline{m}|p}$ *is the direct sum of the matrices from eq.* (B.2) *that occurs when the individual bivariate components of the penalty function is investigated, i.e.*

$$\Sigma_{\boldsymbol{v}|k\ell:\overline{m}|p} \coloneqq \left(\bigoplus_{h=m}^{1} \Sigma_{\breve{\boldsymbol{v}}|\ell k:h|p}\right) \bigoplus \left(\bigoplus_{h=0}^{m} \Sigma_{\boldsymbol{v}|k\ell:h|p}\right). \tag{B.7}$$

Proof. This result follows when the Klimko-Nelson approach is used with the local penalty-function $Q_{\boldsymbol{v}:\breve{\boldsymbol{v}}|k\ell:\overline{m}:n|p}\big(\boldsymbol{\Theta}_{k\ell:\overline{m}|\boldsymbol{b}|p}(\boldsymbol{v},\breve{\boldsymbol{v}})\big)$ from eq. (B.5), and the proof is in essence identical to the proof of (Jordanger and Tjøstheim, 2017, theorem B.23). The three first requirements of the Klimko-Nelson approach follows trivially from the corresponding investigation for the bivariate case, whereas the proof of the fourth requirement must take into account how $m \to \infty$ and $\boldsymbol{b} \to \mathbf{0}^{+}$ as $n \to \infty$.

The investigation of the fourth requirement of the Klimko-Nelson approach can be done in the exact same manner that was employed in Jordanger and Tjøstheim (2017), i.e. first construct a collection of simple random variables whose interaction and asymptotic properties are easy to investigate, then use these basic building blocks to construct a more complicated random variable $\mathfrak{Q}^{n}_{\boldsymbol{v}|\overline{m}}$ that has the same limiting distribution as the estimator of $\sqrt{b_1 b_2}\boldsymbol{\nabla}_{k\ell:\overline{m}|p} Q_{\boldsymbol{v}:\breve{\boldsymbol{v}}|k\ell:\overline{m}:n|p}\big(\boldsymbol{\Theta}_{k\ell:\overline{m}|\boldsymbol{b}|p}(\boldsymbol{v},\breve{\boldsymbol{v}})\big)$ (where $\boldsymbol{\nabla}_{k\ell:\overline{m}|p}$ is obtained by stacking together $\boldsymbol{\nabla}_{k\ell:h|p}$). After this, it is sufficient to use standard methods to prove that the limiting distribution of $\mathfrak{Q}^{n}_{\boldsymbol{v}|\overline{m}}$ is the desired multivariate normal distribution, and the statement for the parameter vectors then follows from the Klimko-Nelson theorem and some linear algebra. □

References

Berentsen, G. D., Kleppe, T. S., Tjøstheim, D. B., Feb. 2014. Introducing `localgauss`, an R Package for Estimating and Visualizing Local Gaussian Correlation. j-J-STAT-SOFT 56 (12).
URL http://www.jstatsoft.org/v56/i12

Brillinger, D. R., 1965. An Introduction to Polyspectra. The Annals of Mathematical Statistics 36 (5), 1351–1374.
URL http://www.jstor.org/stable/2238424

Brillinger, D. R. (Ed.), 1984. The collected works of John W. Tukey. Volume I. Time series: 1949–1964. Wadsworth Statistics/Probability Series. Wadsworth, Pacific Grove, CA, USA, with introductory material by William S. Cleveland and Frederick Mosteller.

Brillinger, D. R., 1991. Some history of the study of higher-order moments and spectra. Statistica Sinica 1 (465-476), 24J.
URL http://www3.stat.sinica.edu.tw/statistica/j1n2/j1n23/..\j1n210\j1n210.htm

Brockwell, P. J., Davis, R. A., 1986. Time Series: Theory and Methods. Springer-Verlag New York, Inc., New York, NY, USA.

Chang, W., Cheng, J., Allaire, J., Xie, Y., McPherson, J., 2017. shiny: Web Application Framework for R. R package version 1.0.3.
URL https://CRAN.R-project.org/package=shiny

Chung, J., Hong, Y., 2007. Model-free evaluation of directional predictability in foreign exchange markets. Journal of Applied Econometrics 22 (5), 855–889.
URL http://dx.doi.org/10.1002/jae.965

Ghalanos, A., 2015a. The rmgarch models: Background and properties. (Version 1.3-0).
URL https://CRAN.R-project.org/web/packages/rmgarch/vignettes/The_rmgarch_models.pdf

Ghalanos, A., 2015b. rmgarch: Multivariate GARCH models. R package version 1.3-0.
URL https://CRAN.R-project.org/package=rmgarch

Hjort, N. L., Jones, M. C., 08 1996. Locally parametric nonparametric density estimation. Ann. Statist. 24 (4), 1619–1647.
URL http://dx.doi.org/10.1214/aos/1032298288

Hong, Y., 1999. Hypothesis Testing in Time Series via the Empirical Characteristic Function: A Generalized Spectral Density Approach. Journal of the American Statistical Association 94 (448), 1201–1220.
URL http://tandfonline.com/doi/abs/10.1080/01621459.1999.10473874

Hong, Y., Tu, J., Zhou, G., 2007. Asymmetries in Stock Returns: Statistical Tests and Economic Evaluation. The Review of Financial Studies 20 (5), 1547–1581.
URL http://www.jstor.org/stable/4494812

Jordanger, L. A., Tjøstheim, D., 2017. Nonlinear spectral analsys via the local Gaussian correlation.
URL http://arxiv.org/abs/1708.02166v1

Klimko, L. A., Nelson, P. I., 05 1978. On Conditional Least Squares Estimation for Stochastic Processes. Ann. Statist. 6 (3), 629–642.
URL http://dx.doi.org/10.1214/aos/1176344207

Otneim, H., Tjøstheim, D., Oct 2016. The locally Gaussian density estimator for multivariate data. Statistics and Computing, 1–22.
URL https://doi.org/10.1007/s11222-016-9706-6

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL https://www.R-project.org/

Teräsvirta, T., Tjøstheim, D., Granger, C. W., et al., 2010. Modelling nonlinear economic time series. OUP Catalogue.

Tjøstheim, D., Hufthammer, K. O., 2013. Local Gaussian correlation: A new measure of dependence. Journal of Econometrics 172 (1), 33 – 48.
URL http://www.sciencedirect.com/science/article/pii/S0304407612001741

Tukey, J. W., 1959. An introduction to the measurement of spectra. In: Grenander, U. (Ed.), Probability and Statistics, The Harald Cramér Volume. Almqvist and Wiksell, Stockholm, Sweden, pp. 300–330.

Wegert, E., Semmler, G., June/July 2011. Phase Plots of Complex Functions: A Journey in Illustration. Notices of the AMS 58 (6), 768–780.
URL http://www.ams.org/journals/notices/201106/rtx110600768p.pdf

Xie, Y., 2015. Dynamic Documents with R and knitr, 2nd Edition. Chapman and Hall/CRC, Boca Raton, Florida, iSBN 978-1498716963.
URL http://yihui.name/knitr/

Xie, Y., 2016. knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.15.1.
URL http://yihui.name/knitr/