econometrics

MDPI

Article

Supplementary Material: Local Gaussian cross-spectrum analysis

Lars Arne Jordanger [1,†,*,‡] **and Dag Tjøstheim** [2,‡]

[1] Department of Computer science, Electrical engineering and Mathematical sciences, Faculty of Engineering and Science, Western Norway University of Applied Sciences, P.B 7030, 5020 Bergen, Norway; lars.arne.jordanger@hvl.no
[2] Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Bergen, P.B. 7803, 5020 Bergen, Norway; dag.tjostheim@uib.no
* Correspondence: lars.arne.jordanger@hvl.no
† Current address: Department of Computer science, Electrical engineering and Mathematical sciences, Faculty of Engineering and Science, Western Norway University of Applied Sciences, P.B 7030, 5020 Bergen, Norway.
‡ These authors contributed equally to this work.

Abstract: This part contains the Supplementary Material for *Local Gaussian cross-spectrum analysis*.

Keywords: Local periodicities; local cospectrum; local quadrature spectrum; local amplitude spectrum; local phase spectrum; heatmap; distance plot; spectral plots

Citation: Jordanger, Lars Arne and Dag Tjøstheim. 2023. Supplementary Material: Local Gaussian cross-spectrum analysis. *Econometrics* 1: 0. https://doi.org/

Academic Editor: Mirjana Rupic

Received:
Accepted:
Published:

SUPPLEMENTARY MATERIAL

This is the supplementary material to the paper *Local Gaussian cross-spectrum analysis*, which extends the univariate theory developed in Jordanger and Tjøstheim (2022) (hereafter referred to as JT22) to the multivariate case.

Section S1 contains an extended version of the asymptotic theory for $\hat{f}^m_{k\ell;\boldsymbol{v}}(\omega)$ (which was only briefly sketched in Section 2.5). The underlying asymptotic theory for the estimates of the local Gaussian parameters are given in Section S2. Note that the multivariate local Gaussian spectral theory is quite similar to the theory developed for the univariate case, and the discussion in Section S2 thus frequently refers to the theory developed in the Supplementary Material of JT22 and it primarily focuses on the adjustments needed to cope with the present multivariate case.

The sensitivity of the tuning parameters (and the effect of varying the point $\boldsymbol{v}$) are discussed in Section S3, whereas Section S4 briefly comments upon the task of selecting these parameters. Note that this discussion in essence is the same as the one given for the univariate case, see JT22 [Appendices D and E], but it has been included here for the convenience of the reader. The scripts needed for the reproduction of these multivariate examples are included in the `R`-package `localgaussSpec`, and an investigator can use these scripts as templates for similar investigations.

Section S5 is devoted to sampling and resampling strategies. The first part of Section S5 applies the univariate local Gaussian sanity testing of parametric models from JT22 to the DAX- and CAC-margins of the `EuStockMarkets` example, and after this new plots are presented that can be used for the sanity testing of the cross-interaction of the multivariate parametric models fitted to the `EuStockMarkets` data. The second part of Section S5 explains the motivation for the *Circular index-based block bootstrap for tuples* resampling strategy that was developed in JT22, and this part also contains a sensitivity analysis of the block length L, which reveals that this tuning parameter plays a minor role for the time series that are long enough for a local Gaussian spectral investigation to be of interest.

All the examples/investigations in this paper can be reproduced by the scripts contained in the `R`-package `localgaussSpec`, and the interested reader will find some practical

information related to that in Section S6. Some further discussions of the examples in Section 3 have also been collected in Section S6 in order to improve the flow of the main part. The interested reader will here find the details related to the bivariate Gaussian test-example, a short explanation of how animations of complex-valued plots can be used to investigate the estimated local Gaussian cross-spectra, and some additional details related to the construction of the local trigonometric examples used for the sanity-testing of the estimation algorithm. Some details are also included at the end related to limitations of the local Gaussian approach for points $\boldsymbol{v}$ that lies in the extreme tails of the observations.

S1. The asymptotic theory for $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$

This part will cover in detail the asymptotic theory for $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$. Section S1.1 gives an extended version of the discussion in Section 2.5 from the main part, and the basic discussion and definitions will be repeated here (in order for this part to be self contained).

The three asymptotic results in Theorems 1 to 3 will for the convenience of the reader also be included here, see Section S1.1.2, whereas the proofs are gathered in Section S1.3. Some technical details needed for the proof of Theorem 1 are covered in Section S2.

S1.1. Asymptotic theory for $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ — all definitions and assumptions

The key idea is that the asymptotic theory for the local Gaussian cross-spectrum $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ follows from a few minor adjustments of the asymptotic theory that was developed in JT22 [Section 2.4] for the local Gaussian auto-spectra. 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 the discussion after Definition 3 (in the main part) for further details.

S1.1.1. Some definitions and an assumption for $\boldsymbol{Y}_t = \left(Y_{1,t}, \ldots, Y_{d,t}\right)$

As for the univariate case in JT22, 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 $\left\{Y_{k,t}\right\}_{t\in\mathbb{Z}}$ and $\left\{Y_{\ell,t}\right\}_{t\in\mathbb{Z}}$. Note that the *folding property* from Item 4 of Algorithm 1 (in the main part) implies that it is sufficient to formulate the assumption based on non-negative values of the lag h. The definition below, repeats Definition 4.

Definition S1. *For a strictly stationary multivariate time series $\{\boldsymbol{Y}_t\}_{t\in\mathbb{Z}}$, with $\boldsymbol{Y}_t = \left(Y_{1,t}, \ldots, Y_{d,t}\right)$, and for a selected pair of indices k and ℓ, define the following bivariate pairs from the univariate marginals $\left\{Y_{k,t}\right\}_{t\in\mathbb{Z}}$ and $\left\{Y_{\ell,t}\right\}_{t\in\mathbb{Z}}$.*

$$\boldsymbol{Y}_{k\ell:h:t} := \left[Y_{k,t+h}, Y_{\ell,t}\right]', \qquad h \geq 0, \tag{S1a}$$

$$\boldsymbol{Y}_{\ell k:h:t} := \left[Y_{\ell,t+h}, Y_{k,t}\right]', \qquad h \geq 1, \tag{S1b}$$

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 key idea for the construction of $f_{k\ell:\boldsymbol{v}}(\omega)$ 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 $\psi(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{k\ell:v:h})$, where ψ is the bivariate Gaussian density function from Equation (4) in the main part. The ψ-function is repeated below for the convenience of the reader,

$$\psi(\boldsymbol{w}; \boldsymbol{\theta}) := \frac{1}{2\pi \cdot \sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{\sigma_2^2(w_1-\mu_1)^2 - 2\sigma_1\sigma_2\rho(w_1-\mu_1)(w_2-\mu_2) + \sigma_1^2(w_2-\mu_2)^2}{\sigma_1^2\sigma_2^2(1-\rho^2)}\right\}.$$

The target of interest is to find a parameter-vector $\boldsymbol{\theta}_{\boldsymbol{v}} = \left[\mu_{1v}, \mu_{2v}, \sigma_{1v}, \sigma_{2v}, \rho_v\right]'$ that gives the best match to a density function $g(\boldsymbol{w})$ in a neighbourhood of $\boldsymbol{v}$, and then extract the correlation component ρ_v from the parameter-vector $\boldsymbol{\theta}_{\boldsymbol{v}}$. The added indices in this section is necessary in order to keep track of the density functions related to $Y_{k,t+h}$ and $Y_{\ell,t}$.

Note that the estimations to be performed for 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 $K_{k\ell:h:\boldsymbol{b}}(\boldsymbol{y}_{k\ell:h} - \boldsymbol{v}) := \frac{1}{b_1 b_2} K\Big(\frac{y_{k,h}-v_1}{b_1}, \frac{y_{\ell,0}-v_2}{b_2}\Big)$, which in turn is used in

$$q_{k\ell:\boldsymbol{v}:h:\boldsymbol{b}} := \int_{\mathbb{R}^2} K_{k\ell:h:\boldsymbol{b}}(\boldsymbol{y}_{k\ell:h} - \boldsymbol{v})[\psi(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}) - \log \psi(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}) g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})] \, \mathrm{d}\boldsymbol{y}_{k\ell:h}, \tag{S2}$$

a minimiser of which should satisfy the vector equation

$$\int_{\mathbb{R}^2} K_{k\ell:h:\boldsymbol{b}}(\boldsymbol{y}_{k\ell:h} - \boldsymbol{v}) \boldsymbol{u}_{k\ell:h}(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h})[g_{k\ell:h}(\boldsymbol{y}_{k\ell:h}) - \psi(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h})] \, \mathrm{d}\boldsymbol{y}_{k\ell:h} = \boldsymbol{0}, \tag{S3}$$

where $\boldsymbol{u}_{k\ell:h}(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}) := \boldsymbol{\nabla}_{k\ell:h} \log \psi(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h})$ is the score function of $\psi(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h})$ (with $\boldsymbol{\nabla}_{k\ell:h} := \partial/\partial\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}$). Under the assumption that there is a bandwidth $\boldsymbol{b}_{k\ell:h:0}$ such that there exists a minimiser $\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h:\boldsymbol{b}}$ of Equation (S2) which satisfies Equation (S3) for any $\boldsymbol{b}$ with $\boldsymbol{0} < \boldsymbol{b} < \boldsymbol{b}_{k\ell:h:0}$, this $\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h:\boldsymbol{b}}$ will be referred to as the population value for the given bandwidth $\boldsymbol{b}$.

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 $\hat{\boldsymbol{\theta}}_{k\ell:\boldsymbol{v}:h}$ (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[1] and a suitably defined local penalty function $Q_{k\ell:\boldsymbol{v}:h:n}(\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h})$ (see Equation (S31) in Section S2.1).

The assumptions to be imposed on $\boldsymbol{Y}_t$ is related to the estimation of Equation (S31), and thus requires a few additional definitions.

Definition S2. *For $\psi(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta})$ 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 $\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h:\boldsymbol{b}}$ the population value that minimises the penalty function $q_{k\ell:\boldsymbol{v}:h:\boldsymbol{b}}$ from Equation (S2), define for all $h \in \mathbb{N}$ and all $q \in \{1, \ldots, 5\}$*

$$u_{k\ell:h:q:\boldsymbol{b}}(\boldsymbol{w}) := \frac{\partial}{\partial\theta_q} \log(\psi(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}))\Big|_{(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}) = (\boldsymbol{w}; \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h:\boldsymbol{b}})}, \tag{S4}$$

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

The following requirements on the kernel function are identical to those given in JT22 [Definition B.9].

[1] 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 JT22 [Appendix B.1] for a more detailed presentation of the Klimko-Nelson approach when a local penalty-function is used.

Definition S3. *From a bivariate, non-negative, and bounded kernel function $K(\boldsymbol{w})$, that satisfies*

$$\int_{\mathbb{R}^2} K(w_1, w_2)\,\mathrm{d}w_1\mathrm{d}w_2 = 1, \tag{S5a}$$

$$\mathcal{K}_{1:k}(w_2) := \int_{\mathbb{R}^1} K(w_1, w_2) w_1^k\,\mathrm{d}w_1 \qquad \text{is bounded for } k \in \{0,1,2\}, \tag{S5b}$$

$$\mathcal{K}_{2:\ell}(w_1) := \int_{\mathbb{R}^1} K(w_1, w_2) w_2^\ell\,\mathrm{d}w_2 \qquad \text{is bounded for } \ell \in \{0,1,2\}, \tag{S5c}$$

$$\int_{\mathbb{R}^2} K(w_1, w_2)\left|w_1^k w_2^\ell\right|\,\mathrm{d}w_1\mathrm{d}w_2 < \infty, \qquad k,\ell \geq 0 \text{ and } k+\ell \leq 2\cdot\lceil\nu\rceil, \tag{S5d}$$

where $\nu > 2$ is from Assumption S1(2) (and $\lceil\cdot\rceil$ is the ceiling function), define

$$K_{h:\boldsymbol{b}}(\boldsymbol{y}_{k\ell:h} - \boldsymbol{v}) := \frac{1}{b_1 b_2} K\left(\frac{y_h - v_1}{b_1}, \frac{y_0 - v_2}{b_2}\right). \tag{S6}$$

Definition S4. *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,*

$$X^{n|\boldsymbol{v}}_{k\ell:h:q:t} := \sqrt{b_1 b_2} K_{h:\boldsymbol{b}}(\boldsymbol{Y}_{k\ell:h:t} - \boldsymbol{v}) u_{k\ell:h:q:\boldsymbol{b}}(\boldsymbol{Y}_{k\ell:h:t}). \tag{S7}$$

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 S1. *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 4 below) the point at which $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$, the estimate of $f_{k\ell:\boldsymbol{v}}(\omega)$, is to be computed.*

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

$$\sum_{j=1}^{\infty} j^a [\alpha(j)]^{1-2/\nu} < \infty \qquad \text{for some } \nu > 2 \text{ and } a > 1 - 2/\nu. \tag{S8}$$

3. *$\mathrm{E}\left[||\boldsymbol{Y}_t||^2\right] < \infty$, 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 Equation (S1), must satisfy the following requirements for a given point $\boldsymbol{v} = (v_1, v_2)$.

4. *$g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})$ 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*

$$g_{k\ell:h}(\boldsymbol{y}_{k\ell:h}) = g_h(\boldsymbol{v}) + \mathfrak{g}_h(\boldsymbol{v})'[\boldsymbol{y}_{k\ell:h} - \boldsymbol{v}] + \mathfrak{R}_h(\boldsymbol{y}_{k\ell:h})'[\boldsymbol{y}_{k\ell:h} - \boldsymbol{v}], \tag{S9}$$
$$\text{where } \mathfrak{g}_h(\boldsymbol{v}) = \left[\frac{\partial}{\partial y_h} g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})\Big|_{\boldsymbol{y}_{k\ell:h}=\boldsymbol{v}}, \frac{\partial}{\partial y_0} g_{k\ell:h}(\boldsymbol{y}_{k\ell:h})\Big|_{\boldsymbol{y}_{k\ell:h}=\boldsymbol{v}}\right]'$$
$$\text{and } \lim_{\boldsymbol{y}_{k\ell:h}\longrightarrow\boldsymbol{v}} \mathfrak{R}_h(\boldsymbol{y}_{k\ell:h}) = 0,$$

   with the same requirement for $g_{\ell k:h}(\boldsymbol{y}_{\ell k:h})$ at the diagonally reflected point $\breve{\boldsymbol{v}} = (v_2, v_1)$.
5. *There exists a bandwidth $\boldsymbol{b}_{k\ell:h:0}$ such that there for every $\boldsymbol{0} < \boldsymbol{b} < \boldsymbol{b}_{k\ell:h:0}$ is a unique minimiser $\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h:\boldsymbol{b}}$ of the penalty function $q_{k\ell:\boldsymbol{v}:h:\boldsymbol{b}}$ from Equation (S2).*

6. *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*[2]

$$\mathbf{0} < \boldsymbol{b}_{k\ell:0} \coloneqq \inf_{h\in\mathbb{Z}} \boldsymbol{b}_{k\ell:h:0}, \tag{S10}$$

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

7. *For* $X^{n|\boldsymbol{v}}_{k\ell:h:q:t}$ *from Definition S4, the related bivariate, trivariate and tetravariate density functions must be such that the expectations* $\mathrm{E}\Big[X^{n|\boldsymbol{v}}_{k\ell:h:q:t}\Big]$, $\mathrm{E}\Big[\Big|X^{n|\boldsymbol{v}}_{k\ell:h:q:t}\Big|^{\nu}\Big]$ *and* $\mathrm{E}\Big[X^{n|\boldsymbol{v}}_{k\ell:h:q:t}\cdot X^{n|\boldsymbol{v}}_{k\ell:i:r:s}\Big]$ *all are finite.*

The present Assumption S1 is in essence identical to JT22 [Assumption 2.1] with some extra indices, so the remarks from JT22 is of intereset here too. In particular, the α-mixing requirement in Item 2 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 6, 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 7 are always trivially satisfied if the required density-functions are finite.

S1.1.2. An assumption for $\left(Y_{k,t}, Y_{\ell,t}\right)$ and the score function $\boldsymbol{u}(\boldsymbol{w};\boldsymbol{\theta})$ of $\psi(\boldsymbol{w};\boldsymbol{\theta})$

The following assumption is in essence identical to JT22 [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 S2. *The collection of local Gaussian parameters* $\{\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}\}$ *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*

1. $\boldsymbol{u}(\boldsymbol{v};\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}) \neq \mathbf{0}$ *for all finite* h.
2. $\lim_{h\to\infty} \boldsymbol{u}(\boldsymbol{v};\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}) \neq \mathbf{0}$.

Note that an inspection of the 5 equations in $\boldsymbol{u}(\boldsymbol{w};\boldsymbol{\theta}) = \mathbf{0}$ can be used to identify when Items 1 and 2 of Assumption S2 might fail, cf. the discussion in JT22 [Section 2.4.2] for further details.

S1.1.3. Assumptions for n, m and $\boldsymbol{b} = (b_1, b_2)$

The following assumption is identical to JT22 [Assumption 2.3]. The internal consistency of it was verified in JT22 [Lemma C.3].

Assumption S3. *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* $h = 1, \ldots, m$ *(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 S1(2). 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.*[3]

1. $\log n / n(b_1b_2)^5 \longrightarrow 0$.
2. $nb_1b_2/m \longrightarrow \infty$.
3. $m^{\delta}(b_1 \vee b_2) \longrightarrow 0$, *where* $\delta = 2 \vee \frac{\nu(a+1)}{\nu(a-1)-2}$.
4. $\sqrt{nm/b_1b_2}\cdot s^{\tau}\cdot\alpha(s-m+1) \longrightarrow \infty$.
5. $m = o\Big((nb_1b_2)^{\tau/(2+5\tau)-\lambda}\Big)$, *for some* $\lambda \in (0, \tau/(2+5\tau))$.

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

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

6. $m = o(s)$.

S1.2. Convergence theorems for $\hat{f}^{m}_{k\ell:\boldsymbol{v}}(\omega)$, $\hat{\alpha}^{m}_{k\ell:\boldsymbol{v}}(\omega)$ and $\hat{\phi}^{m}_{k\ell:\boldsymbol{v}}(\omega)$

This section repeats the asymptotic results given in Section 2.5.2. The proofs are given in Section S1.3.

Theorem S1 (Theorem 1 in the main part)**.** *The estimate $\hat{f}^{m}_{k\ell:\boldsymbol{v}}(\omega) = \hat{c}^{m}_{k\ell:\boldsymbol{v}}(\omega) - i \cdot \hat{q}^{m}_{k\ell:\boldsymbol{v}}(\omega)$ of the local Gaussian cross-spectrum $f_{k\ell:\boldsymbol{v}}(\omega) = c_{k\ell:\boldsymbol{v}}(\omega) - i \cdot q_{k\ell:\boldsymbol{v}}(\omega)$, will under Assumptions S1 to S3 satisfy*

$$\sqrt{n(b_1 b_2)^3/m} \cdot \left(\begin{bmatrix} \hat{c}^{m}_{k\ell:\boldsymbol{v}}(\omega) \\ \hat{q}^{m}_{k\ell:\boldsymbol{v}}(\omega) \end{bmatrix} - \begin{bmatrix} c_{k\ell:\boldsymbol{v}}(\omega) \\ q_{k\ell:\boldsymbol{v}}(\omega) \end{bmatrix} \right) \xrightarrow{d} \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{c|k\ell:\boldsymbol{v}}(\omega) & 0 \\ 0 & \sigma^2_{q|k\ell:\boldsymbol{v}}(\omega) \end{bmatrix} \right), \tag{S11}$$

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

$$\sigma^2_{c|k\ell:\boldsymbol{v}}(\omega) = \lim_{m\to\infty} \frac{1}{m} \left(\tilde{\sigma}^2_{k\ell:\boldsymbol{v}}(0) + \sum_{h=1}^{m} \lambda^2_m(h) \cdot \cos^2(2\pi\omega h) \cdot \left\{ \tilde{\sigma}^2_{k\ell:\boldsymbol{v}}(h) + \tilde{\sigma}^2_{\ell k:\breve{\boldsymbol{v}}}(h) \right\} \right) \tag{S12a}$$

$$\sigma^2_{q|k\ell:\boldsymbol{v}}(\omega) = \lim_{m\to\infty} \frac{1}{m} \left(\sum_{h=1}^{m} \lambda^2_m(h) \cdot \sin^2(2\pi\omega h) \cdot \left\{ \tilde{\sigma}^2_{k\ell:\boldsymbol{v}}(h) + \tilde{\sigma}^2_{\ell k:\breve{\boldsymbol{v}}}(h) \right\} \right), \tag{S12b}$$

with $\tilde{\sigma}^2_{k\ell:\boldsymbol{v}}(h)$ and $\tilde{\sigma}^2_{\ell k:\breve{\boldsymbol{v}}}(h)$ the asymptotic variances related to the estimates $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ and $\hat{\rho}_{\ell k:\breve{\boldsymbol{v}}}(h)$, see Theorem S4 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

$$\sqrt{n(b_1 b_2)^3/m} \cdot \left(\hat{f}^{m}_{k\ell:\boldsymbol{v}}(\omega) - f_{k\ell:\boldsymbol{v}}(\omega) \right) \xrightarrow{d} \mathrm{N}\left(0, \sigma^2_{c|k\ell:\boldsymbol{v}}(\omega)\right), \qquad \omega \in \tfrac{1}{2} \cdot \mathbb{Z}. \tag{S13}$$

The asymptotic results for the local Gaussian amplitude- and phase-spectra are a direct consequence of Theorem S1 and Brockwell and Davis (1986, proposition 6.4.3, p. 211).

Theorem S2 (Theorem 2 in the main part)**.** *Under Assumptions S1 to S3, when $\alpha_{k\ell:\boldsymbol{v}}(\omega) > 0$ and $\omega \notin \frac{1}{2} \cdot \mathbb{Z}$, the estimate $\hat{\alpha}^{m}_{k\ell:\boldsymbol{v}}(\omega) = \sqrt{\left(\hat{c}^{m}_{k\ell:\boldsymbol{v}}(\omega)\right)^2 + \left(\hat{q}^{m}_{k\ell:\boldsymbol{v}}(\omega)\right)^2}$ satisfies*

$$\sqrt{n(b_1 b_2)^3/m} \cdot \left(\hat{\alpha}^{m}_{k\ell:\boldsymbol{v}}(\omega) - \alpha_{k\ell:\boldsymbol{v}}(\omega) \right) \xrightarrow{d} \mathrm{N}\left(0, \sigma^2_{\alpha}(\omega)\right), \tag{S14}$$

where $\sigma^2_{\alpha}(\omega)$ is given relative to $\sigma^2_{c|k\ell:\boldsymbol{v}}(\omega)$ and $\sigma^2_{q|k\ell:\boldsymbol{v}}(\omega)$ (from Equation (S12), Theorem S1) as

$$\sigma^2_{\alpha} = \left(c^2_{k\ell:\boldsymbol{v}}(\omega) \cdot \sigma^2_{c|k\ell:\boldsymbol{v}}(\omega) + q^2_{k\ell:\boldsymbol{v}}(\omega) \cdot \sigma^2_{q|k\ell:\boldsymbol{v}}(\omega) \right) / \alpha^2_{k\ell:\boldsymbol{v}}(\omega). \tag{S15}$$

Theorem S3 (Theorem 3 in the main part)**.** *Under Assumptions S1 to S3, when $\alpha_{k\ell:\boldsymbol{v}}(\omega) > 0$ and $\omega \notin \frac{1}{2} \cdot \mathbb{Z}$, the estimate $\hat{\phi}^{m}_{k\ell:\boldsymbol{v}}(\omega) = \operatorname{args}\left(\hat{c}^{m}_{k\ell:\boldsymbol{v}}(\omega) - i \cdot \hat{q}^{m}_{k\ell:\boldsymbol{v}}(\omega)\right)$ satisfies*

$$\sqrt{n(b_1 b_2)^3/m} \cdot \left(\hat{\phi}^{m}_{k\ell:\boldsymbol{v}}(\omega) - \phi_{k\ell:\boldsymbol{v}}(\omega) \right) \xrightarrow{d} \mathrm{N}\left(0, \sigma^2_{\phi}(\omega)\right), \tag{S16}$$

where $\sigma^2_{\phi}(\omega)$ is given relative to $\sigma^2_{c|k\ell:\boldsymbol{v}}(\omega)$ and $\sigma^2_{q|k\ell:\boldsymbol{v}}(\omega)$ (from Equation (S12), Theorem S1) as

$$\sigma^2_{\phi}(\omega) = \left(q^2_{k\ell:\boldsymbol{v}}(\omega) \cdot \sigma^2_{c|k\ell:\boldsymbol{v}}(\omega) + c^2_{k\ell:\boldsymbol{v}}(\omega) \cdot \sigma^2_{q|k\ell:\boldsymbol{v}}(\omega) \right) / \alpha^4_{k\ell:\boldsymbol{v}}(\omega). \tag{S17}$$

S1.3. The proofs of Theorems 1 to 3

This section presents the proofs of the asymptotic results stated in the main part of the paper (copies of these results are given in Section S1.2). The proof of the result for the m-truncated estimate of the local Gaussian cross-spectrum $f_{k\ell:\boldsymbol{v}}(\omega)$ is in essence identical to the one encountered in JT22 for the local Gaussian auto-spectrum $f_{kk:\boldsymbol{v}}(\omega)$, 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 1 are covered in Section S2.

Proof of Theorem 1.
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 $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$, see Equation (9) in Algorithm 1(4), implies that $\hat{c}^m_{k\ell:\boldsymbol{v}}(\omega)$ and $\hat{q}^m_{k\ell:\boldsymbol{v}}(\omega)$ can be realised as the following inner products,

$$\hat{c}^m_{k\ell:\boldsymbol{v}}(\omega) = \boldsymbol{\Lambda}'_{c|\underline{\overline{m}}}(\omega) \cdot \widehat{\boldsymbol{P}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v}, \breve{\boldsymbol{v}}) \tag{S18a}$$

$$\hat{q}^m_{k\ell:\boldsymbol{v}}(\omega) = \boldsymbol{\Lambda}'_{q|\underline{\overline{m}}}(\omega) \cdot \widehat{\boldsymbol{P}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v}, \breve{\boldsymbol{v}}), \tag{S18b}$$

where $\boldsymbol{\Lambda}'_{c|\underline{\overline{m}}}(\omega)$ and $\boldsymbol{\Lambda}'_{q|\underline{\overline{m}}}(\omega)$ respectively contains the coefficients $\lambda_m(h)\cos(2\pi\omega h)$ and $\lambda_m(h)\sin(2\pi\omega h)$, and where $\widehat{\boldsymbol{P}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v}, \breve{\boldsymbol{v}}) = \left[\hat{\rho}_{\ell k:\breve{\boldsymbol{v}}}(m), \ldots, \hat{\rho}_{\ell k:\breve{\boldsymbol{v}}}(1), \hat{\rho}_{k\ell:\boldsymbol{v}}(0), \ldots, \hat{\rho}_{k\ell:\boldsymbol{v}}(m)\right]'$ contains the $2m+1$ estimates of the local Gaussian cross-correlations (based on the bandwidth $\boldsymbol{b}$). The vector $\widehat{\boldsymbol{P}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v}, \breve{\boldsymbol{v}})$ can by the help of a suitable $(2m+1) \times 5(2m+1)$ matrix $\boldsymbol{E}'_{\underline{\overline{m}}}$ (based on the vectors $\boldsymbol{e}'_5$ that gives $\rho_{k\ell:\boldsymbol{v}}(h) = \boldsymbol{e}'_5 \cdot \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}$) be expressed as

$$\widehat{\boldsymbol{P}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v}, \breve{\boldsymbol{v}}) = \boldsymbol{E}'_{\underline{\overline{m}}} \cdot \widehat{\boldsymbol{\Theta}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v}, \breve{\boldsymbol{v}}), \tag{S19}$$

where $\widehat{\boldsymbol{\Theta}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v}, \breve{\boldsymbol{v}}) = \left[\hat{\boldsymbol{\theta}}_{\ell k:\breve{\boldsymbol{v}}:m:n}, \ldots, \hat{\boldsymbol{\theta}}_{\ell k:\breve{\boldsymbol{v}}:1:n}, \hat{\boldsymbol{\theta}}_{k\ell:\boldsymbol{v}:0:n}, \ldots, \hat{\boldsymbol{\theta}}_{k\ell:\boldsymbol{v}:m:n}\right]'$ is the length $5(2m+1)$ 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

$$\begin{bmatrix} \hat{c}^m_{k\ell:\boldsymbol{v}}(\omega) \\ \hat{q}^m_{k\ell:\boldsymbol{v}}(\omega) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Lambda}'_{c|\underline{\overline{m}}}(\omega) \\ \boldsymbol{\Lambda}'_{q|\underline{\overline{m}}}(\omega) \end{bmatrix} \cdot \boldsymbol{E}'_{\underline{\overline{m}}} \cdot \widehat{\boldsymbol{\Theta}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v}, \breve{\boldsymbol{v}}), \tag{S20}$$

which together with the asymptotic normality result from Theorem S5 (page 10), i.e.,

$$\sqrt{n\left(b_1 b_2\right)^3} \cdot \left(\widehat{\boldsymbol{\Theta}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v}, \breve{\boldsymbol{v}}) - \boldsymbol{\Theta}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v}, \breve{\boldsymbol{v}})\right) \xrightarrow{d} \mathrm{N}\Big(\boldsymbol{0}, \Sigma_{k\ell:\boldsymbol{v}:\underline{\overline{m}}}\Big), \tag{S21}$$

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

$$\sqrt{n\left(b_1 b_2\right)^3 / m} \cdot \left(\begin{bmatrix} \hat{c}^m_{k\ell:\boldsymbol{v}}(\omega) \\ \hat{q}^m_{k\ell:\boldsymbol{v}}(\omega) \end{bmatrix} - \begin{bmatrix} c_{k\ell:\boldsymbol{v}}(\omega) \\ q_{k\ell:\boldsymbol{v}}(\omega) \end{bmatrix}\right) \tag{S22}$$

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

$$\frac{1}{m} \cdot \left(\begin{bmatrix} \boldsymbol{\Lambda}'_{c|\underline{\overline{m}}}(\omega) \\ \boldsymbol{\Lambda}'_{q|\underline{\overline{m}}}(\omega) \end{bmatrix} \cdot \boldsymbol{E}'_{\underline{\overline{m}}} \cdot \Sigma_{k\ell:\boldsymbol{v}:\underline{\overline{m}}} \cdot \boldsymbol{E}_{\underline{\overline{m}}} \cdot \left[\boldsymbol{\Lambda}_{c|\underline{\overline{m}}}, \boldsymbol{\Lambda}_{q|\underline{\overline{m}}}\right]\right). \tag{S23}$$

The specified form of the covariance matrix given in Theorem 1 now follows by the help of some linear algebra, the observation in Theorem S5 that

$$\Sigma_{k\ell:\boldsymbol{v}:\underline{\overline{m}}} := \left(\bigoplus_{h=m}^{1} \Sigma_{\breve{\boldsymbol{v}}|\ell k:h}\right) \bigoplus \left(\bigoplus_{h=0}^{m} \Sigma_{k\ell:\boldsymbol{v}:h}\right), \tag{S24}$$

and the definition $\tilde{\sigma}^2_{k\ell:v}(h) := \boldsymbol{e}_5' \cdot \Sigma_{k\ell:v:h} \cdot \boldsymbol{e}_5$ from Theorem S4.

It is easy to see that both $\sigma^2_{c|k\ell:v}(\omega)$ and $\sigma^2_{q|k\ell:v}(\omega)$ from Equation (S12) 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 and 3 is that they both follow as a consequence of Theorem 1 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.
First observe that the function $h(x_1, x_2) = \sqrt{x_1^2 + x_2^2}$ implies that

$$\hat{\alpha}^m_{k\ell:v}(\omega) - \alpha_{k\ell:v}(\omega) = h(\hat{c}^m_{k\ell:v}(\omega), \hat{q}^m_{k\ell:v}(\omega)) - h(c_{k\ell:v}(\omega), q_{k\ell:v}(\omega)), \tag{S25}$$

and then observe that the asymptotic covariance matrix in Theorem 1

$$\Sigma_{k\ell:v}(\omega) := \begin{pmatrix} \sigma^2_{c|k\ell:v}(\omega) & 0 \\ 0 & \sigma^2_{q|k\ell:v}(\omega) \end{pmatrix}, \tag{S26}$$

obviously is a symmetric non-negative definite matrix.

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

$$\sqrt{n(b_1 b_2)^3/m} \cdot \{h(\hat{c}^m_{k\ell:v}(\omega), \hat{q}^m_{k\ell:v}(\omega)) - h(c_{k\ell:v}(\omega), q_{k\ell:v}(\omega))\} \xrightarrow{d} \mathrm{N}\Big(0, \sigma^2_\alpha(\omega)\Big), \tag{S27}$$

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

$$\boldsymbol{D} = \left[\frac{\partial}{\partial x_1} h(x_1, x_2), \frac{\partial}{\partial x_2} h(x_1, x_2)\right] = \left[x_1/\sqrt{x_1^2 + x_2^2}, x_2/\sqrt{x_1^2 + x_2^2}\right] \tag{S28}$$

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

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

Proof of Theorem 3.
This argument is quite similar to the proof of Theorem 2, 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

$$\left[\frac{\partial}{\partial x_1} h(x_1, x_2), \frac{\partial}{\partial x_2} h(x_1, x_2)\right] = \left[-x_2/\Big(x_1^2 + x_2^2\Big), x_1/\Big(x_1^2 + x_2^2\Big)\right]. \tag{S29}$$

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

$$\sigma^2_\phi(\omega) = \Big(q^2_{k\ell:v}(\omega) \cdot \sigma^2_{c|k\ell:v}(\omega) + c^2_{k\ell:v}(\omega) \cdot \sigma^2_{q|k\ell:v}(\omega)\Big)/\alpha^4_{k\ell:v}(\omega), \tag{S30}$$

which completes the proof. □

S2. The underlying asymptotic results

S2.1. The bivariate case, a brief overview and the $\hat{\rho}_{k\ell:v}(h)$-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 JT22 [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 S1.1, see Definitions S1 to S4, and the first target of interest is to define a suitable bivariate penalty function relative to the requirement of Equation (S3). For a sample of size n from $\left\{\boldsymbol{Y}_{k\ell:h:t}\right\}_{t\in\mathbb{Z}}$, and with the present notation, the local penalty function from Tjøstheim and Hufthammer (2013) can be described as

$$\begin{aligned} Q_{k\ell:\boldsymbol{v}:h:n}(\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}) := &-\sum_{t=1}^{n} K_{k\ell:h:\boldsymbol{b}}(\boldsymbol{Y}_{k\ell:h:t} - \boldsymbol{v}) \log \psi(\boldsymbol{Y}_{k\ell:h:t}; \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}) \\ &+ n \int_{\mathbb{R}^2} K_{k\ell:h:\boldsymbol{b}}(\boldsymbol{y}_{k\ell:h} - \boldsymbol{v}) \psi(\boldsymbol{y}_{k\ell:h}; \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}) \, \mathrm{d}\boldsymbol{y}_{k\ell:h}, \end{aligned} \tag{S31}$$

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,

$$\sqrt{n\left(b_1 b_2\right)^3} \cdot \left(\hat{\boldsymbol{\theta}}_{k\ell:\boldsymbol{v}:h:n} - \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}\right) \xrightarrow{d} \mathrm{N}(\boldsymbol{0}, \Sigma_{k\ell:\boldsymbol{v}:h}). \tag{S32}$$

See Tjøstheim and Hufthammer (2013, Th. 3) for the details.

Notice that it only is the correlation component of $\hat{\boldsymbol{\theta}}_{k\ell:\boldsymbol{v}:h:n} - \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}$ that is of interest for the present paper. The relevant result for that part is stated below in order to give a reference for the statements in Theorem 1.

Theorem S4. *Under Assumptions S1 to S3, the following univariate asymptotic result holds for the estimates $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ of the local Gaussian cross-correlations $\rho_{k\ell:\boldsymbol{v}}(h)$.*

$$\sqrt{n\left(b_1 b_2\right)^3} \cdot \left(\hat{\rho}_{k\ell:\boldsymbol{v}}(h) - \rho_{k\ell:\boldsymbol{v}}(h)\right) \xrightarrow{d} \mathrm{N}\Big(0, \tilde{\sigma}^2_{k\ell:\boldsymbol{v}}(h)\Big), \tag{S33}$$

where $\tilde{\sigma}^2_{k\ell:\boldsymbol{v}}(h) := \boldsymbol{e}_5' \cdot \Sigma_{k\ell:\boldsymbol{v}:h} \cdot \boldsymbol{e}_5$.

Proof. With $\boldsymbol{e}_5'$ the unit vector that picks out the correlation part from $\hat{\boldsymbol{\theta}}_{k\ell:\boldsymbol{v}:h:n} - \boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}$, it follows from Brockwell and Davis (1986, proposition 6.4.2, p. 211) that

$$\sqrt{n\left(b_1 b_2\right)^3} \cdot \left(\hat{\rho}_{k\ell:\boldsymbol{v}}(h) - \rho_{k\ell:\boldsymbol{v}}(h)\right) \xrightarrow{d} \mathrm{N}\Big(0, \boldsymbol{e}_5' \cdot \Sigma_{k\ell:\boldsymbol{v}:h|5} \cdot \boldsymbol{e}_5\Big). \tag{S34}$$

□

S2.2. The asymptotic result for $\widehat{\boldsymbol{\Theta}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v}, \breve{\boldsymbol{v}})$

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 JT22 [Appendix B].

Definition S5. *For each bivariate penalty function $Q_{k\ell:\boldsymbol{v}:h:n}(\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h})$ (as given in Equation (S31)), denote by $\tilde{Q}_{k\ell:\boldsymbol{v}:h:n}(\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h})$ the extension of it from a function of $\boldsymbol{Y}_{k\ell:h:t} := \left[Y_{k,t+h}, Y_{\ell,t}\right]'$ to a function of $\left[Y_{k,t+m}, \ldots, Y_{k,t}, Y_{\ell,t+m}, \ldots, Y_{\ell,t}\right]'$. Use these extensions do define the new penalty function*

$$Q_{\boldsymbol{v}:\breve{\boldsymbol{v}}|k\ell:\underline{\overline{m}}:n}\Big(\boldsymbol{\Theta}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v}, \breve{\boldsymbol{v}})\Big) := \sum_{h=m}^{1} \tilde{Q}_{\ell k:\breve{\boldsymbol{v}}:h:n}(\boldsymbol{\theta}_{\ell k:\breve{\boldsymbol{v}}:h}) + \sum_{h=0}^{m} \tilde{Q}_{k\ell:\boldsymbol{v}:h:n}(\boldsymbol{\theta}_{k\ell:\boldsymbol{v}:h}). \tag{S35}$$

The $2m+1$ bivariate components in the sum that defines $Q_{\boldsymbol{v}:\check{\boldsymbol{v}}|k\ell:\underline{\overline{m}}:n}\Big(\widehat{\boldsymbol{\Theta}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v},\check{\boldsymbol{v}})\Big)$ have no common parameters, so the optimisation of the parameters for the different summands can be performed independently. The optimal parameter vector $\widehat{\boldsymbol{\Theta}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v},\check{\boldsymbol{v}})$ for the penalty function $Q_{\boldsymbol{v}:\check{\boldsymbol{v}}|k\ell:\underline{\overline{m}}:n}$ (for a given sample) can thus be constructed by stacking on top of each other the parameter vectors $\hat{\boldsymbol{\theta}}_{k\ell:\boldsymbol{v}:h}$ and $\hat{\boldsymbol{\theta}}_{\ell k:\check{\boldsymbol{v}}:h}$ that optimise the individual summands in Equation (S35).

The Klimko-Nelson approach can now be used on the penalty function from Equation (S35) i.e., four requirements related to the penalty function must be verified before the desired asymptotic result for the paramater-vector $\widehat{\boldsymbol{\Theta}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v},\check{\boldsymbol{v}})$ is obtained. The following cross-spectrum analogue of JT22 [Theorem B.23] can now be stated for the present case of interest.

Theorem S5. *Under Assumptions S1 to S3, the following asymptotic behaviour holds for the estimated parameters* $\widehat{\boldsymbol{\Theta}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v},\check{\boldsymbol{v}}) = \big[\hat{\boldsymbol{\theta}}_{\ell k:\check{\boldsymbol{v}}:m:n}{}', \ldots, \hat{\boldsymbol{\theta}}_{\ell k:\check{\boldsymbol{v}}:1:n}{}', \hat{\boldsymbol{\theta}}_{k\ell:\boldsymbol{v}:0:n}{}', \ldots, \hat{\boldsymbol{\theta}}_{k\ell:\boldsymbol{v}:m:n}{}'\big]'$,

$$\sqrt{n(b_1 b_2)^3}\cdot\Big(\widehat{\boldsymbol{\Theta}}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v},\check{\boldsymbol{v}}) - \boldsymbol{\Theta}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v},\check{\boldsymbol{v}})\Big) \xrightarrow{d} \mathrm{N}\Big(\mathbf{0}, \Sigma_{k\ell:\boldsymbol{v}:\underline{\overline{m}}}\Big), \tag{S36}$$

where the matrix $\Sigma_{k\ell:\boldsymbol{v}:\underline{\overline{m}}}$ *is the direct sum of the matrices from Equation (S32) that occurs when the individual bivariate components of the penalty function are investigated, i.e.,*

$$\Sigma_{k\ell:\boldsymbol{v}:\underline{\overline{m}}} := \left(\bigoplus_{h=m}^{1}\Sigma_{\check{\boldsymbol{v}}|\ell k:h}\right)\bigoplus\left(\bigoplus_{h=0}^{m}\Sigma_{k\ell:\boldsymbol{v}:h}\right). \tag{S37}$$

Proof. This result follows when the Klimko-Nelson approach is used with the local penalty-function $Q_{\boldsymbol{v}:\check{\boldsymbol{v}}|k\ell:\underline{\overline{m}}:n}\Big(\boldsymbol{\Theta}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v},\check{\boldsymbol{v}})\Big)$ from Equation (S35), and the proof is in essence identical to the proof of JT22 [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 JT22, 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 $\mathfrak{Q}^n_{\boldsymbol{v}|\underline{\overline{m}}}$ that has the same limiting distribution as the estimator of $\sqrt{b_1 b_2}\boldsymbol{\nabla}_{k\ell:\underline{\overline{m}}}Q_{\boldsymbol{v}:\check{\boldsymbol{v}}|k\ell:\underline{\overline{m}}:n}\Big(\boldsymbol{\Theta}_{k\ell:\underline{\overline{m}}|\boldsymbol{b}}(\boldsymbol{v},\check{\boldsymbol{v}})\Big)$ (where $\boldsymbol{\nabla}_{k\ell:\underline{\overline{m}}}$ is obtained by stacking together $\boldsymbol{\nabla}_{k\ell:h}$). After this, it is sufficient to use standard methods to prove that the limiting distribution of $\mathfrak{Q}^n_{\boldsymbol{v}|\underline{\overline{m}}}$ is the desired multivariate normal distribution, and the statement for the parameter vectors then follows from the Klimko-Nelson theorem and some linear algebra. □

S3. Sensitivity analysis of the tuning parameters

An investigation of how sensitive the estimates of the local Gaussian auto-spectra $\hat{f}^m_{\ell\ell:\boldsymbol{v}}(\omega)$ are to changes in the tuning parameters (and the point $\boldsymbol{v}$) can be found in JT22 [Appendix D]. This section presents a similar investigation for the estimates of the local Gaussian cross-spectra $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$, and it is seen here that the features observed for the univariate case are also present in the multivariate case.

Different plots based on estimates of the local Gaussian cross-correlations and the corresponding local Gaussian cross-spectra, i.e., $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ and $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$, will be encountered, and some of these plots are based on the distance function D from JT22 [Appendix D.1]. The definition of D, and a short discussion of it, is for the convenience of the reader repeated in Section S3.1 in the present Supplementary Material.

Sections S3.2 and S3.3 respectively consider the sensitivity of the point $\boldsymbol{v}$ and the bandwidth $\boldsymbol{b}$, whereas the sensitivity of the truncation level m is discussed in Section S3.4.

The effect the value of the block length L has upon the bootstrap-based pointwise confidence intervals is discussed in Section S5.3, since that gives the most natural flow.

The scripts required for the replication of the results in this section are contained in the R-package `localgaussSpec`, and these scripts can be used as templates for those that would like to investigate other time series in a similar manner. See Section S6.1 for details.

S3.1. Sensitivity analysis - the distance function

The purpose of this section is to repeat *the motivation for* and *the definition of* the distance function D given in JT22 [Appendix D.1]. The discussion of possible alternatives will not be included here.

Motivation: An investigation of the sensitivity of the different tuning parameters requires a tool that can measure the differences that occur when these tuning parameters are adjusted. The distance function D used in this paper is the one inherited from the complex Hilbert space of Fourier series on the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$, cf. e.g. Brockwell and Davis (1986, Ch. 2.8), i.e., for $f(\omega) = \sum_{h=-\infty}^{\infty} \rho(h) e^{-2\pi i h}$ the norm is defined by $||f(\omega)||^2 = \int_{-1/2}^{1/2} f(\omega)\overline{f(\omega)}\, \mathrm{d}\omega = \sum_{h=-\infty}^{\infty} \rho(h)^2$. This motivates the following definition.

Definition S6. *Given two spectra* $f_1(\omega) = \sum_{h=-\infty}^{\infty} \rho_1(h) e^{-2\pi i h}$ *and* $f_2(\omega) = \sum_{h=-\infty}^{\infty} \rho_2(h) e^{-2\pi i h}$, *the distance between them is denoted by*

$$D(f_1(\omega), f_2(\omega)) := \sqrt{\sum_{h=-\infty}^{\infty} \big(\rho_1(h) - \rho_2(h)\big)^2}. \tag{S38}$$

Furthermore: The notation $D(f_1(\omega))$ *will be interpreted as* $D(f_1(\omega), 0)$, *which implies that* $D(f_1(\omega), f_2(\omega))$ *also can be written as* $D(f_1(\omega) - f_2(\omega))$ *(which is used in Figure S3.10).*

The obvious adjustment must be applied when the distance function D is used on m-truncated estimates $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$, i.e., $\rho(h)$ should be replaced with $\lambda_m(h) \cdot \hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ when $|h| \leq m$, and with 0 when $|h| > m$.

Regarding the frequency-dimension: The distance measure D given in Definition S6 does not contain any information about the frequencies, and completely different spectral densities can have the same distance-value. It is thus, for the purpose of sensitivity analysis, important to combine distance-based plots with plots that reveal something about the frequency-component too.

S3.2. Sensitivity analysis: The point $\boldsymbol{v}$

The point $\boldsymbol{v}$, contrary to the bandwidth $\boldsymbol{b}$ and the truncation level m, is not a tuning parameter of the estimation algorithm — but it is natural to investigate how the local Gaussian spectra $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ varies with $\boldsymbol{v}$. Due to this it is also of interest to consider some plots that can reveal how $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ behaves as a function of the selected point $\boldsymbol{v}$.

Figures 2, 4 and 7 of the main part used heatmap- and distance-plots to show how $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ varied with the selected point $\boldsymbol{v}$ (restricted to the diagonal), in particular Figures 2 and 4 considered the *bivariate local trigonometrical* examples used for the sanity testing of the implemented estimation algorithm, whereas Figure 7 looked at the DAX- and CAC-components of the `EuStockMarkets`-data. These plots are extensions to the multivariate case of the corresponding heatmap- and distance-plots that was introduced for the univariate case in JT22.

S3.2.1. Revisiting the univariate case

For a local Gaussian investigation of a multivariate sample, like the `EuStockMarkets`-data, it is of course important to investigate both the local Gaussian cross-spectra $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ and the marginal local Gaussian auto-spectra $\hat{f}^m_{kk:\boldsymbol{v}}(\omega)$. Figure 7 from the main-part should

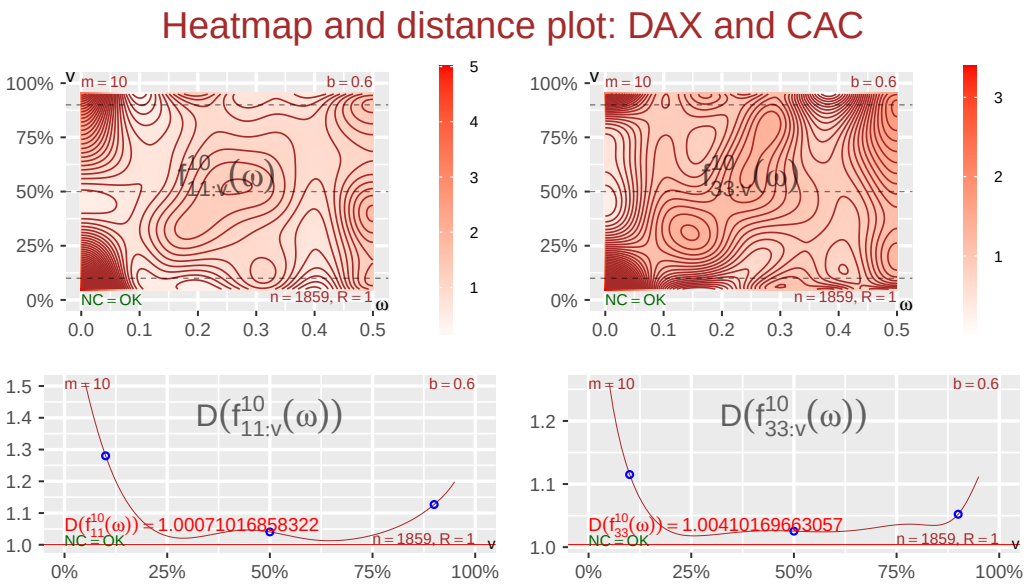

Figure S3.1. Heatmap- and distance-plots for the marginal DAX- and CAC-components of the `EuStockMarkets`-data, the DAX-component to the left. Investigation with focus on how the local Gaussian auto-spectra $\hat{f}^{10}_{kk:\boldsymbol{v}}(\omega)$ changes when the point $\boldsymbol{v}$ varies along the diagonal. The three points used in Figure 9 have been highlighted with lines/points.

thus be accompanied with the marginal heatmap- and distance-plots of the DAX- and CAC-components of `EuStockMarkets`, as seen in Figure S3.1.

Note that the points $\boldsymbol{v} = (v_1, v_2)$ seen in Figure S3.1 only varies along the diagonal, i.e., $v_1 = v_2$. This was used in JT22 since it simplified the graphical inspection (the local Gaussian auto-spectra are always real-valued on the diagonal). The points $\boldsymbol{v} = (v_1, v_2)$ can of course also vary along non-diagonal lines, but the resulting visualisation for the local Gaussian auto-spectra must then deal with complex-valued results instead of real-valued results.

Figure S3.1 shows how heatmap-plots can be used to see how the local Gaussian auto-spectra $\hat{f}^{m}_{kk:\boldsymbol{v}}(\omega)$ varies with diagonal points $\boldsymbol{v}$ (for a fixed bandwidth $\boldsymbol{b}$ and a fixed truncation level m). The distance function D from Definition S6 has been used to create the distance-based plots, and these show how the norms $D\left(\hat{f}^{m}_{kk:\boldsymbol{v}}(\omega)\right)$ varies with $\boldsymbol{v}$. Note that the scales are different for the DAX- and CAC-components in Figure S3.1, with the DAX-component (left side) having the highest values.

The distance-parts of Figure S3.1 also includes the distance-based value computed from the ordinary (m-truncated) autospectra (red horizontal line). A comparison of the distances from the ordinary spectra and the local Gaussian spectra can provide an indication of the presence of non-Gaussian dependency structures in the time series under investigation. Note that a comparison of the global spectra and the local Gaussian spectra is not possible based on the heatmaps seen in Figure S3.1 (since they only reveal information based on the local Gaussian spectra), and it is thus also necessary to consider plots like those seen in Figures 1, 3, 5, 9 and 12 in the main part.

The points $\boldsymbol{v}$ in Figure S3.1 ranges from the 5% percentile to the 95% percentile of the standard normal distribution, increasing in steps of 0.5% (altogether 91 different points). This percentile based selection implies that the corresponding points are not equally spaced along the actual diagonal, and the plots in Figure S3.1 (and all similar $\boldsymbol{v}$-investigating plots) have thus used the option that the points $\boldsymbol{v}$ have been presented according to their underlying percentile-values — which implies that these plots primarily reveals information about the copula-structure of the time series under investigation.

The features seen for the DAX- and CAC-parts of Figure S3.1 are similar to those observed for the `dmbp`-example[4] investigated in JT22. In particular, the local Gaussian auto-spectra $\hat{f}^m_{kk:\boldsymbol{v}}(\omega)$ are rather flat near near the 50% percentile, and in this region the spectra are quite similar to those encountered from a sample from an i.i.d. white noise situation. Moreover, it is also here seen (as it was for the `dmbp`-case) that there is a clear asymmetric behaviour between the lower tail and upper tail, and this asymmetry is easier to see from the distance-parts of Figure S3.1.[5]

The observed asymmetry, with a higher peak at the lower tail, are in agreement with the asymmetry between a *bear market* (going down) and a *bull market* (going up).

It must be added that the 5% and 95% percentiles are quite far out in the tails of the distribution, and it is thus natural to assume that the selected bandwidth in those cases might fail to work properly — and the small sample variation of the points closest to the point $\boldsymbol{v}$ might then render the estimated local Gaussian autocorrelations rather dubious. It is possible to counter this problem by selecting a larger bandwidth for percentiles in the tails, but it is then important to keep in mind that a too large bandwidth might completely miss the desired local structure at the point of investigation.

S3.2.2. Extension to the multivariate case

The restriction to diagonal points, as seen in Figure S3.1, ensures that the resulting local Gaussian auto-spectra $\hat{f}^m_{kk:\boldsymbol{v}}(\omega)$ always are real-valued. The local Gaussian cross-spectra $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ will however be complex-valued, and some modifications of the plotting procedure are thus needed when these plots are extended to the multivariate case.

The distance function D from Definition S6 works equally well for real-valued and complex-valued spectra, so the distance-part of the plots remains exactly the same.[6] The heatmap-part of the plots must however be updated, and a natural strategy for this endeavour is to split the complex-valued local Gaussian cross-spectra $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ into a combination of the related real-valued spectra given in Definition 2 in the main part

A complex number z can be represented in a Cartesian form, i.e., $z = x + i \cdot y$, where x and y respectively are the real and imaginary parts of the number. From the definition of the `Co`-spectrum $c_{k\ell:\boldsymbol{v}}(\omega)$ and `Quad`-spectrum $q_{k\ell:\boldsymbol{v}}(\omega)$, it is clear that $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega) = c_{k\ell:\boldsymbol{v}}(\omega) - i \cdot q_{k\ell:\boldsymbol{v}}(\omega)$. This implies that one way to adjust the heatmap-plot to the bivariate case is to present a pair of plots, where one plot represent the `Co`-spectrum and the other represent the `Quad`-spectrum. This approach was used in Figures 2, 4 and 7 in the main part.

The complex number z can also be represented in a polar form, i.e., $z = \alpha \cdot e^{i \cdot \phi}$, where α and ϕ respectively are the modulus and the phase of the complex number. From the definition of the `Amplitude`-spectrum $\alpha_{k\ell:\boldsymbol{v}}(\omega)$ and `Phase`-spectrum $\phi_{k\ell:\boldsymbol{v}}(\omega)$, it is clear that $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega) = \alpha_{k\ell:\boldsymbol{v}}(\omega) \cdot e^{i \cdot \phi_{k\ell:\boldsymbol{v}}(\omega)}$. This implies that another possible extension of the heatmap-plot to the bivariate case is to use a pair of plots, where one plot represent the `Amplitude`-spectrum and the other plot represent the `Phase`-spectrum.

Figure S3.2 shows a heatmap- and distance plot for the local Gaussian cross-spectrum for the DAX- and CAC-components of the `EuStockMarkets`-data, where the heatmap-part use the `Amplitude + Phase` decomposition. A comparison with Figure 7 from the main part reveals that the information gained from the `Amplitude`-spectrum part of Figure S3.2 in this case is similar to the one gained from the `Co`-spectrum part of Figure 7. A similar

[4] The Deutschemark/British pound Exchange Rate (`dmbp`) data from Bollerslev and Ghysels (1996), which is a common benchmark data set for GARCH-type models. The data used in JT22 was found in the `R`-package `rugarch`, see Ghalanos (2022), where the following description was given: "The daily percentage nominal returns computed as $100[\ln(P_t) - \ln(P_t - 1)]$, where P_t is the bilateral Deutschemark/British pound rate constructed from the corresponding U.S. dollar rates."

[5] It can also be seen that the DAX-component (left side of Figure S3.1) have a slightly more extreme behaviour in the lower tail than the one observed for the CAC-component (right side).

[6] It is worth noticing that the lowest value that can occur for the distance is 1 when an auto-spectrum is investigated, whereas it is 0 when a cross-spectrum is investigated, i.e., $D\left(\hat{f}^m_{kk:\boldsymbol{v}}(\omega)\right) \geq 1$ and $D\left(\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)\right) \geq 0$ when $k \neq \ell$. The values seen in the distance-plot thus gives an investigator an idea with regard to how far away from an i.i.d. white noise situation the resulting auto- and cross-spectra are.

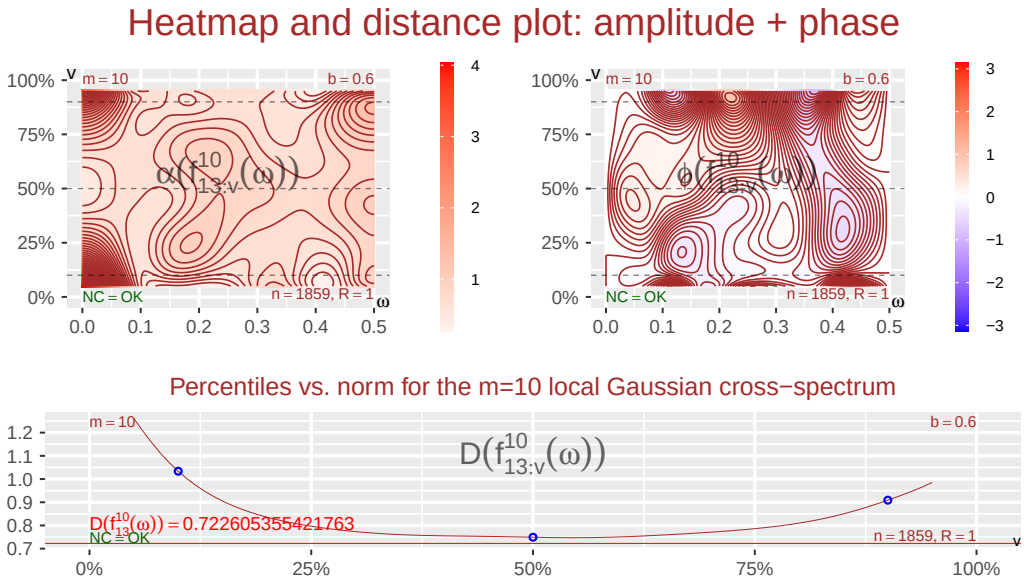

Figure S3.2. Heatmap and corresponding distance-based plots based on the DAX- and CAC-components of the `EuStockMarkets`-data, showing how the complex-valued local Gaussian cross-spectrum $\hat{f}^{10}_{k\ell:v}(\omega)$ changes when when the point v varies along the diagonal. Investigation based on the `Amplitude`- and `Phase`-spectra, cf. Figure 7 the main part for the `Co`- and `Quad`-spectra version. The three points used in Figure 9 have been highlighted with lines/points.

comparison of the `Quad`-spectrum and the `Phase`-spectrum from these reveals that these also are quite similar (at least after taking into account that the definition of the `Quad`-spectrum is minus one times the imaginary part of the cross-spectrum explains why positive and negative regions have been inverted).

Based on the details seen in Figure S3.2 and Figure 7, it might be tempting to conclude that it hardly matters what kind of decomposition (Cartesian or polar) that is used for the inspection of the complex-valued local Gaussian spectra, but it should be noted that the `Quad`-spectrum in this particular case have much smaller values than the `Co`-spectrum, and that could be the source of the similarity.

The two cases investigated in Figures 2 and 4, i.e., the *bivariate local trigonometrical* examples used for the sanity testing of the implemented estimation algorithm, can also be investigated by the help of a polar composition of the complex-valued local Gaussian cross-spectrum, and the resulting plots are presented in Figures S3.3 and S3.4.

An inspection of the `Co`- and `Quad`- spectra in Figures 2 and 4 clearly reveals that the two *bivariate local trigonometrical* examples have different local properties, i.e., the peaks and troughs occur at different frequencies for different points v. The distance-parts of Figures 2 and 4 did, however, not show this difference — and this highlights why it is important to include a visualisation of the frequency component.

As seen in the previous paragraph, the Cartesian-type decomposition of the local Gaussian cross-spectrum, i.e., $\hat{f}^{m}_{k\ell:v}(\omega) = c_{k\ell:v}(\omega) - i \cdot q_{k\ell:v}(\omega)$, made it easy to spot the differences between the two *bivariate local trigonometrical* examples. What about the strategy based on the polar-type decomposition, i.e., $\hat{f}^{m}_{k\ell:v}(\omega) = \alpha_{k\ell:v}(\omega) \cdot e^{i \cdot \phi_{k\ell:v}(\omega)}$?

The complex-valued data used for the construction of Figures 2 and 4 was also used for the construction of Figures S3.3 and S3.4, but it is quite a bit harder to use the latter pair of plots to spot the differences between the two *bivariate local trigonometrical* examples. In fact the `Amplitude`-spectra parts of Figures S3.3 and S3.4 are so similar that it is hard to tell them apart. The `Phase`-spectra parts are similarly quite hard to investigate, in particular since the values in this case are restricted to the range $(-\pi, \pi]$. Unless the `Phase`-spectra values have opposite signs, it is very hard to compare the colour-gradients in the `Phase`-spectra

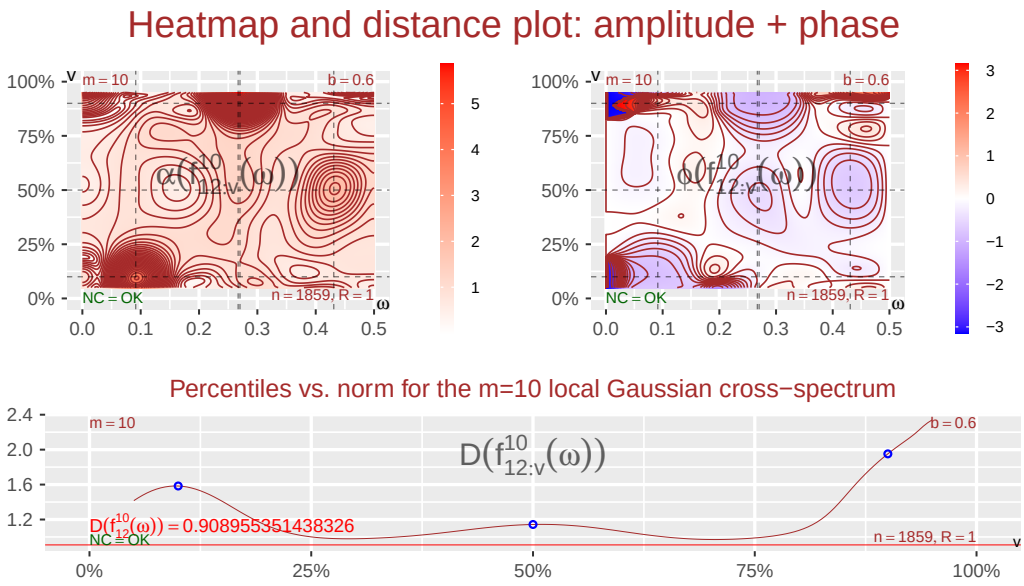

Figure S3.3. Heatmap and corresponding distance-based plots based on the *local trigonometric constant phase-adjustment* example, showing how the complex-valued local Gaussian cross-spectrum $\hat{f}^{10}_{k\ell:v}(\omega)$ changes when when the point v varies along the diagonal. Investigation based on the `Amplitude`- and `Phase`-spectra, cf. Figure 2 in the main part for the `Co`- and `Quad`-spectra version. The three points used in Figure 3 have been highlighted with lines/points. Vertical lines have been added to the heatmap-plots to highlight the frequencies used to generate the *local trigonometric example*.

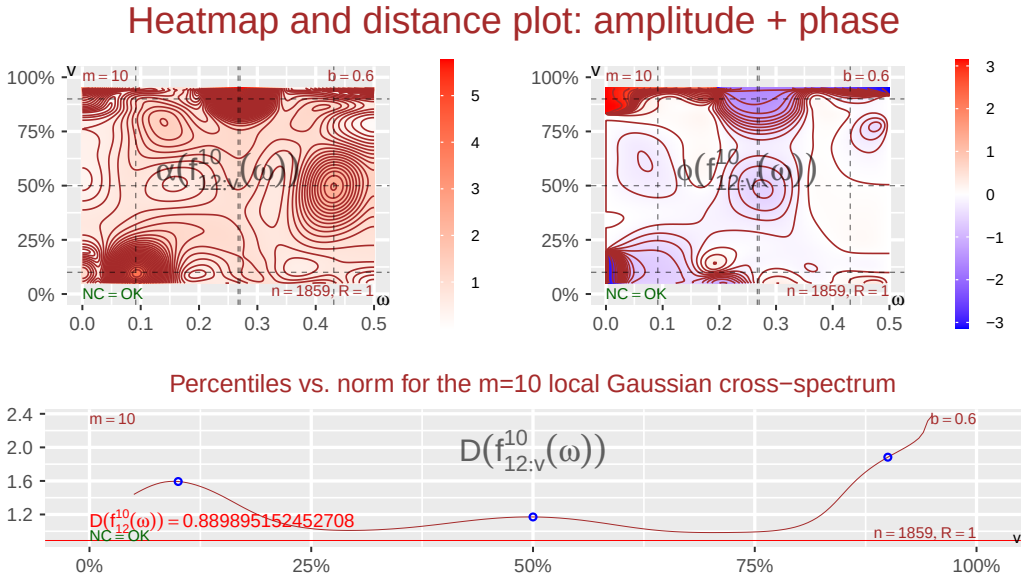

Figure S3.4. Heatmap and corresponding distance-based plots based on the *local trigonometric different phases-adjustment* example, showing how the complex-valued local Gaussian cross-spectrum $\hat{f}^{10}_{k\ell:v}(\omega)$ changes when when the point v varies along the diagonal. Investigation based on the `Amplitude`- and `Phase`-spectra, cf. Figure 4 in the main part for the `Co`- and `Quad`-spectra version. The three points used in Figure 5 have been highlighted with lines/points. Vertical lines have been added to the heatmap-plots to highlight the frequencies used to generate the *local trigonometric example*.

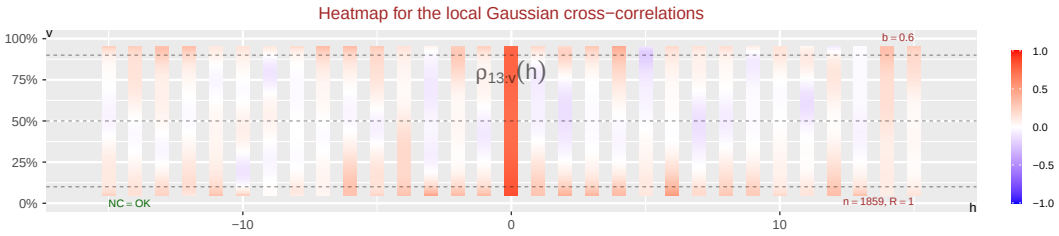

Figure S3.5. Heatmap of the local Gaussian cross-correlations $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ based on the DAX- and CAC-components of the `EuStockMarkets`-data, showing how these estimates change when the point $\boldsymbol{v}$ varies along the diagonal. The three points used in Figure 9 have been highlighted with lines.

parts of Figures S3.3 and S3.4 and conclude that there are significant differences between the two *bivariate local trigonometrical* examples that are investigated in these plots.

The preceding discussion shows that it is easier to digest the information stored in the complex-valued local Gaussian cross-spectrum when the heatmap-parts are decomposed into `Co`- and `Quad`-spectra — and this motivated the use of this version for the heatmap- and distance-plots included in the main part, i.e., Figures 2, 4 and 7.

S3.2.3. Heatmap-plots for the estimates $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$

The computation of the complex-valued local Gaussian cross-spectra $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ requires the computation of the $2m+1$ local Gaussian cross-correlations $\{\hat{\rho}_{k\ell:\boldsymbol{v}}(h)\}_{h=-m}^{m}$, and it is thus (as a supplement to the plots seen in Figures 2, 4 and 7 in the main part and Figures S3.2 to S3.4 in the preceding discussion) also possible to create heatmap-based plots that can visualise how these estimates change as the point $\boldsymbol{v}$ moves along the diagonal from the 5% to the 95% percentile. Figure S3.5 presents a heatmap-plot for the local Gaussian cross-correlation between the DAX- and CAC- components of the `EuStockMarkets`-data.

Note that the values $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ by construction lies in the interval $[-1,1]$, and the colour-gradient has thus been predefined to use this range. It is quite easy to see from Figure S3.5 if the $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$-values are close to zero (white) or if they are closer to the endpoints ± 1. It can from such an investigation also be seen if some lags behave differently from the others, and it is also possible to see if it might be an asymmetric situation between the behaviour in lower and upper tails. Finally, it can also help reveal if the estimation of the values in the tails might be somewhat dubious, since that often will reveal itself by values that becomes quite close to ± 1, or suddenly changes from one extreme to the other.

S3.3. Sensitivity analysis: The bandwidth $\boldsymbol{b}$

The bandwidth $\boldsymbol{b} = (b_1, b_2)$ is a bivariate parameter, but it was noted in JT22 that it is quite natural to assume $b_1 = b_2$ for univariate time series $\{Y_t\}_{t=1}^n$. This is due to the local Gaussian autocorrelations $\rho_{kk:\boldsymbol{v}}(h)$ being computed based on (pseudo-normalised versions of the) lag h pairs (Y_{t+h}, Y_t), and the marginal time series in this case thus coincide with the one we started with.

The justification from the univariate case for $b_1 = b_2$ does not work when the local Gaussian cross-correlations $\rho_{k\ell:\boldsymbol{v}}(h)$ are to be estimated for a multivariate time series, but in this case it is possible to use the pseudo-normalisation of the marginals as a justification for the assumed equality of b_1 and b_2.

With this restriction it follows that the sensitivity of the complex-valued local Gaussian cross-spectrum $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$, due to changes in the bandwidth $\boldsymbol{b}$, can be investigated in a similar manner to the one used in the preceding section for the diagonal points $\boldsymbol{v}$. Figures S3.6 and S3.7 show the results of such an investigation (at the point `10%::10%`) for the complex-valued local Gaussian cross-spectrum $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ based on the DAX- and CAC-components of the `EuStockMarkets`-data. Figure S3.6 gives a Cartesian decomposition into the `Co`- and `Quad`-spectra, whereas Figure S3.7 presents the polar decomposition into the `Amplitude`- and `Phase`-spectra. As mentioned before: Since the scale of the imaginary part of $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ is much smaller than the real part in this particular case, the `Amplitude`-spectrum is in

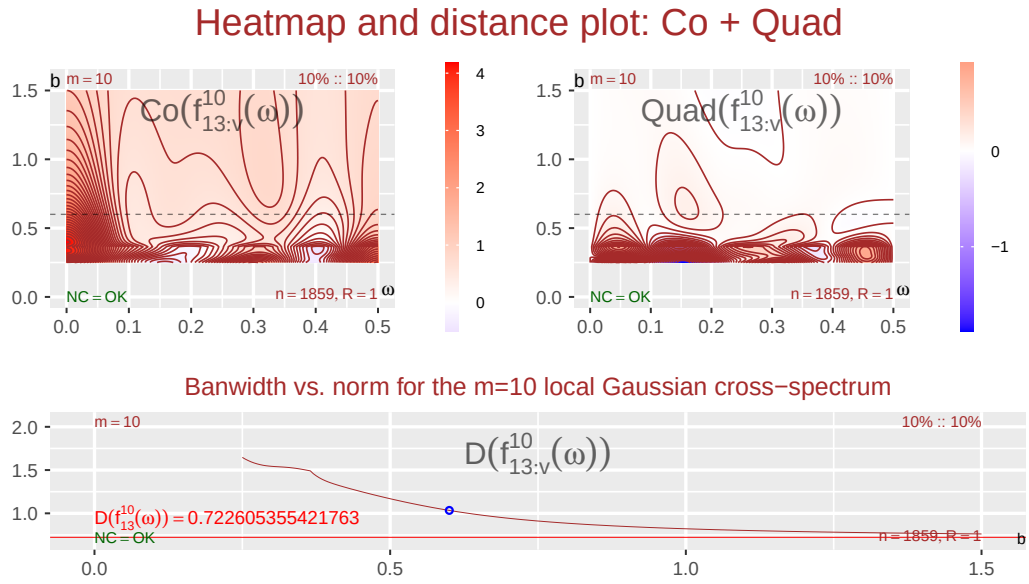

Figure S3.6. Heatmap and corresponding distance-based plots based on the DAX- and CAC-components of the `EuStockMarkets`-data, showing how the complex-valued local Gaussian cross-spectrum $\hat{f}^{10}_{k\ell:v}(\omega)$ changes when when the bandwidth $\boldsymbol{b}$ varies. Investigation based on `Co`- and `Quad`-spectra, cf. Figure S3.7 for the the `Amplitude`- and `Phase`-spectra version. The bandwidth $b = 0.6$ used in Figure 9 has been highlighted with a line/point.

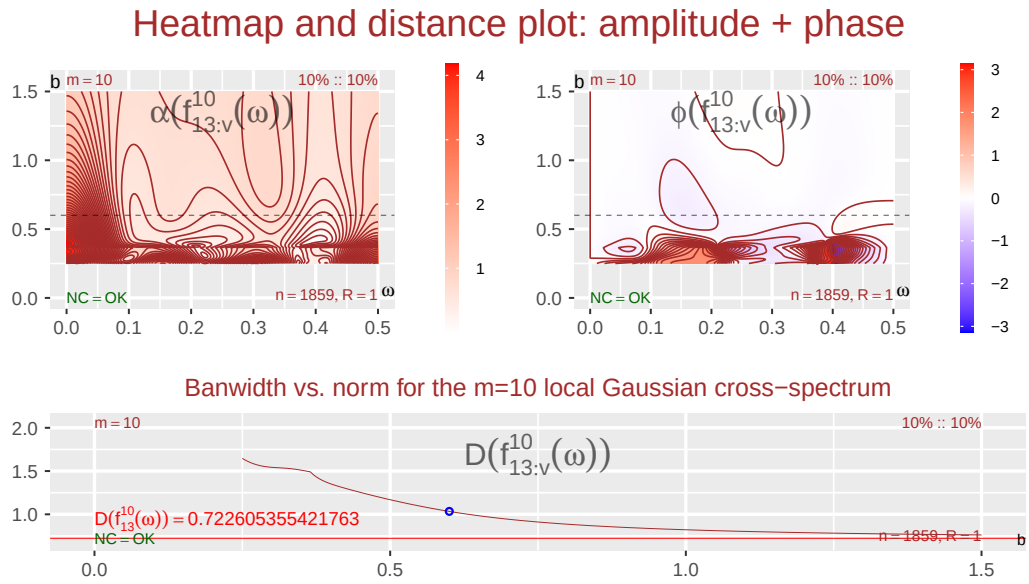

Figure S3.7. Heatmap and corresponding distance-based plots based on the DAX- and CAC-components of the `EuStockMarkets`-data, showing how the complex-valued local Gaussian cross-spectrum $\hat{f}^{10}_{k\ell:v}(\omega)$ changes when when the bandwidth $\boldsymbol{b}$ varies. Investigation based on the `Amplitude`- and `Phase`-spectra, cf. Figure S3.6 for the `Co`- and `Quad`-spectra version. The bandwidth $b = 0.6$ used in Figure 9 has been highlighted with a line/point.

essence the absolute value of the `Co`-spectrum — and since the `Co`-spectrum is positive for this example it follows that the plots of the `Amplitude`-spectrum is very similar to the plot of the `Co`-spectrum in this particular case.

As mentioned in JT22 [Appendix D.3]: "The bandwidth $\boldsymbol{b}$ should be selected according to the Goldilocks principle, i.e. it should neither be 'too low' nor 'too high', it must be 'just right'." An inspection of the heatmap and distance-based plots in Figures S3.6 and S3.7 can

provide some information with regard to what kind of problems that can occur when the bandwidth does not belong to the "just right" region. The bandwidths used in Figures S3.6 and S3.7 starts at 0.25 and ends at 1.5 (increasing in steps of 0.005, altogether 251 different bandwidths), and the bandwidth $b = 0.6$ has been highlighted since it was that value that was used in Figures 7 to 9 in the main part.

The problems occurring for the multivariate case are the same as those discussed in JT22 [Appendix D.3] for the univariate case. If $\boldsymbol{b}$ becomes too large, then the estimated local Gaussian cross-correlations $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ will no longer capture the local structure of interest, and the corresponding estimated local Gaussian spectral density $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ (which no longer deserves to be referred to as "local") will then be indistinguishable from the ordinary spectral density. It is clear from Figures S3.6 and S3.7 that a bandwidth greater than $b = 1.0$ will be too large for the investigation of the `EuStockMarkets`-data.

On the other side, it is expected that a too low bandwidth will trigger a degeneration of the estimated local Gaussian cross-correlations, i.e., $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ will tend towards either $+1$ or -1 regardless of the actual structure of the underlying density distributions. The reason for this is (as previously explained in JT22 for the local Gaussian auto-correlation case) that $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ will, due to the kernel function from the density estimation algorithm, become increasingly sensitive to the position of the (pseudo-normalised) h-lagged pairs $\left(Y_{k,t+h}, Y_{\ell,t}\right)$ that lies nearest to the point $\boldsymbol{v} = (v_1, v_2)$. To clarify, for a given point $\boldsymbol{v}$ there will be a collection of Euclidean distances to the (pseudo-normalised versions of the) h-lagged pairs $\left(Y_{k,t+h}, Y_{\ell,t}\right)$ in the sample, and these distances could (after a re-indexing) be sorted in ascending order $\{d_i\}_{i=1}^{n-h}$.

Under the assumption that it is the product normal kernel that is used, the contribution from a lag-h pair $\left(Y_{k,t+h}, Y_{\ell,t}\right)$ that lies a distance of d_i from $\boldsymbol{v}$ will be weighted by $w_{i:\boldsymbol{b}} := \frac{1}{2\pi b^2} e^{-d_i^2/2b^2}$ — and it is now natural to consider the set of all the weights $\mathcal{W}_{\boldsymbol{v}:\boldsymbol{b}} := \{w_{i:\boldsymbol{b}}\}_{i=1}^{n-h}$.

The primary detail of interest is how much larger the weights are for the pairs that are closest to $\boldsymbol{v}$, and it is thus necessary to consider the fraction $r_{ij:\boldsymbol{b}} := w_{i:\boldsymbol{b}}/w_{j:\boldsymbol{b}} = \left(e^{d_j^2 - d_i^2}\right)^{1/b^2}$. The number $r_{ij:\boldsymbol{b}}$ will, when $d_i < d_j$, grow to ∞ when $\boldsymbol{b} \to \boldsymbol{0}^+$, and this implies that the estimation algorithm for small b-values will become increasingly sensitive to the h-lagged pairs that are closest to the point $\boldsymbol{v}$ when the bandwidth shrinks — and in the end it would thus be natural to have a degeneration of the estimated value $\rho_{\boldsymbol{v}}(h)$ to either $+1$ or -1.

Note that the corresponding $D\left(\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)\right)$ will grow when this degeneration happens, as can be seen for $b = 0.25$ in the distance-part plots in Figures S3.6 and S3.7.

Heatmap-plots for the estimates $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$: It is here, as it was for the investigation of the diagonal points $\boldsymbol{v}$, possible to also consider a heatmap-based investigation of the underlying estimates $\{\hat{\rho}_{k\ell:\boldsymbol{v}}(h)\}_{h=-m}^{m}$. Such a plot is given in Figure S3.8, and it can there be observed how some of the $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$-estimates changes from having a value close to zero to a value that must be close to $+$ or -1 as the bandwidth $\boldsymbol{b}$ decreases. This is, as mentioned above, not a surprising observation — since this behaviour is expected when the bandwidth $\boldsymbol{b}$ has shrunk to a level where the kernel function in the local penalty function gives quite high weights to the few observations $\left(Y_{k,t+h}, Y_{\ell,t}\right)$ nearest $\boldsymbol{v}$, and very low weights elsewhere.

Note that Figures S3.6 to S3.8 consider the situation where $\boldsymbol{v}$ is the "lower tail" diagonal point at `10%::10%`. Similar heatmap- and distance-plots could have been included for the cases where $\boldsymbol{v}$ corresponds to either the center or the upper tail.[7] A comparison of the distance-based plots for these three points is presented in Figure S3.9, in order to show how the bandwidth-sensitivity of $\hat{f}^m_{\boldsymbol{v}}(\omega)$ also depends on the selected point $\boldsymbol{v}$. A common

[7] The interested reader can use the scripts in the `R`-package `localgaussSpec` to get access to these plots for the center and upper tail, cf. Section S6.1 for details.

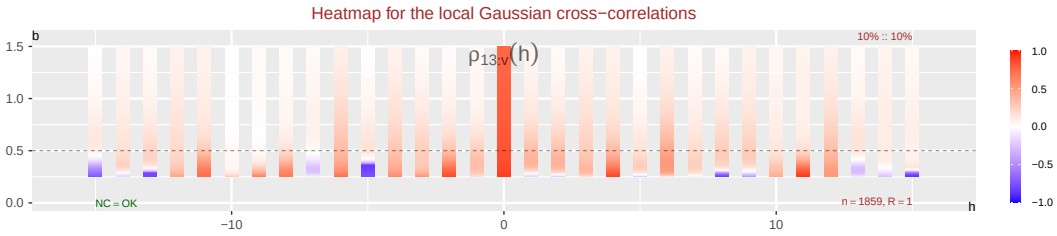

Figure S3.8. Heatmap of the local Gaussian cross-correlations $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ based on the DAX- and CAC-components of the `EuStockMarkets`-data, showing how these estimates change when the bandwidth $\boldsymbol{b}$ varies. The bandwidth $b = 0.6$ used in Figure 9 has been highlighted with a line.

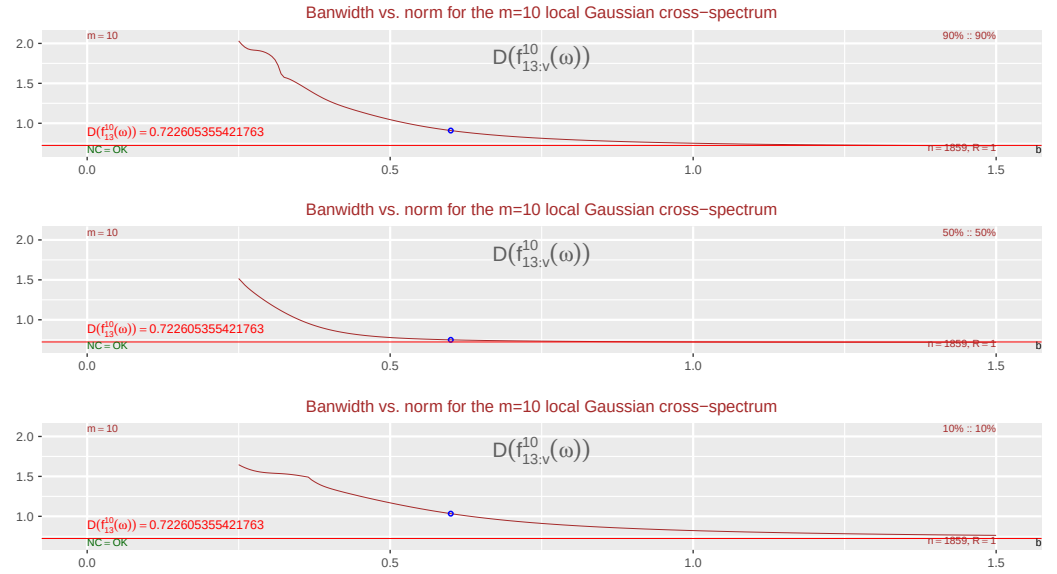

Figure S3.9. Distance-based plots based on the DAX- and CAC-components of the `EuStockMarkets`-data, that shows how the distance $D\left(\hat{f}^{10}_{k\ell:\boldsymbol{v}}(\omega)\right)$ varies with the bandwidth $\boldsymbol{b}$ for the three percentiles used in Figure 9, i.e., 10%, 50% and 90%. The bandwidth $b = 0.6$ used in Figure 9 has been highlighted with a point.

scale has been used for the three subplots in order to emphasise the asymmetry between the lower and upper tail.

The center plot of Figure S3.9 reveals that the "too low bandwidth problem" occurs a bit slower in a high density region, but it will even there eventually create a situation where the estimated local Gaussian autocorrelations degenerate towards either $+1$ or -1. This is the exact same behaviour that was observed in JT22 for the local Gaussian auto-spectra $f^m_{kk:\boldsymbol{v}}(\omega)$, and this is as expected since the higher concentration of (pseudo-normalised) points at the center implies that the degeneration of the estimates first will occur at shorter bandwidths.

The heatmap and distance-based plots in Figures S3.6 to S3.9 can detect the clearly undesirable regions for the bandwidth $\boldsymbol{b}$, but they do not reveal what the "just right" value for the bandwidth should be. Nevertheless, it is still possible to gain some insight into how sensitive the estimate of $f^m_{k\ell:\boldsymbol{v}}(\omega)$ will be for minor variations of the bandwidth $\boldsymbol{b}$, and that can be useful with regard to the selection of a few bandwidths that can be used when e.g. a bootstrap-investigation is to be performed.

The framework used in the `R`-package `localgaussSpec` ensures that it is trivial to compute and investigate/compare a wide range of bandwidths simultaneously, and the key idea is that knowledge of the local dependency structure can still be obtained even if the selected bandwidths are not spot on the "just right" value for the bandwidth.

S3.4. Sensitivity analysis: The truncation level m

This section is in essence the same as JT22 [Appendix D.4]. The main differences are primarily (1) that the notation have been updated in order to reflect that the target of the investigation now is the local Gaussian cross-spectrum $f^m_{k\ell:\boldsymbol{v}}(\omega)$ (instead of the local Gaussian auto-spectrum), and (2) that the sample under investigation in this case is the multivariate `EuStockMarkets`-data instead of the univariate `dmbp`-data.

The shape of $f^m_{k\ell:\boldsymbol{v}}(\omega)$ for a low truncation level can be different from the shape seen when a higher truncation level is used. It is thus of interest to investigate how sensitive the estimates $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ are to changes in the truncation level m.

This issue can easily be probed by performing an initial investigation with a high value for the maximum lag to be computed, since the computational cost is not too large when only a single sample (like the `EuStockMarkets`-data) is investigated. It did e.g. not take a long time to estimate $\rho_{k\ell:\boldsymbol{v}}(h)$ for $h = -200, \ldots, -1, 0.1, \ldots, 200$, which was needed for the construction of Figure 8 in the main document — and with these estimates it is trivial to compare $\hat{f}^m_{k\ell}(\omega)$ and $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ for m up to 200, since the integrated `shiny`-application of the `R`-package `localgaussSpec` can animate the changes that occur in the spectra when m grows from 0 to 200.

The computational costs can become rather large when it is necessary to find point-wise confidence intervals, since a high number of replicates then must be investigated with the same configuration of tuning parameters. It is then important to figure out a sufficient truncation level m, and restrict the attention to the estimates of $\rho_{k\ell:\boldsymbol{v}}(h)$ for $h = -m, \ldots, -1, 0, 1, \ldots, m$.

A drawback with the `shiny`-based approach in `localgaussSpec` is that it requires an inspection of many different plots. It could thus be of interest to also consider summary-plots that either use the distance function D from Definition S6, or some heatmap-based alternative visualisation of $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$, similar to those used for $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ in Figures S3.5 and S3.8.

Distance plots: It is possible to investigate the m-sensitivity by distance-based plots, but those plots are less useful in this case. One reason for this is that the norms $D\Big(\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)\Big)$ are monotonically increasing as functions of m. This can easily be seen by first recalling (cf. Algorithm 1(4)) that the estimates $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ are given by

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

and then keeping in mind that the lag-window function $\lambda_m(h)$ satisfies $\lambda_{m+1}(h) \geq \lambda_m(h)$. It follows that $D\Big(\hat{f}^{m+1}_{k\ell:\boldsymbol{v}}(\omega)\Big) \geq D\Big(\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)\Big)$, which does not provide any useful new information.

Instead of a plot of the norms $D\Big(\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)\Big)$, it is slightly more interesting to consider a plot that shows $D\Big(\hat{f}^{m+1}_{k\ell:\boldsymbol{v}}(\omega) - \hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)\Big)$, i.e., the distances between $\hat{f}^{m+1}_{k\ell:\boldsymbol{v}}(\omega)$ and $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ in the Hilbert space of Fourier series. This idea is shown in Figure S3.10 for the three diagonal points and 200 lags that was included in Figure 8. Note that Figure S3.10 takes into account the scaling due to the lag-window function $\lambda_m(h)$, and as such it does provide some new information compared to that contained in the plot showing the estimated local Gaussian cross-correlations.

The three subplots of Figure D.8 shows that $D\Big(\hat{f}^{m+1}_{k\ell:\boldsymbol{v}}(\omega) - \hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)\Big)$ rather quickly starts to decrease monotonically, which is as expected given the presence of the lag-window function $\lambda_m(h)$. This decrease implies that the effect of a change in the truncation level from m to $m+1$ becomes smaller as m grows, and the sensitivity is thus largest when m is small.

Figure D.8 might indicate that the $m = 10$ used in the main part is a bit too small. However, the purpose of that particular truncation level was simply to show that even a low truncation level could be used to detect the presence of nonlinear dependency structures in the time series under investigation, i.e., structures not detected by the ordinary spectrum.

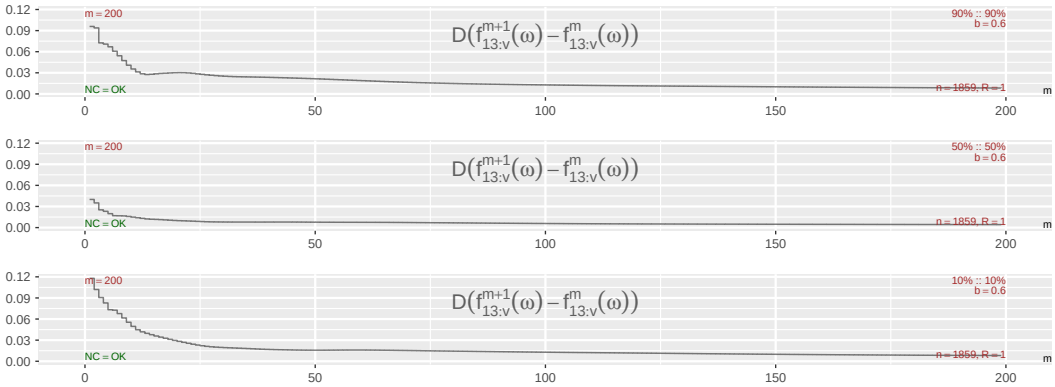

Figure S3.10. Distances between successive m-truncations of the local cross-spectra $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ between the DAX- and CAC-components of the `EuStockMarkets`-data.

It is natural to assume that two *successive* local Gaussian spectra $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ and $\hat{f}^{m+1}_{k\ell:\boldsymbol{v}}(\omega)$ should be similar in shape when m has grown a bit, but this does not imply that the *accumulated* changes to $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ are negligible. It is thus important to also inspect the frequency-dimension, and this can as mentioned above easily be done by the interactive `shiny`-application in the `localgaussSpec`-package.

Heatmap plots: The truncation level m is a discrete tuning parameter, and an inspection based on a heatmap-based approach could thus follow the setup used for the estimated $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$-values seen in Figures S3.5 and S3.8. The `R`-package `localgaussSpec` contains a script that can be used to create such a heatmap-based plot for $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$, with the frequencies ω along one axis and the truncation levels m along the other.

The resulting heatmap-based plot clearly showed that the peak seen in the `Co`-spectra part of Figure 9 at $\omega = 0$ (for $m = 10$ and a point either in the lower or upper tail) became even more dominating as m increased, and the peak dominated to such an extent that the heatmap-based plot did not reveal anything about the other frequencies. This plot has thus not been included here, but the script is available in `localgaussSpec`, cf. Section S6.1 for details.

S4. How to select the tuning parameters?

The sensitivity analysis in Section S3 considered the effect of minor changes to the tuning parameters $\boldsymbol{b}$ and m, and it did also discuss the sensitivity of $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ that is due to the position of the point $\boldsymbol{v}$ — which is of interest to know when a given sample/model is to be investigated. But what should the tuning parameters be for an actual investigation?

This topic was discussed in JT22 [Appendix E] for the local Gaussian auto-spectra $\hat{f}^m_{kk:\boldsymbol{v}}(\omega)$, and this discussion extends without any changes to the present multivariate case. The proposed strategy has been summarised below.

First of all, the `R`-package `localgaussSpec` can compute $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ for a wide range of tuning parameters, and for a large selection of different points $\boldsymbol{v}$. The computation will then also provide estimates of the univariate components, and all the relevant plots (of local Gaussian auto- and cross-correlations, and local Gaussian auto- and cross-spectra) can be investigated interactively by the `shiny`-application that is a part of `localgaussSpec`.

The computational cost for a single estimate $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ is normally not too large (assuming the truncation level m is not very high), but the total time can become quite large when a wide range of points $\boldsymbol{v}$ and bandwidths $\boldsymbol{b}$ are to be investigated at the same time. The creation of pointwise confidence intervals might on top of this require a large number of replicates R (or bootstrapped samples).

An investigation can start by looking at a single sample (real data, or a sample generated by a parametric model). It is reasonable to first look at a coarse grid based on a low number of points $\boldsymbol{v}$. The bandwidths $\boldsymbol{b}$ should also be restricted to only a few values in this

part. The truncation level m can be rather low (range 10-30), since it seems to be the case that interesting features can be detected even for this range of truncation levels.

If the initial investigation detects the presence of non-Gaussian dependency structures in the sample, then the next step is to select a (reasonable) bandwidth $\boldsymbol{b}$ and a (not too high) truncation level m, and then create heatmap- and distance-based plots of the estimates $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ for a wide range of points $\boldsymbol{v}$.

Select a collection of points $\boldsymbol{v}$ that it could be of particular interest to investigate further, and then investigate e.g. $R = 100$ replicates (or bootstrap samples) in order to produce pointwise confidence intervals for the resulting estimates of the local Gaussian auto- and cross-spectra. (See the discussion in Section S5 for details related to the bootstrapping.)

Some comments regarding the bandwidth $\boldsymbol{b}$: The discussion about the $\boldsymbol{b}$-sensitivity of the estimated local Gaussian cross-spectra $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ in Section S3.3, see Figures S3.6 to S3.9, indicate that the value used for the bandwidth $\boldsymbol{b}$ might not be that critical, unless it is "too short" or "too long". Insofar the purpose of the investigation is the detection of non-Gaussian dependency structures, it seems to be the case that a reasonable range of bandwidths will return the same conclusion. This implies that it instead of a "select the optimal bandwidth"-algorithm, should be sufficient to perform an investigation for a few bandwidths in order to see if any findings are consistent across those.

The interested reader can find a more in detail discussion related to the selection of the bandwidth $\boldsymbol{b}$ in JT22 [Appendix E.2], including a short discussion of some papers that have discussed the bandwidth selection for the estimation of the local Gaussian correlation $\rho_{\boldsymbol{v}}$.

S5. Regarding sampling and resampling

The section corresponds to JT22 [Appendix F], and the topic of interest is how sampling and resampling strategies can be used for different local Gaussian spectral investigations of a time series.

Section S5.1 focuses on the parametric bootstrap approach, where repeated independent simulations are done from given parametric models fitted to a given sample. It is here seen how the univariate local Gaussian sanity testing of parametric models from JT22 [Appendix F.2] can be extended to cover the multivariate case. This part includes an investigation of the DAX- and CAC-margins of the `EuStockMarkets` example, and then new plots are presented for the sanity testing of the cross-interaction of the multivariate parametric models fitted to the `EuStockMarkets` data.

The discussion in Sections S5.2 and S5.3 focuses on the resampling problem, which is critical for the production of the pointwise confidence intervals based on a given sample. Section S5.2 gives a summary of the discussion in JT22 [Appendix F] that motivated the introduction of the *Circular index-based block bootstrap for tuples*, with some additional comments about the role of the block length L. Finally, Section S5.3 presents a sensitivity analysis of the block length L for a multivariate example, which shows that this parameter have a minor effect on the resulting estimated pointwise confidence intervals.

S5.1. Parametric bootstrap and local sanity-testing of models

The discussion in JT22 [Appendix F.2] (motivated by a similar approach in Birr et al. (2019)) explained how a parametric bootstrap approach could be used to investigate models fitted to real data. The key idea is that a parametric model first is fitted to the original sample, and samples from that particular fitted model is thereafter used to produce the local Gaussian estimates needed for the comparison with the actual sample under investigation. This enables a local sanity-test (based on the m-truncated local Gaussian spectra) of the fitted model, since it can be used to identify if there are any points/frequencies with a clear mismatch between the local structures detected in the original sample and those seen in samples from the fitted model.

A key restriction in the above mentioned comparison is that the number of observations simulated from the fitted model should equal the number of observations in the sample. It is of course also possible to compare different fitted models against each other

directly, in which case the length of the simulated samples can be selected freely. Such an approach might be useful in order to check if different model-fitting approaches might have deviating local Gaussian dependency structures. The plots from the main part (and Section S3) can be used when comparing different fitted models.

It might also be of interest to simulate samples of different lengths from one particular known parametric model, and then fit different parametric models to those samples. The plots discussed in Sections S5.1.2 and S5.1.3 might be helpful for such an investigation.

S5.1.1. The example under consideration

The univariate example in JT22 [Appendix F.2] used the `dmbp`-data as the known sample. The fitted model was a GARCH-type model, more precisely an *asymmetric power ARCH-model* (apARCH) of order $(2,3)$, with parameters based on a fitting to the `dmbp`-data.[8] The univariate approach will now be extended to the multivariate case, with a repetition of the setup from the univariate case in Section S5.1.2 (since any investigation of a multivariate parametric model fitted to a multivariate sample should include a comparison of the univariate components).

The multivariate example considered in this paper uses the `EuStockMarkets`-data as the known sample, and in particular the restriction to the bivariate subset based on the DAX- and CAC-components (see Figures S5.1 to S5.6). As described in Section 3.2.2, the fitted parametric model in this case is a basic multivariate copula GARCH-model — which was selected only in order to provide a proof of concept. From this starting point it is natural to expect that the present *local Gaussian sanity investigation* should reveal that this particular multivariate copula GARCH-model does not properly capture the local Gaussian structures seen in the original sample.

S5.1.2. The univariate marginals

Figures S5.1 and S5.2 consider respectively the univariate DAX- and CAC-components of the `EuStockMarkets`-data, and these use the comparison of fitted models and sample that was introduced in JT22 [Figure F.1]. The description below is in essence the same as the one given in JT22 [Appendix F.2], and it explains how these plots are to be interpreted.

The key idea in Figures S5.1 and S5.2 is that estimates of the m-truncated local Gaussian auto-spectra $f^m_{kk:\boldsymbol{v}}(\omega)$, based on the original sample, can be superimposed on the plots based on parametric bootstrapping from the fitted model, and this makes it easy to compare them.

Nine different points $\boldsymbol{v} = (v_1, v_2)$ are considered in Figures S5.1 and S5.2, and these are based on the combinations that can be created when v_1 and v_2 varies over the 10%, 50% and 90% percentiles of the standard normal distribution. The corresponding plots are ordered in a grid in accordance with the position of these nine points in the plane, as can be seen by the information about $\boldsymbol{v}$ in the upper right corner of the respective plots.

The estimates $f^m_{kk:\boldsymbol{v}}(\omega)$ are always real-valued on the diagonal (since $\rho_{kk:\boldsymbol{v}}(-h) = \rho_{kk:\boldsymbol{v}}(h)$ when the point $\boldsymbol{v}$ is on the diagonal), and the three diagonal plots are thus marked with $f^m_{kk:\boldsymbol{v}}(\omega)$ in order to signal that the content is the local Gaussian autospectrum. The solid lines represents the estimates based on the original sample, whereas the dotted lines and the 95 % pointwise confidence intervals are based on samples from the fitted model.

The estimates $f^m_{kk:\boldsymbol{v}}(\omega)$ are complex-valued for the six off-diagonal points, and in this case the `R`-package `localgaussSpec` follows the convention used for the complex-valued cross-spectra, viz., $\mathrm{Co}\big(f^m_{kk:\boldsymbol{v}}(\omega)\big) = \mathrm{Re}\big(f^m_{kk:\boldsymbol{v}}(\omega)\big)$ and $\mathrm{Quad}\big(f^m_{kk:\boldsymbol{v}}(\omega)\big) = -\mathrm{Im}\big(f^m_{kk:\boldsymbol{v}}(\omega)\big)$.

The off-diagonal points are symmetric around the diagonal, i.e., both $\boldsymbol{v} = (v_1, v_2)$ and its diagonal reflection $\breve{\boldsymbol{v}} = (v_2, v_1)$ are present. The property $f_{kk:\boldsymbol{v}}(\omega) = \overline{f_{kk:\breve{\boldsymbol{v}}}(\omega)}$, implies that it is sufficient to plot $\mathrm{Co}\big(f^m_{kk:\boldsymbol{v}}(\omega)\big)$ on one side of the diagonal and $\mathrm{Quad}\big(f^m_{kk:\boldsymbol{v}}(\omega)\big)$ on the other side.

[8] The `R`-package `rugarch`, Ghalanos (2022) was used to find the parameters of several GARCH-models, and the asymmetric power ARCH model with the best fit was then selected. See JT22 [Section 3.4.3] for further details.

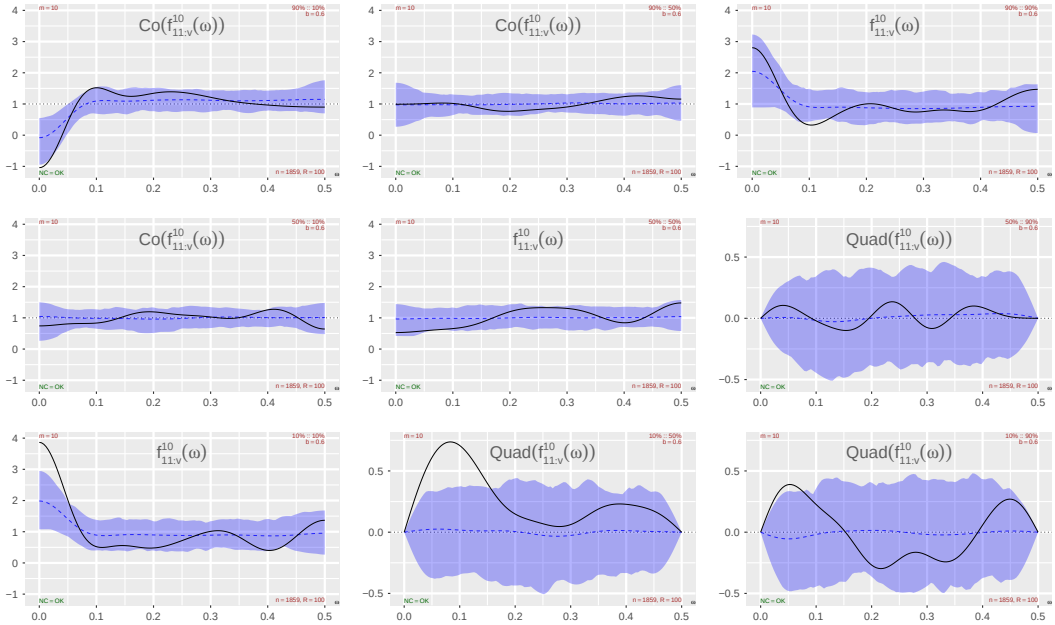

Figure S5.1. Marginal investigation: cGARCH-model fitted to `EuStockMarkets`, the part corresponding to the DAX-component.

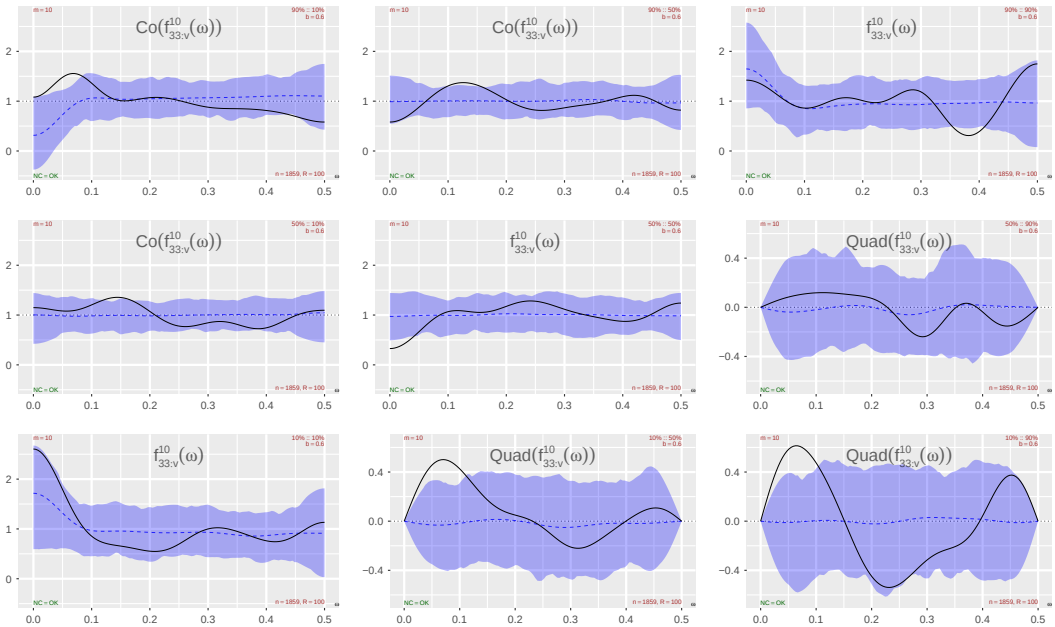

Figure S5.2. Marginal investigation: cGARCH-model fitted to `EuStockMarkets`, the part corresponding to the CAC-component.

Finally, the same scale is used for all plots showing real values, whereas another scale is used for the plots related to the imaginary parts. This distinction is natural since the scale needed for the imaginary part can be much smaller, as can be seen in Figures S5.1 and S5.2.

It was in JT22 [Appendix F.2] seen that the local Gaussian dependency structure of the fitted apARCH$(2,3)$-model agreed quite well with the one observed in the `dmbp`-data. This did not come as a surprise, since the selected apARCH$(2,3)$-model in that case was obtained after a testing procedure that tried out several thousand different variations of the GARCH-type models implemented in the `rugarch`-package.

The situation in this paper is completely contrary to the one in JT22, since a default model from the `rmgarch`-package, see Ghalanos (2022), was fitted to the `EuStockMarkets`-data in order to simply provide a proof-of-principle parametric model for how this kind of investigation can be performed. It is thus natural to expect that an investigation of Figures S5.1 and S5.2 would indicate that the local Gaussian dependency structures of the fitted marginal models do not properly match those seen in the DAX- and CAC-components of the `EuStockMarkets`-data — and this is indeed what is observed when the dashed and solid lines in Figures S5.1 and S5.2 are compared against each other.

The plots related to the real parts in Figures S5.1 and S5.2 show that there is a mismatch for the lower tail for low frequencies, in particular for the DAX-component as seen in Figure S5.1. The plots related to the imaginary parts also indicate that the fitted models do not properly reflect the dependency structure from the DAX- and CAC-components — but it should here be noted that different scales are used for the two groups of plots (real-valued versus complex valued-parts), and it is more critical when a deviation is observed within the plots having the dominating scale.

The observed mismatch between the fitted marginal models and the DAX- and CAC-components of the `EuStockMarkets`-data seen in Figures S5.1 and S5.2 clearly indicates that the selected marginal models should be replaced. This is (as mentioned above) not a surprising revelation, since the present marginal models were selected simply in order to present a proof-of-principle for the methods discussed in this paper.

S5.1.3. The multivariate interdependency structure

This section provides an extension to the multivariate case of the univariate local Gaussian sanity testing seen in Section S5.1.2. The target of interest in this case is the cross-temporal interdependency structure of the parametric model that has been fitted to a given multivariate sample. The idea is again to use visual inspections to see if clear deviations can be seen when the local Gaussian cross-spectra based on the fitted model are compared with the local Gaussian cross-spectra from the original sample.

The univariate investigation in Figures S5.1 and S5.2 focused on a grid of nine points $\boldsymbol{v} = (v_1, v_2)$, where v_1 and v_2 varied over the 10%, 50% and 90% percentiles of the standard normal distribution — and the multivariate investigation will use the exact same grid (other points can of course be used if it is of interest to focus on a particular region).

The "diagonal folding" seen in Figures S5.1 and S5.2 can not be used for the investigation of the cross-temporal case, since the local Gaussian cross-spectrum $f^m_{k\ell:\boldsymbol{v}}(\omega)$ are complex valued even for the diagonal points. This implies that it, similarly to the situation encountered for the heatmap- and distance plots (see discussion in Section S3.2), can be natural to include pair of plots that either focus on the *Cartesian decomposition* or the *polar decomposition* of the complex-valued local Gaussian cross-spectrum $f^m_{k\ell:\boldsymbol{v}}(\omega)$.

A visual investigation based on the *Cartesian decomposition* are given in Figures S5.3 and S5.4, which respectively present the `Co`- and `Quad`-spectra components of $f^m_{k\ell:\boldsymbol{v}}(\omega)$. These figures gathers local Gaussian spectra from the original data (shown with solid lines) and local Gaussian spectra based on simulations from the fitted models (shown with dashed lines and pointwise confidence intervals). The same internal scale is used for all the subplots, since that simplifies the task of detecting different behaviour between the points.

The scales for the `Co`-plots in Figure S5.3 are much larger than those for the `Quad`-plots in Figure S5.4, and it is thus natural to focus on the `Quad`-plots in this case. An inspection

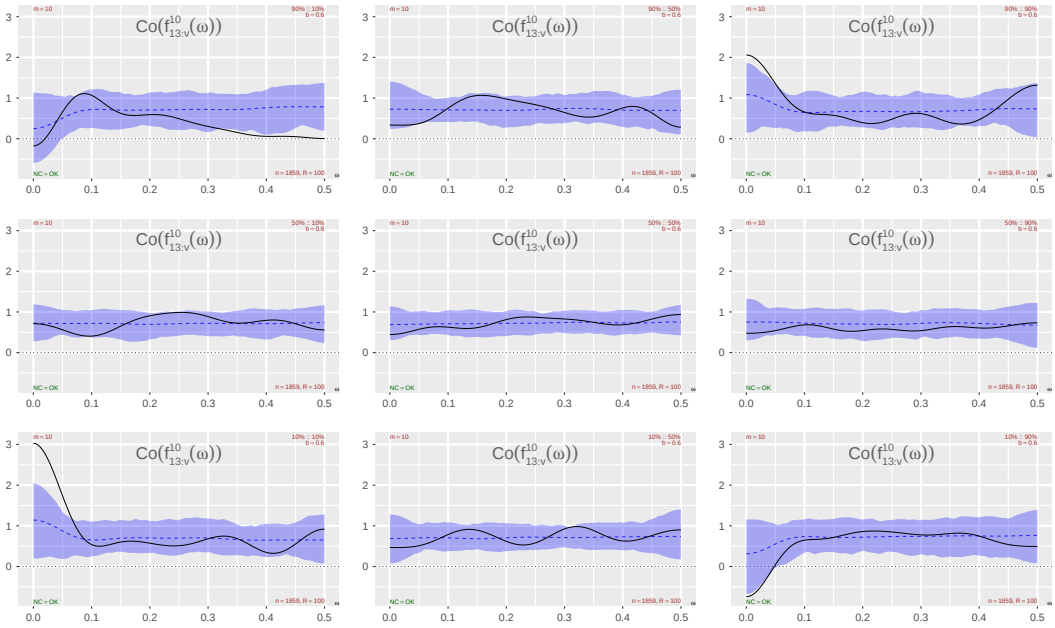

Figure S5.3. `Co`-spectra based local Gaussian investigation of the cGARCH-model fitted to the `EuStockMarkets`-data. Focus on the interaction between the DAX- and CAC-components of `EuStockMarkets`

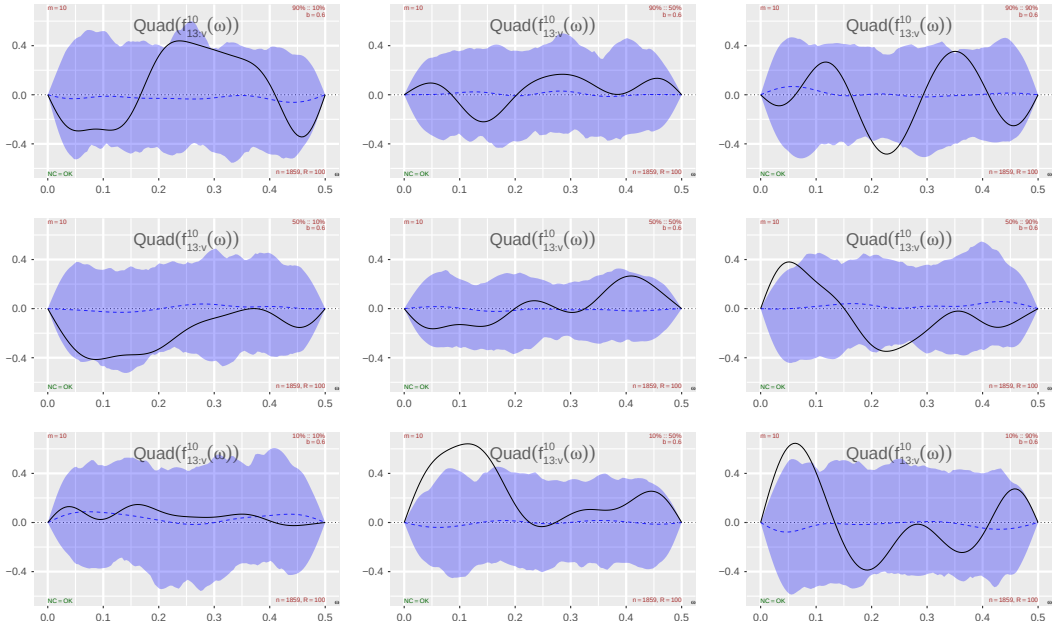

Figure S5.4. `Quad`-spectra based local Gaussian investigation of the cGARCH-model fitted to the `EuStockMarkets`-data. Focus on the interaction between the DAX- and CAC-components of `EuStockMarkets`

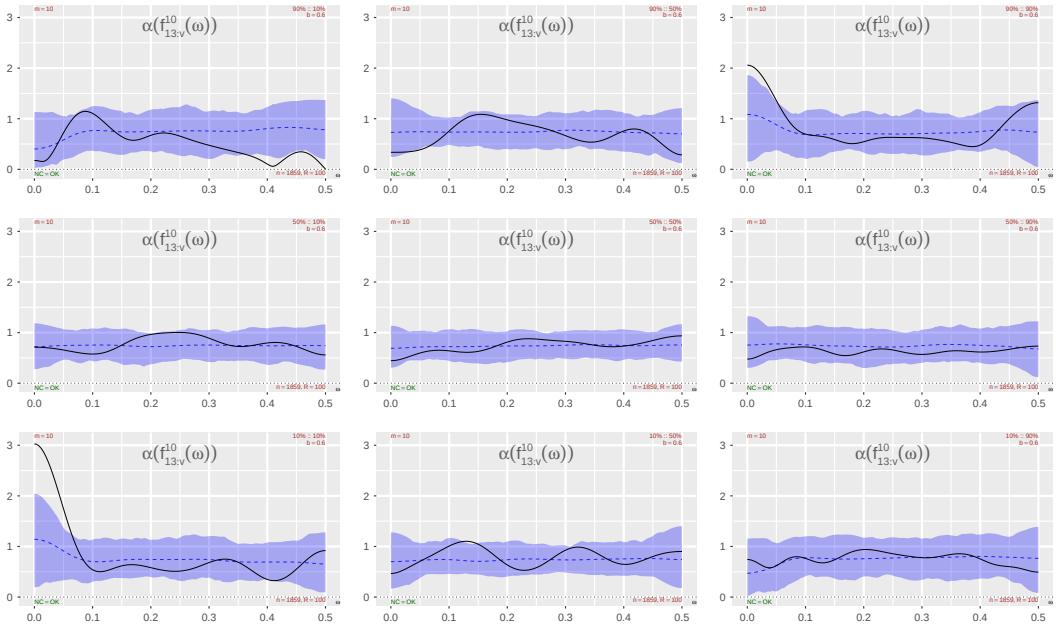

Figure S5.5. `Amplitude`-spectra based local Gaussian investigation of the cGARCH-model fitted to the `EuStockMarkets`-data. Focus on the interaction between the DAX- and CAC-components of `EuStockMarkets`

of Figure S5.3, in particular the points in the tails, shows that the local Gaussian cross-spectra based on the fitted parametric model deviates quite a bit from those computed from the original `EuStockMarkets`-sample. This is not surprising at all, since the most basic parametric model was used simply in order to provide a proof-of-concept for the present investigation.

The *polar decomposition* of the complex-values local Gaussian cross-spectra $f^m_{k\ell:\boldsymbol{v}}(\omega)$ into `Amplitude`- and `Phase`-spectra are given in Figures S5.5 and S5.6. The `Amplitude`-plots in Figure S5.5 are quite similar to the `Co`-plots seen in Figure S5.3, which is as expected since the scales of the `Co`-plots are much larger than those for the `Quad`-plots. For this particular example it follows that the same conclusions can be drawn from Figures S5.3 and S5.5, i.e. that it would be preferable to look for a better model than the trivial one used as a proof-of-concept example in this paper.

Note that the similarity between Figures S5.3 and S5.5 is due to the real part of $f^m_{k\ell:\boldsymbol{v}}(\omega)$ are the dominating component, and it could be that would not always be the case. Furthermore, the `Phase`-spectra seen in Figure S5.6 can be quite hard to interpret if the observations of $f^m_{k\ell:\boldsymbol{v}}(\omega)$ lies in the third and fourth quadrant of the complex plane. Due to this it seems natural to consider a `Co`- and `Quad`- based investigation of $f^m_{k\ell:\boldsymbol{v}}(\omega)$ instead of an `Amplitude`- an `Phase`-investigation when a local Gaussian sanity testing is performed for a parametric model of the the cross-temporal interdependency structure for a given multivariate sample.

S5.2. Nonparametric bootstrapping techniques

The local Gaussian estimates $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$, with pointwise confidence intervals, are straightforward to construct when a given parametric model is investigated — simply generate a sufficiently high number of samples from the model, compute $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ for each sample, and then (for each frequency ω) use suitable quantiles (like, e.g., the 5% and 95%) to construct the pointwise confidence intervals.

Pointwise confidence intervals are also of interest when the local Gaussian estimates $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ are based on a given sample (like the `EuStockMarkets`-data), and such confidence intervals can, e.g., be seen in Figure 9 in the main part. These intervals are again (for each

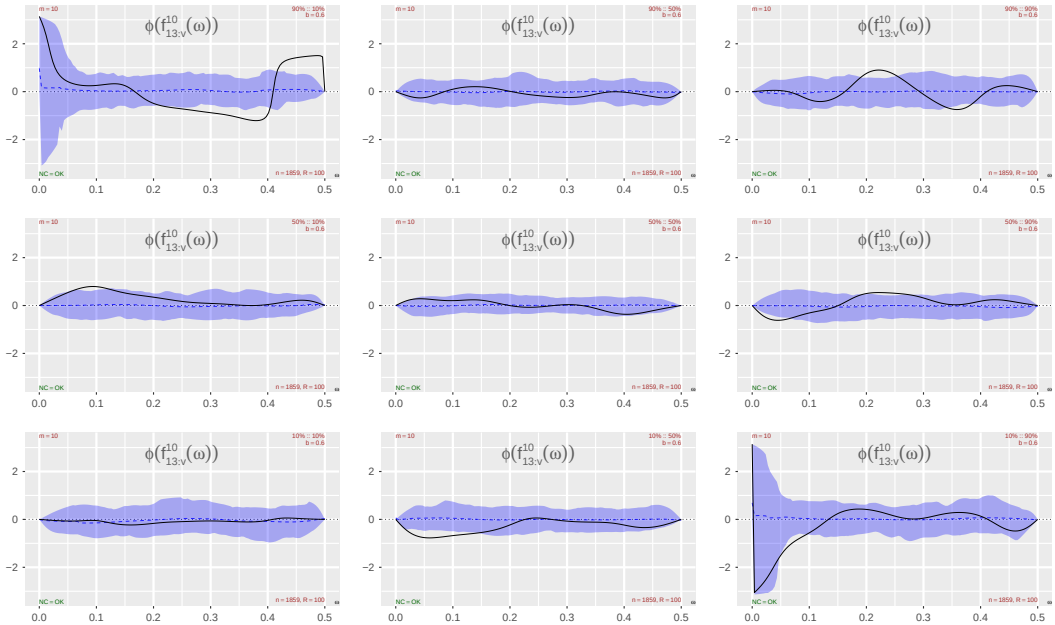

Figure S5.6. `Phase`-spectra based local Gaussian investigation of the cGARCH-model fitted to the `EuStockMarkets`-data. Focus on the interaction between the DAX- and CAC-components of `EuStockMarkets`

frequency ω) based on quantiles of $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ from a collection of samples, but these samples must now be constructed by a suitable resampling strategy.

The initial resampling strategy was the block bootstrap from Künsch (1989). The block bootstrap can (as explained below) be applied for the estimators underlying the local Gaussian estimates $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$, but it was early on observed that it was necessary with rather heuristic arguments regarding the selection of the block length argument L for the time series samples of interest (cases where Y_t and Y_{t+h} are dependent but not correlated).

Comments received during the review-process of JT22 initiated an investigation of how estimates based on the block bootstrap method can suffer from edge-effect noise when it is used on smaller sample sizes — and this resulted in the development of the *circular index-based block bootstrap for tuples* resampling strategy in JT22 [Algorithm F.4]

A summary of this story will be given below. The interested reader should consult JT22 [Appendices F.3 and F.4] for an in depth discussion of different bootstrapping methods, with emphasis on the problems that occur due to edge-effect noise. The revised resampling algorithm is based on a tweaking of the block bootstrap, and it thus inherits the block length argument L. Note that the block length L plays a different role in this algorithm, and that the sensitivity analysis (based on the `EuStockMarkets`-data) in Section S5.3 shows that the resulting pointwise confidence intervals for $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ are quite stable.

S5.2.1. Summary of JT22 [Appendix F.3]

The discussion in JT22 [Appendix F.3] started out with an explanation of why resampling based on the block bootstrap should be an acceptable strategy for the estimators needed for the computation of m-truncated estimates of the local Gaussian spectra $f^m_{k\ell:\boldsymbol{v}}(\omega)$, then it addressed the problem that edge-effects between the blocks could potentially create serious problems for the estimates $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ when used on the "short" `dmbp`-sample (length 1974) that was investigated in JT22. Potential solutions to these problems were addressed, including the block-of-blocks bootstrap (where edge-effects do not occur), but these solutions did not provide an optimal approach for a statistic computed by means of an algorithm that contains a kernel function.

Justification for the block bootstrap: The estimates $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ of the local Gaussian spectral densities are based on linear combinations of the estimates $\{\hat{\rho}_{k\ell:\boldsymbol{v}}(h)\}_{h=-m}^{m}$ of the

local Gaussian cross-correlations. These are all estimated by a local likelihood approach — and the asymptotic properties of the estimates $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ were developed in JT22 using the procedure from Klimko and Nelson (1978). A statistic obtained from the Klimko-Nelson procedure was explicitly mentioned by Künsch as an example for which the block bootstrap method would be applicable, cf. Künsch (1989, Example 2.4, p. 1219-20), and a resampling based on the block bootstrap was thus initially used for the construction of the pointwise confidence intervals for the `dmbp`-example seen in JT22 [Figure 9].

Edge-effect noise: The blocks in the block bootstrap algorithm will introduce noise into the estimation procedure, since sequences will occur in the resampled data that was not present in the original data. The following paragraphs from JT22 [Appendix F.3.2] introduce the "corrupt tuple"-concept and some notation needed later on.

> For example, if a time series $\{Y_t\}_{t=1}^n$ of length n is given, then an estimate of $\rho_{\boldsymbol{v}}(h)$ will be based on the bivariate set $\mathcal{Y}_h := \{(Y_{t+h}, Y_t)\}_{t=1}^{n-h}$ of size $n-h$. When the block bootstrap is used with some block length L, then there will be a resampled sequence $\{Y_t^*\}_{t=1}^n$ and the idea is that an estimate of $\rho_{\boldsymbol{v}}(h)$ now should be computed based on the bivariate set $\mathcal{Y}^*_{h:L} := \{(Y^*_{t+h}, Y^*_t)\}_{t=1}^{n-h}$.
>
> However, the set $\mathcal{Y}^*_{h:L}$ will contain *corrupt tuples* that do not exist in $\mathcal{Y}_h$, i.e., the first and second component of (Y^*_{t+h}, Y^*_t) can belong to different blocks, and this will add a bit of *edge-effect noise* into the estimation process. The edge-effect noise is negligible in the asymptotic situation (very large sample sizes n and large block lengths L), but it can make an impact when smaller samples are investigated.

The example used in JT22 was the `dmbp`-data ($n = 1974$ unique observations), and the plots of the estimated local Gaussian autospectra $\hat{f}^m_{\boldsymbol{v}}(\omega)$ used the truncation level $m = 10$. It was thus of particular interest to compute (for different block lengths L) the expected amount of *corrupt tuples* when estimating the local Gaussian autocorrelations $\rho_{\boldsymbol{v}}(h)$ for $h = 1, \ldots, 10$ — and a copy of the resulting table from JT22 is given in Table S1.

$L \setminus h$	1	2	3	4	5	6	7	8	9	10
25	4.0%	7.9%	11.9%	15.8%	19.8%	23.8%	27.8%	31.7%	35.7%	39.7%
100	1.0%	1.9%	2.9%	3.9%	4.8%	5.8%	6.8%	7.7%	8.7%	9.7%

Table S1. Copy of JT22 [Table 2]. The expected fraction of corrupt tuples when $\rho_{\boldsymbol{v}}(h)$ are estimated from block bootstrap replicates of the `dmbp`-data ($n = 1974$), when $L \in \{25, 100\}$ and $h \in \{1, \ldots, 10\}$.

Table S1 shows that the amount of noise due to edge-effects can be severe when the block bootstrap is used with a too short block length, and this was also seen when attempts were made at making pointwise confidence intervals for $f^m_{\boldsymbol{v}}(\omega)$ based on this. The resulting "confidence intervals" for small block lengths L did in some cases (for low frequencies ω) not even contain the actual estimate $\hat{f}^m_{\boldsymbol{v}}(\omega)$ based on the sample under investigation.

The observations in Table S1 indicated that an adjusted resampling technique was needed, preferably one that completely (or at least partially) removed the corrupt tuples from the estimation algorithm. Two different approaches that completely avoids the corrupt tuples was discussed in JT22, but they had some issues that made them less interesting to implement — as briefly explained below.

Resampling from the tuples in $\mathcal{Y}_h := \{(Y_{t+h}, Y_t)\}_{t=1}^{n-h}$, for $h = 1$ to $h = m$, was considered in JT22 [Appendix F.3.3] as one option for finding an estimate $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ of the m truncated local Gaussian autospectrum $f^m_{k\ell:\boldsymbol{v}}(\omega)$. The resampling must be done in a manner that properly takes into account the time-connection between the bivariate pairs (Y_{t+h}, Y_t) for different h-values, but it is easy to find an algorithm that does this for a selected truncation level m: Simply consider the bivariate tuples in $\mathcal{Y}_h$ as projections from the $(m+1)$-variate tuples $\mathcal{Y}_{m:0} = \{(Y_{t+m}, \ldots, Y_{t+1}, Y_t)\}_{t=1}^{n-m}$, and start out by sampling the desired number of elements from $\mathcal{Y}_{m:0}$. A shortcoming with this approach is that the resulting estimates $\hat{\rho}_{\boldsymbol{v}}(h)$ will depend on the selected m-value — and this dependency on the truncation level m implies that all the local Gaussian autocorrelations must be

recomputed if the initial truncation level m is changed to $m+1$, and this is computational costly and not desirable to implement.

The block-of-blocks bootstrap introduced in Politis and Romano (1992) was discussed in JT22 [Appendix F.3.4]. This method completely avoids the edge-effect issue that occurs for the block bootstrap, since the statistic of interest are computed on the individual blocks. However, as mentioned in JT22:

> This restriction to individual blocks can be an excellent idea for many statistics of interest, but it is a somewhat questionable approach for the estimates $\hat{\rho}_{\boldsymbol{v}}(h)$ of the local Gaussian autocorrelations. The reason for this is that the bandwidth argument $\boldsymbol{b}$ in the kernel function $K_{\boldsymbol{b}}(\boldsymbol{w}-\boldsymbol{v})$ must be much larger if the estimation algorithm is to be used on only a subset of the observations — and the local structures of interest might then not be detected at all.
>
> It would of course be of interest to implement the block-of-block bootstrap for the estimates of the local Gaussian spectra if very large samples are encountered, i.e., when the individual blocks contains several thousand consecutive observations — but for shorter samples (like the `dmbp`-example) it seems better to use something else.

S5.2.2. Summary of JT22 [Appendix F.4]

The *circular index-based block bootstrap for tuples* was introduced in JT22 [Appendix F.4]. This is a minor adjustment of the ordinary block bootstrap, and it inherits the block length argument L as a tuning parameter (see Section S5.3 for a sensitivity analysis). The key motivation for the revised approach is that an estimate $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ of the m-truncated local Gaussian autospectrum $f^m_{k\ell:\boldsymbol{v}}(\omega)$ should be based on m "linked samples" from $\mathcal{Y}_h := \left\{\left(Y_{t+h}, Y_t\right)\right\}_{t=1}^{n-h}$, for $h=1$ to $h=m$, and that these "linked samples" should be used for the estimation of the local Gaussian autocorrelations $\rho_{\boldsymbol{v}}(h)$.

This approach avoids the computational cost problem that made the method mentioned in JT22 [Appendix F.3.3]) less interesting to implement, i.e., the estimates $\hat{\rho}_{\boldsymbol{v}}(h)$ do not depend on the truncation level m — but the price to pay for this is that some "corrupt tuples" will still be present, and the issue with noise due to edge-effects does not completely disappear. However, the expected amount of this noise for the revised resampling strategy (see Table S2 on page 31) is reduced a lot from the one seen for the resampling based on the block bootstrap (see Table S1 on page 29).

The following toy-example are copied from JT22 in order to explain the underlying motivation for the *circular index-based block bootstrap for tuples* resampling algorithm, i.e., that it is the indices of the observations that are the key detail of interest to keep track of.

> It will be a bit easier to digest the definitions and the algorithm that are given later on in this section, if a simple toy-example is investigated first: Consider a situation with a time series having five unique observations Y_1, Y_2, Y_3, Y_4, Y_5 and assume that there is an interest for an estimate based on the four lag-1 tuples in $\mathcal{Y}_1 = \left\{\left(Y_{t+1}, Y_t\right)\right\}_{t=1}^{4}$. If a block bootstrap with block length $L=2$ is used, the resampled time series might e.g. look like $Y_1^* = Y_4, Y_2^* = Y_5, Y_3^* = Y_3, Y_4^* = Y_4, Y_5^* = Y_2$, and the corresponding set of lag-1 tuples would be $\mathcal{Y}^*_{1:2} = \left\{\left(Y^*_{t+1}, Y^*_t\right)\right\}_{t=1}^{4}$. It is easy to see that $\mathcal{Y}^*_{1:2}$ in this case will contain the two corrupt tuples $\left(Y_3, Y_5\right)$ and $\left(Y_2, Y_4\right)$, i.e., tuples that are not present in $\mathcal{Y}_1$.
>
> The key idea in the adjusted algorithm is to move the focus to the indices of the original sample, i.e., $1,2,3,4,5$, and then use the block bootstrap to sample from these. The resampled set of indices for the example above would be $4,5,3,4,2$, and from these it is possible to construct the *cyclically* $h=1$ *shifted* set of indices $5,1,4,5,3$. The method is simply to add the lag $h=1$ to all the resampled indices — and to start back on 1 if a value exceeds $n=5$. The four desired lag-1 tuples $\mathcal{Y}^\sharp_{1:2} = \left\{\left(Y^\sharp_{t+1}, Y^\sharp_t\right)\right\}_{t=1}^{4}$ are now created by using the resampled set of

indices in the $Y_t^\sharp$-component, whereas the cyclically $h = 1$ shifted indices are used for the $Y_{t+h}^\sharp$-component. This results in the following four tuples, $\mathcal{Y}_{1:2}^\sharp = \{(Y_5, Y_4), (Y_1, Y_5), (Y_4, Y_3), (Y_5, Y_4)\}$, and it is easy to see that the only corrupt tuple in $\mathcal{Y}_{1:2}^\sharp$ is (Y_1, Y_5). Note: It could in principle now also be added a fifth tuple (Y_3, Y_2) to $\mathcal{Y}_{1:2}^\sharp$, but that is not of interest since there are only four tuples in $\mathcal{Y}_1$.

The adjusted resampling algorithm is, as seen above, based on a simple tweaking of the ordinary block bootstrap. The formal definition of the algorithm requires some rather technical steps, and these will not be repeated in this paper. The interested reader can consult JT22 [Appendix F.4] for further details, see in particular JT22 [Definitions F.1 to F.3 and algorithm F.4].

Note that the resampling algorithm from JT22 [Algorithm F.4] by construction will return the same results as those obtained from the ordinary block bootstrap when the sample size n and the block length L are large. The situation is different for smaller sample sizes, since the adjusted approach then will remove the majority of the corrupt tuples that adds edge-effect noise into the estimation of the local Gaussian autocorrelations $\rho_v(h)$. This can be seen clearly by comparing the expected amount of corrupt tuples when the block bootstrap (see Table S1 on page 29) with the corresponding numbers for the revised resampling algorithm, as given in Table S2.

$L \setminus h$	1	2	3	4	5	6	7	8	9	10
25	0.002%	0.006%	0.012%	0.020%	0.030%	0.043%	0.057%	0.073%	0.092%	0.112%
100	0.001%	0.002%	0.003%	0.005%	0.008%	0.011%	0.014%	0.019%	0.023%	0.028%

Table S2. Copy of JT22 [Table 3]. The expected amount of corrupt tuples when $\rho_v(h)$ are estimated for the `dmbp`-data by the *circular index-based block bootstrap for tuples*, cf. JT22 [Algorithm F.4].

A comparison of the entries in Tables S1 and S2 reveals that it, for a time series of length $n = 1974$, is quite a drastic reduction in the number of corrupt tuples when the block bootstrap is replaced with the *circular index-based block bootstrap for tuples* resampling strategy in JT22 [Algorithm F.4]. From the $h = 10$ column it can be seen that the expected amount have been reduced from 39.7% to 0.112% when $L = 25$, and it has been a reduction from 9.7% to 0.028% when $L = 100$. The edge-effect noise are thus in this case rather negligible when the adjusted resampling strategy is used.

Table S2 indicates that the expected amount of corrupted tuples will be low even for short block lengths L. This implies that the pointwise confidence intervals for the m-truncated local Gaussian autocorrelations $f_v^m(\omega)$ should not be too sensitive to the block length L when the *circular index-based block bootstrap for tuples* resampling strategy is used.

S5.3. Sensitivity analysis: The block length L

The block length sensitivity for the adjusted resampling strategy from JT22 [Algorithm F.4] was investigated for two different cases in JT22 [Appendix F.5]. The first case was based on the `dmbp`-data, whereas the second case was based on resampling of a single sample generated by the local trigonometric model initially used for the sanity testing of the estimation algorithm. For each case the $m = 10$ truncated local Gaussian spectral density was estimated, and then pointwise confidence intervals were generated for each case based on resampling using the *circular index-based block bootstrap for tuples* strategy. The exact same results for both of the cases, i.e., the resulting pointwise confidence intervals hardly changed when the block length L increased from 10 to 69.

The topic of interest for the present paper is the extension of the local Gaussian spectral theory to the multivariate case, and the script used for the univariate `dmbp`-case in JT22 [Appendix F.5] was thus modified to cope with the multivariate `EuStockMarkets`-case — and the results are shown in Figures S5.7 and S5.8. Note that these figures are strikingly similar to those encountered in JT22 [Appendix F.5], and the discussion below is with some minor adjustments the same as the one found there.

Figure S5.7 contains two sequences of boxplots, indexed by the block length L which increases in steps of 1 from $L = 10$ to $L = 69$. The distance function D mentioned in

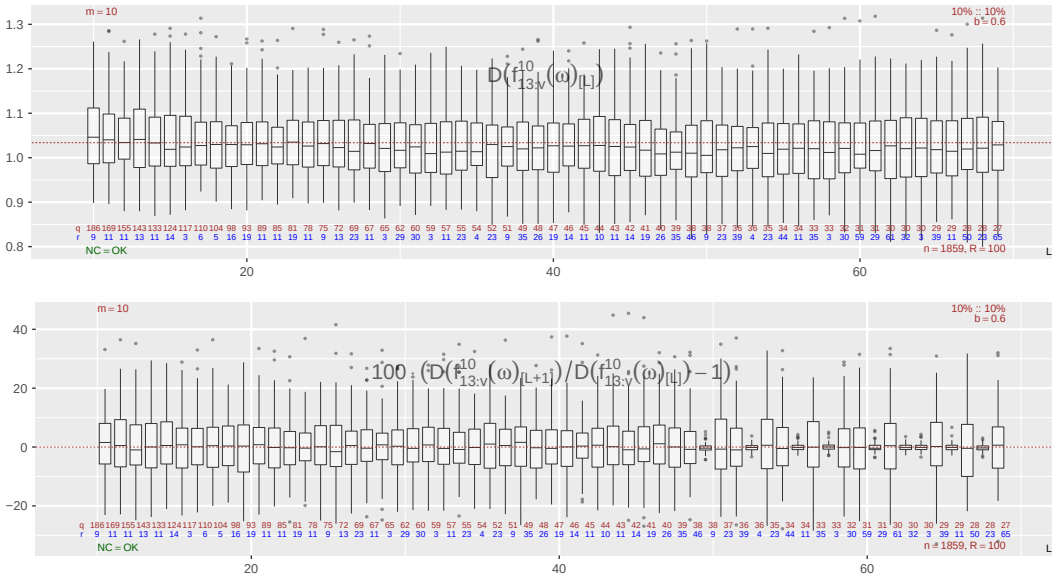

Figure S5.7. Distance based box-plots for the investigation of the sensitivity of the block length L for the adjusted resampling strategy from JT22 [Algorithm F.4]. The numbers in the two bottom rows show $q = \lceil n/L \rceil$ and $r = n - (q-1) \cdot L$, i.e., the number of blocks and the length of the last block.

Section S3.1 is used to generate the values the boxplots are based on. Keep in mind that the distance function does not reveal anything about the frequency-component of the cases under investigation, so it is also necessary to include a plot that focus on that aspect for a few of the block lengths L, as seen in Figure S5.8.

The panel at the top of Figure S5.7: A box-plot for the $D\Big(f^{10}_{k\ell:\boldsymbol{v}}(\omega)_{[L]}\Big)$-values (based on $R = 100$ replicates) is given for each block length L. A horizontal red dashed line has been added that shows the $D\big(f^{10}_{k\ell:\boldsymbol{v}}(\omega)\big)$-value for the original sample. It can be seen that the medians of the box-plots tend to be slightly smaller than the horizontal line that corresponds to the value based on the original sample, but these medians are based on $R = 100$ replicates — and another realisation might thus look slightly different. It does not seem to be any pattern here with regard to how these box-plots changes when L increases.

The panel at the bottom of Figure S5.7: These box-plots shows the percent-wise changes in the distances when the block length goes from L to $L+1$, and everything else is kept identical, i.e., $100 \cdot \Big(D\Big(f^{10}_{k\ell:\boldsymbol{v}}(\omega)_{[L+1]}\Big) / D\Big(f^{10}_{k\ell:\boldsymbol{v}}(\omega)_{[L]}\Big) - 1\Big)$. This is possible to do since the reproducibility setup enables a tracking for each individual realisation.

A horizontal red dashed line has been added at 0, and it is clear that the median-part of these box-plots are quite close to this horizontal line. It can also be observed that some of these box-plots are more compact than the other ones, and a simple investigation of the numbers given at the bottom of the plots reveals that this phenomenon occurs when an increase from L to $L+1$ does not reduce the number of blocks that are needed, i.e., they occur when $\lceil n/L \rceil = \lceil n/(L+1) \rceil$.

For the individual bootstrapped time series, this indicates that the changes are minimal when the number of blocks remains the same — whereas the changes are much larger when the increase of L triggers a reduction in the number of blocks. However, as is evident from an inspection of the panel at the top of Figure S5.7, this effect is only on the level of the individual replicates, and it is averaged away when a collection of replicates is considered.

Note that the effect noticed in the bottom panel of Figure S5.7 also is present for the global spectral densities (based on these bootstrapped samples), so this phenomenon is thus not an artefact of the way the local Gaussian spectral densities are estimated.

The frequency-component: Figure S5.7 indicates that the block length sensitivity, as measured by $D\big(f^{m}_{k\ell:\boldsymbol{v}}(\omega)\big)$, for the *circular index-based block bootstrap for tuples* resampling strategy from JT22 [Algorithm F.4] is rather small. But does this imply that these block

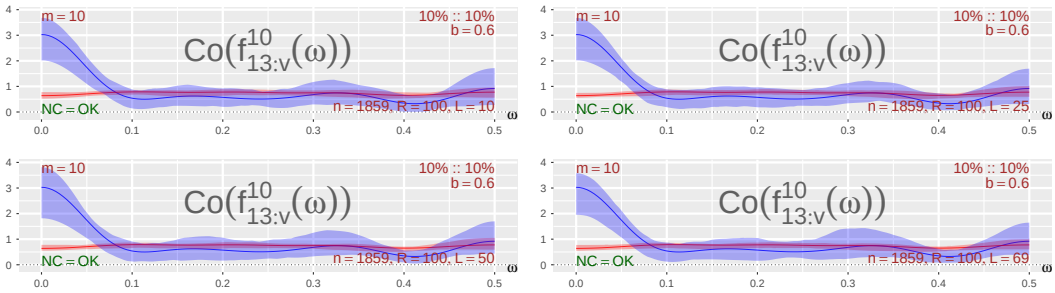

Figure S5.8. Four different block lengths L (from those investigated in Figure S5.7) have here been used in the *circular index-based block bootstrap for tuples* resampling strategy given in second part of Section S5. The values of L are 10, 25, 50 and 69, and this information is plotted at the lower right corner of the plots. The observed differences are tiny.

lengths should be considered *equally good* or *equally bad*? That can not be concluded from Figure S5.7 alone, and it is thus necessary to also consider a plot that takes the frequency-dimension into account. Figure S5.8 does this, using the `Co`-spectra component, for the four block lengths $L \in \{10, 25, 50, 69\}$.

It is clear from Figure S5.8 that the differences between these estimates are rather small, and it is necessary to look closely in order to see that the pointwise confidence intervals are slightly wider near $\omega = 0$ for the case $L = 50$.

Conclusion: These results, are as previously mentioned, in complete agreement with the corresponding results from the univariate case in JT22 [Appendix F.5], i.e., the effect of changing the block length L is minuscule. The interested reader can in JT22 [Appendix F.4.3] find an in depth discussion of why the block length only plays a role as a tuning parameter in the *circular index-based block bootstrap for tuples* resampling strategy.

Reproducibility: The scripts needed for the reproduction of the plots in this section are included in the `R`-package `localgaussSpec` (see Section S6.1 for further details), and the interested reader can there easily adjust the range of the block lengths to be used. It is also possible to adjust all the other tuning parameters needed for the estimation of $f^m_{k\ell:v}(\omega)$, and it is even possible to perform the computations with the ordinary block bootstrap if so should be desired.

S6. Scripts and details related to the examples

The reproducibility of all the examples in this paper can be done by the scripts contained in the `R`-package `localgaussSpec`, and Section S6.1 explains how the interested reader can extract these scripts. This section also contains some further details/discussions related to the examples in Section 3 in the main paper. Details related to the bivariate Gaussian test-example are given in Section S6.2, whereas Section S6.3 presents the R-code used to estimate the parameters used for the cGARCH-example seen in Figure 12.

Section S6.4 contains a short explanation of how animations of complex-valued plots can be used to investigate the estimated local Gaussian cross-spectra, a part that was moved here in order to improve the flow of the main part.

Section S6.5 investigates the *local bivariate trigonometric* example seen in Figures 2 to 5 of Section 3.1 in the main part. This is quite similar to the corresponding discussion for the univariate case in JT22 [Appendix G.4], it adjusts to the bivariate case the theoretical investigation of the general construction of which the *local trigonometric* examples are particular realisations, and it also briefly summarises the heuristic arguments from JT22 that was used to sanity test the implemented estimation algorithm.

Finally, once more mirroring the discussion in JT22 [Appendix G], the last part of Section S6.5 verifies that it for a large sample is possible to detect an elusive component that only occurs with probability 0.05 in the *local bivariate trigonometric* example, and it ends with some comments related to issues that can occur (under specific circumstances) when

the local Gaussian machinery is used on a time series whose global spectrum does not look like white noise.

S6.1. The scripts in the R-package `localgaussSpec`

All the examples in this paper (and all the examples in JT22) can be reproduced by the scripts in the R-package `localgaussSpec`. This R-package can be installed by using "`remotes::install_github("LAJordanger/localgaussSpec")`". The simplest way to extract the scripts from the internal storage of this R-package is to use the R-function "`LG_extract_scripts()`" after the package has been installed.

These scripts can either be used as they are (reproduction of the examples in this paper), or they can be used as templates for similar investigations of other samples/models that the user would like to investigate.

The reproduction of the figures requires two different scripts. The first scripts contain the code needed for the estimation of the local Gaussian auto-correlations $\rho_{kk:\boldsymbol{v}}(h)$ and cross-correlations $\rho_{k\ell:\boldsymbol{v}}(h)$ for all the specified combinations of the tuning parameters, whereas the second scripts contain the code that creates the particular visualisations seen in the figures in this paper. Note that it is sufficient to use the first type of scripts in order to use the integrated `shiny`-application that enables an easy interactive investigation of the resulting estimates. The second type of scripts is first needed when one wants to put many figures into one larger grid.

S6.2. The bivariate Gaussian white noise example

This section contains the details related to the bivariate Gaussian white noise example, used for the sanity testing of the estimation algorithm for the local Gaussian cross-spectra $f_{k\ell:\boldsymbol{v}}(\omega)$, that was referred to in the initial discussion of Section 3 in the main part.

It is natural to start with a Gaussian example, since it is the one case where it is explicitly known what the true value of the the local Gaussian cross-spectra $f_{k\ell:\boldsymbol{v}}(\omega)$ is, in particular, it is known from Lemma 1(1) that the local Gaussian cross-spectrum coincide with the ordinary cross-spectrum when the time series under investigation is Gaussian.

The plots in Figure S6.1 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 Figure S6.1 shows the situation for a point off the diagonal, whereas the right column 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[9] are the same for each row of Figure S6.1.

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 S6.1 shows that the red and blue dotted lines, that respectively represents the estimates of the global and local m-truncated spectra,[10] 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.

S6.3. The cGARCH-example in Figure 12

The cGARCH-example seen in Figure 12 (see also Figures 10 and 11) had coefficients that were fitted to the pseudo-normalised log-returns of the `EuStockMarkets`-data by the help of the `rmgarch`-package Ghalanos (2022). The "c" in cGARCH implies that a

[9] If you have a black and white copy of this paper, then read "red" as "dark" and "blue" as "light".

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

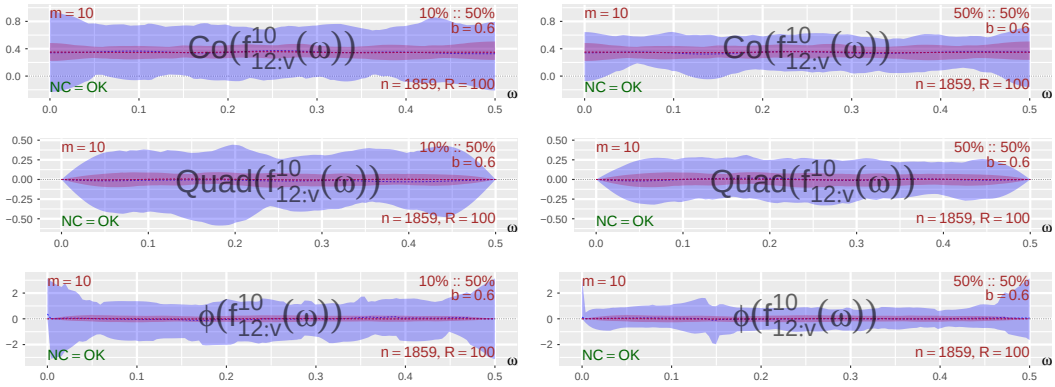

Figure S6.1. `Co`-, `Quad`- and `Phase`-plots for two points: Samples from the bivariate Gaussian white noise model, for which it is known that the local Gaussian spectra are equal to the ordinary spectra. The equality of the estimates are clear from this figure, the wider pointwise confidence intervals for the local Gaussian spectra is a consequence of the kernel-function in the estimation algorithm.

Copula-based approach is used for the modelling of the interdependency-structure, and the following steps can be used to fit a Copula-GARCH model to a multivariate sample:

1. For each univariate subset of the sample, create a `uGARCHspec`-object (univariate GARCH specification) by using the `ugarchspec`-function from the `rugarch`-package.
2. Create a `uGARCHmultispec`-object (univariate multiple GARCH specification) by using the `multispec`-function from the `rugarch`-package on the list of univariate GARCH specifications obtained in the first step.
3. Create a `cGARCHspec`-object (Copula-GARCH specification) by using the `cgarchspec`-function from the `rmgarch`-package on the `uGARCHmultispec`-object obtained in the second step.
4. Create a `cGARCHfit`-object (Copula-GARCH fit object) by using the `cgarchfit`-function from the `rmgarch`-package. The `spec`-argument should be the `cGARCHspec`-object from the third step, and the `data`-argument should be the multivariate sample under investigation.

The model and parameters stored in the `cGARCHfit`-object can now be used as the `fit`-argument of the `cgarchsim`-function from the `rmgarch`-package. This implies that samples from the fitted model can be generated, from which it is possible to find estimates and pointwise confidence intervals for the m-truncated local Gaussian auto- and cross-spectra for this particular model.

Note that the `ugarchspec`, `uGARCHmultispec` and `cgarchfit` all have many arguments, and tweaking of these arguments enables a plethora of different GARCH-type models to be investigated. The cGARCH-model used for the creation of Figure 12 used the default arguments in the previously mentioned functions, and this was done since the desired target was to produce a simple proof-of-concept model for the local Gaussian sanity testing presented in Section S5.1. It was no surprise that Figures S5.1 to S5.6 revealed that the local Gaussian spectra based on this "trivial" cGARCH-model had some clear deviations from the local Gaussian spectra based on the `EuStockMarkets` data.

The full list of default arguments for the `R`-functions `ugarchspec`, `uGARCHmultispec` and `cgarchfit` will not be included here, since that would require several pages of code and explanations. The interested reader can look these details up in the documentation of the `rugarch`- and `rmgarch`-packages.

A final comment: It is of course possible to follow the strategy used in JT22, where an apARCH(2,3)-model was fitted to the `dmbp`-data after a procedure that tested several thousand different variations of the GARCH-type models implemented in the `rugarch`-package. This strategy was not adopted in the present paper since the multivariate nature of the `EuStockMarkets`-data, and the amount of arguments to tweak in the functions `ugarchspec`,

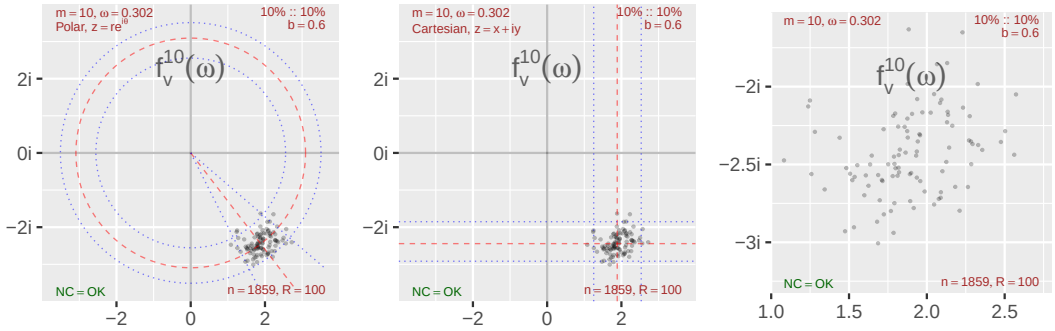

Figure S6.2. Complex-valued representation of 100 samples of $f_{12:v}(\omega)$ from Equation (20), at the peak frequency $\omega = 0.302$. Left panel: Pointwise 90% confidence bands based on polar representation. Center panel: Pointwise 90% confidence bands based on Cartesian representation. Right panel: Zoomed in plot.

`uGARCHmultispec` and `cgarchfit`, would have required a rather big computational investment for a sample that ended out being used simply due to it being available in R.

S6.4. Plots of the complex-valued spectra

This section contains a discussion related to animations of complex-valued plots, and how these might be used to investigate the estimated local Gaussian cross-spectra. This was initially contained in the main part, but it was then moved here in order to improve the flow of that part. Before reading on, recall the following from the main part:

> **The reference case:** The heuristic argument needed for the bivariate case is identical in structure to the one used in the univariate case, and for the present investigation the reference for the plots later on is based on the following simple bivariate model,
>
> $$Y_{1,t} = \cos(2\pi\alpha t + \phi) + w_{1,t} \text{ and } Y_{2,t} = \cos(2\pi\alpha t + \phi + \theta) + w_{2,t},$$
>
> where $w_{i,t}$ is Gaussian white noise with mean zero and standard deviation σ, with $w_{1,t}$ and $w_{2,t}$ independent, 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 Figure 1, 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.

It can be enlightening to compare the `Co`-, `Quad`- and `Phase`-plots in Figure 1 with a plot that shows the underlying estimates upon which the pointwise confidence intervals were based. Such a plot is shown in Figure S6.2, where the left panel presents the complex-valued estimates of the local Gaussian cross-spectrum at the frequency $\omega = \alpha$, and where means and quantiles relative to a polar representation, i.e., $z = re^{i\theta}$, have been added to the plot.

The center panel of Figure S6.2 shows the same estimated values, but this time the means/quantiles are based on a Cartesian representation $z = x + iy$. These two panels gives a geometrical view to the observations presented above. The third panel of Figure S6.2 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 2.5.2.

S6.5. The local bivariate trigonometric examples in Figures 2 to 5

This section will discuss some topics related to the *local bivariate trigonometric* examples, whose local Gaussian spectral density were investigated in Figures 2 to 5 of Section 3.1 in the main part. This discussion is quite similar to the one given for the univariate *local trigonometric* examples in JT22 [Appendix G.4], and the interested reader should consult this source for some of the in depth discussions.

The *local bivariate trigonometric* examples are a special case of a more general bivariate construction, which follows from an extension of the corresponding univariate construction given in JT22 [Appendix G.4.1]. The details are slightly technical, and they have for the benefit of the reader been included in Section S6.5.1.

A brief discussion of the heuristic arguments that explains why the *local bivariate trigonometric* examples can be used to sanity test the implemented estimation algorithm are given in Section S6.5.2. An in depth discussion can be found in JT22 [Appendix G.4.2].

The *local bivariate trigonometric* examples that was used as the basis for the plots in Figures 2 to 5, contained four components, of which one occurred with a lower probability. This setup was inherited from the *local unvariate trigonometric* examples in JT22, and it was there seen (cf. Figure G.3 in JT22 [Appendix G.4.3]) that the "lower tail"-component could be detected when the sample size was large enough. Section S6.5.4 presents a similar investigation for the *local bivariate trigonometric* examples.

The discussion in JT22 also considered issues that could occur if the m-truncated local Gaussian spectrum was estimated for samples from a deterministic function perturbed by very low random fluctuations, and it was there seen (cf. JT22 [Appendix G.4.4]) that the local Gaussian machinery did not work well for such cases. This discussion will not be duplicated in this paper, but the key observations can be mentioned here: It might be safest to apply the local Gaussian spectral machinery to time series with rather flat global spectra. If a "clear" global structure is present, then it might be preferable to use the local Gaussian spectral machinery on the residuals that are left after some suitable model have been fitted to the data.

S6.5.1. Some properties of the general construction

The *local bivariate trigonometric* examples considered in Section 3.1 are (as mentioned there) particular cases of a general construction that is based on the key idea that an artificial bivariate time series $\left\{\left(Y_{1,t}, Y_{2,t}\right)\right\}_{t\in\mathbb{Z}}$ can be constructed by the following scheme:

1. Select $r \geq 2$ bivariate time series $\left\{\left(C_{1,i}(t), C_{2,i}(t)\right)\right\}_{i=1}^{r}$.
2. Select a random variable I with values in the set $\{1,\ldots,r\}$, and use this to sample a collection of indices $\{I_t\}_{t\in\mathbb{Z}}$ (that is, for each t an independent realization of I is taken). Let $p_i := \mathrm{P}(I_i = i)$ denote the probabilities for the different outcomes.
3. Define Y_t by means of the equation

$$Y_{1,t} := \sum_{i=1}^{r} C_{1,i}(t) \cdot \mathbb{1}\{I_t = i\}, \tag{S39a}$$

$$Y_{2,t} := \sum_{i=1}^{r} C_{2,i}(t) \cdot \mathbb{1}\{I_t = i\}. \tag{S39b}$$

The basic properties of $\left\{\left(Y_{1,t}, Y_{2,t}\right)\right\}_{t\in\mathbb{Z}}$ can be expressed relatively those of $\left\{\left(C_{1,i}(t), C_{2,i}(t)\right)\right\}$ as seen in the following result:

Lemma S1. *With* $\left\{\left(Y_{1,t}, Y_{2,t}\right)\right\}_{t\in\mathbb{Z}}$ *as defined above, it follows that:*

1. $\mathrm{E}\left[Y_{k,t}\right] = \sum_{i=1}^{r} p_i \cdot \mathrm{E}\left[C_{k,i}(t)\right]$

2. $\mathrm{E}\Big[Y_{k,t+h} \cdot Y_{\ell,t}\Big] = \begin{cases} \sum_{i=1}^{r}\sum_{j=1}^{r} p_i \cdot p_j \cdot \mathrm{E}\Big[C_{k,i}(t+h) \cdot C_{\ell,j}(t)\Big] & h \neq 0 \\ \sum_{i=1}^{r} p_i \cdot \mathrm{E}\Big[C_{k,i}(t) \cdot C_{\ell,j}(t)\Big] & h = 0 \end{cases}$

3. $\mathrm{Cov}\Big(Y_{k,t+h}, Y_{\ell,t}\Big) = \begin{cases} \sum_{i=1}^{r}\sum_{j=1}^{r} p_i \cdot p_j \cdot \mathrm{Cov}\Big(C_{k,i}(t+h), C_{\ell,j}(t)\Big) & h \neq 0 \\ \sum_{i=1}^{r} p_i \cdot \mathrm{E}\Big[C_{k,i}(t) \cdot C_{\ell,i}(t)\Big] - & \\ \quad \Big(\sum_{i=1}^{r} p_i \cdot \mathrm{E}\Big[C_{k,i}(t)\Big]\Big) \cdot \Big(\sum_{j=1}^{r} p_j \cdot \mathrm{E}\Big[C_{\ell,j}(t)\Big]\Big) & h = 0 \end{cases}$

4. *For $k, \ell \in \{1,2\}$, the additional assumption that $\Big(C_{1,i}(t), C_{2,i}(t)\Big)$ and $\Big(C_{1,j}(t), C_{2,j}(t)\Big)$ are independent when $i \neq j$, simplifies the $h \neq 0$ case to:* $\mathrm{Cov}\Big(Y_{k,t+h}, Y_{\ell,t}\Big) = \sum_{i=1}^{r} p_i^2 \cdot \mathrm{Cov}\Big(C_{k,i}(t+h), C_{\ell,i}(t)\Big)$.

Proof. The results in Items 1 to 4 are in essence the same as those seen in JT22 [Lemma G.1], with some extra indices k and ℓ added to it. The proof is short, and it is repeated below.

The random variable I_t that produces the set of indices is independent of $C_{k,i}(t)$ and $C_{\ell,j}(t)$, and Item 1 thus follows without further ado. For the $h \neq 0$ case of Item 2 it is sufficient to note that I_{t+h} and I_t then are independent, and it follows that $\mathrm{E}\big[\mathbb{1}\{I_{t+h} = i\} \cdot \mathbb{1}\{I_t = j\}\big] = \mathrm{E}\big[\mathbb{1}\{I_{t+h} = i\}\big] \cdot \mathrm{E}[\mathbb{1}\{I_t = j\}] = \mathrm{P}(I_{t+h} = i) \cdot \mathrm{P}(I_t = j) = p_i \cdot p_j$. For the $h = 0$ case of Item 2 it is enough to note that $\mathbb{1}\{I_t = i\} \cdot \mathbb{1}\{I_t = j\} = 0$ when $i \neq j$, which together with $\mathbb{1}\{I_t = i\} \cdot \mathbb{1}\{I_t = i\} = \mathbb{1}\{I_t = i\}$ gives the required expression. The statements in Items 3 and 4 follows trivially from those in Items 1 and 2. □

The key idea in the local bivariate trigonometric example is that the r bivariate time series $\Big(C_{1,i}(t), C_{2,i}(t)\Big)$ all should be "connected cosine-pairs with some noise", since this implies (given a reasonable parameter configuration) that it should be possible to present a decent guesstimate with regard to the expected shape of the m-truncated local Gaussian cross-spectrum density (for some carefully selected tuning parameters of the estimation algorithm). The global cross-spectrum in this case will not be flat, but it will for low truncation levels be "flat enough" for the purpose of showing that the global spectrum does not detect the underlying frequencies whereas the local Gaussian cross-spectrum can do that task.

The following result reiterates the $\Big(C_{1,i}(t), C_{2,i}(t)\Big)$-definition used in the local bivariate trigonometric example, and it presents some basic properties related to this definition.

Lemma S2. *Let the bivariate random variables $\Big(C_{1,i}(t), C_{2,i}(t)\Big)$, for $i = 1, \ldots, r$, be defined by $C_{1,i}(t) = L_i + A_i(t) \cdot \cos\big(2\pi\alpha_i t + \phi_i\big)$ and $C_{2,i}(t) = L_i + A_i(t) \cdot \cos\big(2\pi\alpha_i t + \phi_i + \theta_i\big)$ in the following manner: L_i and α_i are constants that respectively defines the horizontal base-line and the frequency. The amplitude $A_i(t)$ are for each t uniformly distributed on an interval $[a_i, b_i]$, and $A_i(t+h)$ and $A_i(t)$ are independent when $h \neq 0$. The phase-adjustment ϕ_i are uniformly drawn (one time for each realisation) from the interval between 0 and 2π, whereas the phases θ_i are constants. It is also assumed that the stochastic processes ϕ_i and $A_i(t)$ are independent of each other.*

1. $\mathrm{E}\Big[C_{k,i}(t)\Big] = L_i$, *for* $k \in \{1,2\}$.
2. $\mathrm{E}\Big[C_{k,i}(t+h)C_{\ell,j}(t)\Big] = L_i L_j$, *for* $k, \ell \in \{1,2\}$ *when* $i \neq j$.
3. $\mathrm{Cov}\Big(C_{k,i}(t+h), C_{\ell,j}(t)\Big) = 0$, *for* $k, \ell \in \{1,2\}$ *when* $i \neq j$.
4. $\mathrm{E}\Big[C_{k,i}(t+h) \cdot C_{k,i}(t)\Big] = \begin{cases} L_i^2 + \frac{1}{4} \cdot \big(a_i^2 + 2a_i b_i + b_i^2\big) \cdot \cos(2\pi\alpha_i h) & h \neq 0, k \in \{1,2\} \\ L_i^2 + \frac{1}{3} \cdot \big(a_i^2 + a_i b_i + b_i^2\big) & h = 0, k \in \{1,2\} \end{cases}$
5. $\mathrm{Cov}\Big(C_{k,i}(t+h), C_{k,i}(t)\Big) = \begin{cases} \frac{1}{4} \cdot \big(a_i^2 + 2a_i b_i + b_i^2\big) \cdot \cos(2\pi\alpha_i h) & h \neq 0, k \in \{1,2\} \\ \frac{1}{3} \cdot \big(a_i^2 + a_i b_i + b_i^2\big) & h = 0, k \in \{1,2\} \end{cases}$
6. $\mathrm{E}\Big[C_{k,i}(t+h) \cdot C_{\ell,i}(t)\Big] = \begin{cases} L_i^2 + \frac{1}{4} \cdot \big(a_i^2 + 2a_i b_i + b_i^2\big) \cdot \cos(2\pi\alpha_i h + \theta_i) & h \neq 0, k = 1, \ell = 2 \\ L_i^2 + \frac{1}{4} \cdot \big(a_i^2 + 2a_i b_i + b_i^2\big) \cdot \cos(2\pi\alpha_i h - \theta_i) & h \neq 0, k = 2, \ell = 1 \\ L_i^2 + \frac{1}{3} \cdot \big(a_i^2 + a_i b_i + b_i^2\big) \cdot \cos(\theta_i) & h = 0, k, \ell \in \{1,2\} \end{cases}$

7. $$\mathrm{Cov}\Big(C_{k,i}(t+h), C_{\ell,i}(t)\Big) = \begin{cases} \frac{1}{4}\cdot\left(a_i^2 + 2a_ib_i + b_i^2\right)\cdot\cos(2\pi\alpha_i h + \theta_i) & h \neq 0, k = 1, \ell = 2 \\ \frac{1}{4}\cdot\left(a_i^2 + 2a_ib_i + b_i^2\right)\cdot\cos(2\pi\alpha_i h - \theta_i) & h \neq 0, k = 2, \ell = 1 \\ \frac{1}{3}\cdot\left(a_i^2 + a_ib_i + b_i^2\right)\cdot\cos(\theta_i) & h = 0, k, \ell \in \{1, 2\} \end{cases}$$

Proof. The independence of $A_i(t)$ and ϕ_i, together with some simple results related to the expectations of functions based on the uniformly distributed random variables A_i and ϕ is needed in order to obtain these results.

The assumption that A_i is uniformly distributed on the interval (a_i, b_i) implies that its density function is given by $\frac{1}{b_i - a_i}$ on this interval (and 0 outside of it). From this it is easy to find $\mathrm{E}\big[A_i(t)\big] = \frac{1}{2}\cdot\left(a_i + b_i\right)$ and $\mathrm{E}\big[A_i^2(t)\big] = \frac{1}{3}\cdot\left(a_i^2 + a_ib_i + b_i^2\right)$. Moreover, since $A_i(t+h)$ and $A_i(t)$ are independent when $h \neq 0$, it also follows that $\mathrm{E}\big[A_i(t+h)A_i(t)\big] = \frac{1}{4}\left(a_i^2 + 2a_ib_i + b_i^2\right)$ in this case.

Similarly, the assumption that ϕ_i is uniformly distributed on the interval $(0, 2\pi)$ implies that its density function is given by $\frac{1}{2\pi}$ on the interval $(0, 2\pi)$ (and 0 outside of it). From this it follows that $\mathrm{E}\big[\cos(\phi_i)\big] = \mathrm{E}\big[\sin(\phi_i)\big] = \mathrm{E}\big[\cos(\phi_i)\sin(\phi_i)\big] = 0$ and $\mathrm{E}\Big[\cos(\phi_i)^2\Big] = \mathrm{E}\Big[\sin(\phi_i)^2\Big] = \frac{1}{2}$. In addition to this, since ϕ_i and ϕ_j are independent when $i \neq j$, it also follows that $\mathrm{E}\Big[\cos(\phi_i)\sin\Big(\phi_j\Big)\Big] = 0$. From this it follows that $\mathrm{E}\big[\cos\big(2\pi\alpha_i t + \phi_i\big)\big] = \mathrm{E}\big[\cos\big(2\pi\alpha_i t + \phi_i + \theta_i\big)\big] = 0$.

The trigonometric formula $\cos(u \pm v) = \cos(u)\cos(v) \mp \sin(u)\sin(v)$ can be used to show that the expectation of a product of the form $\cos\big(u_1 + \phi_i\big)\cos\Big(u_2 + \phi_j\Big)$ is equal to 0 when $i \neq j$, whereas it is equal to $\frac{1}{2}\cos\big(u_1 - u_2\big)$ when $i = j$. From this last observation it is easy to conclude the following:

$$\mathrm{E}[\cos(2\pi\alpha_i(t+h) + \phi_i)\cos(2\pi\alpha_i t + \phi_i)] = \frac{1}{2}\cos(2\pi\alpha_i h) \tag{S40a}$$

$$\mathrm{E}[\cos(2\pi\alpha_i(t+h) + \phi_i + \theta_i)\cos(2\pi\alpha_i t + \phi_i + \theta_i)] = \frac{1}{2}\cos(2\pi\alpha_i h) \tag{S40b}$$

$$\mathrm{E}[\cos(2\pi\alpha_i(t+h) + \phi_i)\cos(2\pi\alpha_i + \phi_i + \theta_i)] = \frac{1}{2}\cos(2\pi\alpha_i h - \theta_i) \tag{S40c}$$

$$\mathrm{E}[\cos(2\pi\alpha_i(t+h) + \phi_i + \theta_i)\cos(2\pi\alpha_i + \phi_i)] = \frac{1}{2}\cos(2\pi\alpha_i h + \theta_i) \tag{S40d}$$

From these observations it is trivial to verify Items 1 and 2, and Item 3 follows immediately from these two. The statements in Items 4 and 6 are easily verified by multiplying the relevant expressions for $C_{1,i}(t) = L_i + A_i(t)\cdot\cos\big(2\pi\alpha_i t + \phi_i\big)$ and $C_{2,i}(t) = L_i + A_i(t)\cdot\cos\big(2\pi\alpha_i t + \phi_i + \theta_i\big)$, and then using the above mentioned results related to $A_i(t)$ and ϕ_i to establish the expected values. The covariance-results in Items 5 and 7 now follows trivially. □

Finally, the *local bivariate trigonometric* example is obtained by using r bivariate time series $\Big(C_{1,i}(t), C_{2,i}(t)\Big)$, of the form given in Lemma S2, in the construction of the bivariate time series $\Big(Y_{1,t}, Y_{2,t}\Big)$, i.e.,

$$Y_{1,t} := \sum_{i=1}^{r} \mathbb{1}\{I_t = i\}\cdot[L_i + A_i(t)\cdot\cos(2\pi\alpha_i t + \phi_i)], \tag{S41a}$$

$$Y_{2,t} := \sum_{i=1}^{r} \cdot\mathbb{1}\{I_t = i\}\cdot[L_i + A_i(t)\cdot\cos(2\pi\alpha_i t + \phi_i + \theta_i)], \tag{S41b}$$

where it furthermore is assumed that the i-indexed stochastic variables $A_i(t)$ and φ_i are independent of the j-indexed variants when $i \neq j$. It now follows from Lemmas S1 and S2

that the $h \neq 0$ auto-correlations of the bivariate time series $\left(Y_{1,t}, Y_{2,t}\right)$ in Equation (S41) is given by

$$\rho_{11}(h) = \rho_{22}(h) = \frac{\frac{1}{4} \cdot \sum_{i=1}^{r} p_i^2 \cdot \left(a_i^2 + 2a_i b_i + b_i^2\right) \cdot \cos(2\pi\alpha_i h)}{\sum_{i=1}^{r} p_i \cdot \left[L_i^2 + \frac{1}{3} \cdot \left(a_i^2 + a_i b_i + b_i^2\right)\right] - \left(\sum_{i=1}^{r} p_i \cdot L_i\right)^2}. \tag{S42}$$

The cross-correlations $\rho_{Y_{12}}(h)$ and $\rho_{Y_{12}}(h)$ can for $h \neq 0$ similarly be written out as

$$\rho_{Y_{12}}(h) = \frac{\frac{1}{4} \cdot \sum_{i=1}^{r} p_i^2 \cdot \left(a_i^2 + 2a_i b_i + b_i^2\right) \cdot \cos(2\pi\alpha_i h + \theta_i)}{\sum_{i=1}^{r} p_i \cdot \left[L_i^2 + \frac{1}{3} \cdot \left(a_i^2 + a_i b_i + b_i^2\right)\right] - \left(\sum_{i=1}^{r} p_i \cdot L_i\right)^2}, \tag{S43a}$$

$$\rho_{Y_{21}}(h) = \frac{\frac{1}{4} \cdot \sum_{i=1}^{r} p_i^2 \cdot \left(a_i^2 + 2a_i b_i + b_i^2\right) \cdot \cos(2\pi\alpha_i h - \theta_i)}{\sum_{i=1}^{r} p_i \cdot \left[L_i^2 + \frac{1}{3} \cdot \left(a_i^2 + a_i b_i + b_i^2\right)\right] - \left(\sum_{i=1}^{r} p_i \cdot L_i\right)^2}, \tag{S43b}$$

whereas the $h = 0$ case is given by:

$$\rho_{Y_{12}}(0) = \rho_{21}(0) = \frac{\sum_{i=1}^{r} p_i \cdot \left[L_i^2 + \frac{1}{3} \cdot \left(a_i^2 + a_i b_i + b_i^2\right)\right] \cdot \cos(\theta_i) - \left(\sum_{i=1}^{r} p_i \cdot L_i\right)^2}{\sum_{i=1}^{r} p_i \cdot \left[L_i^2 + \frac{1}{3} \cdot \left(a_i^2 + a_i b_i + b_i^2\right)\right] - \left(\sum_{i=1}^{r} p_i \cdot L_i\right)^2}. \tag{S44}$$

An inspection of Equations (S42) and (S43) reveals that it is fairly easy to find a parameter configuration for which the numerator is rather small compared to the denominator. This is of course not white noise, but the key idea is that it is close enough to white noise to make it impossible to deduce anything about the underlying frequencies α_i and the phase-differences θ_i based on the ordinary cross-spectrum.

S6.5.2. The heuristic argument underlying the local bivariate trigonometric example

The purpose of this section is to give a short summary of the corresponding discussion from JT22 [Appendix G.4.2], i.e., to briefly explain the reasoning that motivated the construction of the *local bivariate trigonometric* example in Section S6.5.1. Only the general ideas will be described here, and the interested reader should consult JT22 [Appendix G.4.2] for an in depth discussion.

The key motivation for the construction of the *local bivariate trigonometric* models is the sanity testing of the estimation algorithm for the m-truncated local Gaussian cross-spectrum $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$, i.e., to construct models in such a manner that heuristic arguments can be used to give a "decent guesstimate of the result" when the m-truncated estimates $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ are created based on pseudo-normalised samples from these models.

In order for this to work, it is necessary to select several parameters in a structured way, both on the side of the *local bivariate trigonometric* models (probabilities p_i, amplitudes A_i, base-lines L_i, frequencies ϕ_i and phases θ_i) and on the side of the local Gaussian estimation algorithm (truncation level m, bandwidth $\boldsymbol{b}$ and point of investigation $\boldsymbol{v}$).

The problem is to figure out how the parameters p_i, A_i and L_i should "map to" points $\boldsymbol{v}$ and bandwidths $\boldsymbol{b}$ when the samples are pseudo-normalised. This problem was discussed in detail in JT22 [Appendix G.4.2], and this discussion will not be repeated here.

The following observation is the pivotal one for the heuristic argument that motivates the sanity testing. For a *local bivariate trigonometric* random variable $\left(Y_{1,t}, Y_{2,t}\right)$, as given in Equation (S41), the idea goes as follows: For specified indices i and j, it should be possible to find a "good combination" of point $\boldsymbol{v}$ and bandwidth $\boldsymbol{b}$, such that the estimated m-truncated local Gaussian spectrum $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ looks similar to the spectrum that would occur if only the $C_{k,i}(t)$ and $C_{\ell,j}(t)$ components had been used to generate the sample under investigation.

The argument for this part is actually quite simple: For the "correct combination" of $\boldsymbol{v}$ and $\boldsymbol{b}$, the local kernel function in the estimation algorithm of the local Gaussian

cross-correlations $\rho_{k\ell:\boldsymbol{v}}(h)$ will in essence "ignore" the contributions from other components than $C_{k,i}(t)$ and $C_{\ell,j}(t)$ when the lag h pairs from $Y_{k,t}$ and $Y_{\ell,t}$ are considered (since low weights are assigned to the lag h pairs that lies "far away" from the point $\boldsymbol{v}$). The lag h pairs that contribute most to the estimated value $\rho_{k\ell:\boldsymbol{v}}(h)$, can be seen as a "randomly selected subset" of those lag h pairs that would have occurred if all the observations had been from the $C_{k,i}(t)$ and $C_{\ell,j}(t)$ components.

If this "randomly selected subset" from $C_{k,i}(t)$ and $C_{\ell,j}(t)$ is sufficiently large, then it can be expected that the local Gaussian cross-correlation $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ estimated from this subset should be "fairly close to" the one estimated from the full set. This argument does not depend on the lag h value, so the expected result is that the estimates $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$ all should be "fairly close to" to those estimated from the full set.

Since the estimate $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$ of the local Gaussian cross-spectrum $f^m_{k\ell:\boldsymbol{v}}(\omega)$ is a linear combination of the estimates $\hat{\rho}_{k\ell:\boldsymbol{v}}(h)$, it now seems reasonable to expect that the shape of it should share some clear similarities with the one based on the full set. This is the reasoning that enabled *local bivariate trigonometric* models to be used for the sanity testing of the estimation algorithm for the m-truncated local Gaussian cross-spectrum $\hat{f}^m_{k\ell:\boldsymbol{v}}(\omega)$, as seen in Section 3.1.2 in the main part.

Strictly speaking, neither $f_{k\ell}(\omega)$ nor $f_{k\ell:\boldsymbol{v}}(\omega)$ are well defined for the *local bivariate trigonometric* times series, but this is not a problem when the m-truncated spectra $f^m_{k\ell}(\omega)$ and $f^m_{k\ell:\boldsymbol{v}}(\omega)$ are investigated. The important detail is that it (for suitably selected combinations of point $\boldsymbol{v}$ and bandwidth $\boldsymbol{b}$) is possible to "predict" that the m-truncated estimates $f^m_{k\ell:\boldsymbol{v}}(\omega)$ seen in Figure 3 should look a bit like like the "global bivariate cosine combination" seen in Figure 1.

S6.5.3. What points to investigate for the bivariate local trigonometric examples?

The *local bivariate trigonometric* models used for the sanity testing in the main part used points located along the diagonal, like seen in Figure 3. The diagonal points were selected since it for them was "fairly clear" that the resulting local Gaussian spectra would share some similarities with the basic reference example seen in Figure 1, which enabled an easy visual comparison of the results.

The heatmap+distance plot seen in Figure 2 also focused on points along the diagonal, and from this it can be seen that the spectra will (as expected per construction) change a lot as the point $\boldsymbol{v}$ varies along the diagonal. It is of course possible to consider points outside of the diagonal too, and an example of this can be seen in Figure S6.3. The same sample that was used for Figure 3, has here been investigated for the three points `10%::90%`, `10%::50%` and `50%::90%` (percentages relates to quantiles of the standard normal distribution). This plot is, as was done for Figure 3, based on a presentation of the `Co`-, `Quad`- and `Phase`-plots of the $m = 10$ truncated local Gaussian cross-spectrum.

The plots in Figure S6.3 shows that the `Co`-, `Quad`- and `Phase`-plots at the point `10%::90%` (middle column) looks more like the i.i.d. white noise that was encountered in Figure S6.1 (see page 35), whereas the plots for the two points `10%::50%` and `50%::90%` do 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/peaks 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 Figure 3 supports the idea that there are local features in the data that depends on these α-values. These features are not detected by the $m = 10$ truncated ordinary cross-spectrum $f^m_{k\ell}(\omega)$, but they are detected by the local Gaussian spectrum $f^m_{k\ell:\boldsymbol{v}}(\omega)$.

Note that it is possible to create new "global trigonometric examples", like the one seen in Figure 1, in order to see if the figures in Figure S6.3 are similar to those — but that is only part of the work to do. The bandwidth $\boldsymbol{b} = 0.6$ used for he investigation in Figure 3 might not be the "correct one" for the points investigated in Figure S6.3. If the bandwidth $\boldsymbol{b}$ is too big, then "contamination" might occur, i.e, the resulting local Gaussian spectra will

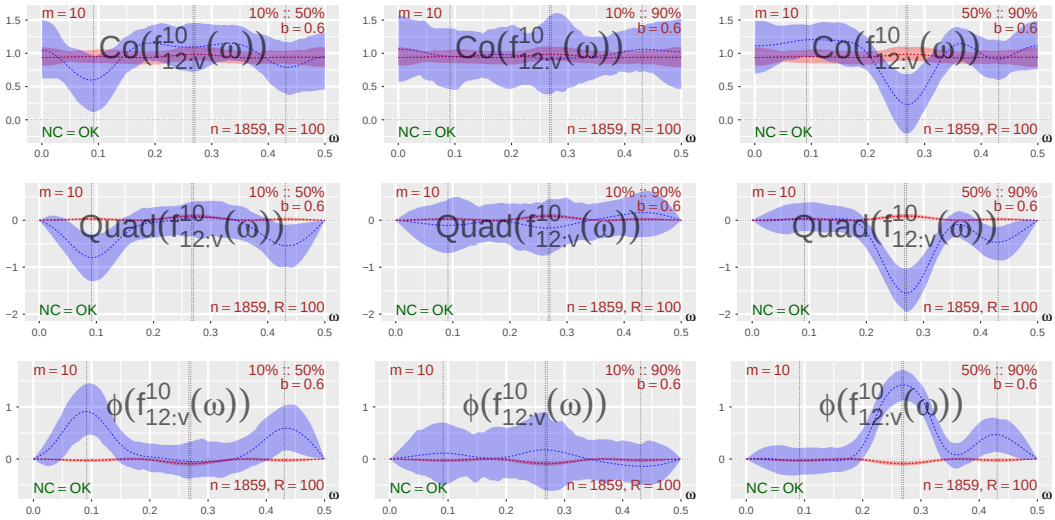

Figure S6.3. `Co`-, `Quad`- and `Phase`-plots for points away from the diagonal: three diagonal points: The *bivariate local trigonometric* model in Equation (S39), constant phase-changes $\theta_i = \pi/3, i = 1, 2, 3, 4$. The frequencies α_i shown as vertical lines. The local Gaussian spectra detects structures that are not detected by the ordinary spectrum.

pick up features from the other components used in the construction of the *local bivariate trigonometric* models, see the discussion in JT22 [Appendix G.4.2] for further details.

S6.5.4. Heatmap+distance plots, bivariate local trigonometric and extreme tail

Another aspect that can be discussed based on the *local bivariate trigonometric* examples is the capability of the local Gaussian estimation machinery to pick up local structures that occurs far out in the tails. This section will mirror the discussion in JT22 [Appendix G.4.3], and it will also here be investigated if a known local structure in the periphery of the data might be detected if the sample size is large.

The *local bivariate trigonometric* example in this paper is based on a simple bivariate extension of the corresponding *local trigonometric* example from JT22. It contains 4 different components $\left(C_{1,i}(t), C_{2,i}(t)\right)$, $i = 1, \ldots, 4$, of which the first component $\left(C_{1,1}(t), C_{2,1}(t)\right)$ occurs with probability $p_1 = 0.05$. For a sample of length 1859 (length inherited from the investigation of the `EuStockMarkets`-example) it is then expected that $0.05 \cdot 1859 = 92.95$ observations are from this first component — and these observations will (by construction) all be in the lower tail.

As explained in JT22 [Appendix G.4.3], the "border" between the observations from the $i = 1$ and $i = 2$ components should occur near the 5% percentile, but it is necessary to "zoom in" on a diagonal point v that lies farther out in the tail than $p_1/2 = 0.025$. This requirement occurs since the estimate should avoid "contamination" from the observations from the $i = 2$ component.

The parameters selected for the present investigation are all inherited from the following discussion in JT22 [Appendix G.4.3]:

> Based on the idea that it might be necessary to go all the way out to the 1%, it seemed natural to attempt an investigation based on $n = 25000$ observations. Since the point v now is far out in the lower tail, e.g. the 1% percentile of the standard normal distribution is -2.326, it seemed reasonable to use the bandwidth $\boldsymbol{b} = (0.4, 0.4)$.
>
> The heatmap and distance plots in Figure G.2 is based on an investigating of a single realisation, that included percentiles based on values starting from 2 bandwidths below the 5% percentile and ending at 1/2 bandwidth below the 5%

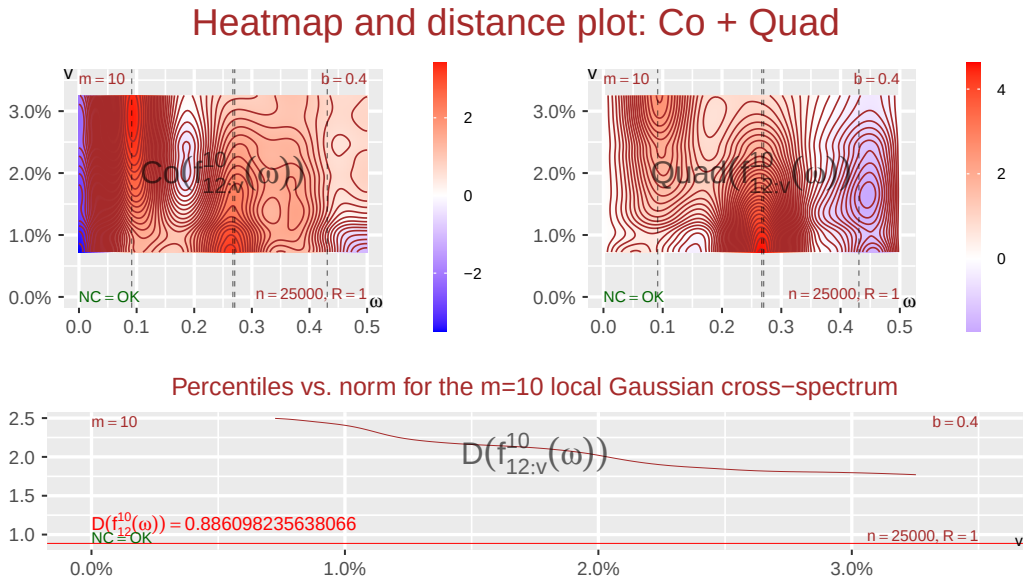

Figure S6.4. `Co`- and `Quad`-heatmaps, distance-plot: Sample from the *bivariate local trigonometric* model in Equations (S39) and (S41), constant phase-changes $\theta_i = \pi/3, i = 1,2,3,4$. The frequencies α_i shown as vertical lines. This can be used to search for an "optimal" percentile that can reveal the elusive α_1 frequency in the lower tail.

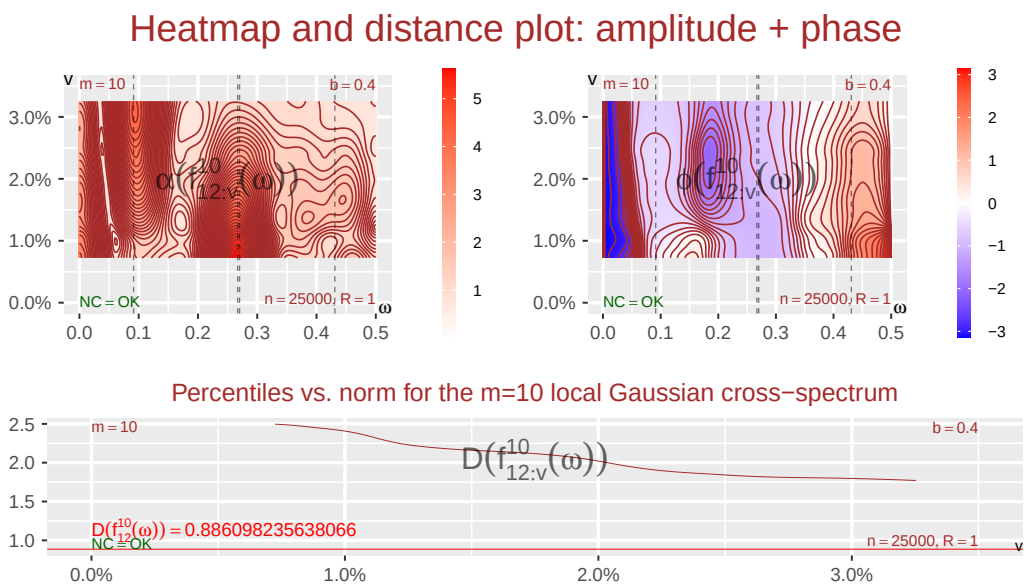

Figure S6.5. `Amplitude`- and `Phase`-heatmaps, distance-plot: Sample from the *bivariate local trigonometric* model in Equations (S39) and (S41), constant phase-changes $\theta_i = \pi/3, i = 1,2,3,4$. The frequencies α_i shown as vertical lines. This can be used to search for an "optimal" percentile that can reveal the elusive α_1 frequency in the lower tail.

percentile, i.e., the diagonal points starts at approximately the 0.72% percentile and ends at the 3.25% percentile.

Note that the reference to Figure G.2 in this quoted text is to JT22 [Figure G.2]. The corresponding heatmap and distance plots for the *local bivariate trigonometric* example are given in Figure S6.4 (`Co`- and `Quad`-version) and Figure S6.5 (`Amplitude`- and `Phase`-version).

The heatmap-parts of Figures S6.4 and S6.5 reveal that the contribution from the $i = 2$ component completely dominates at the 3.25% percentile, and it can also be seen that it is

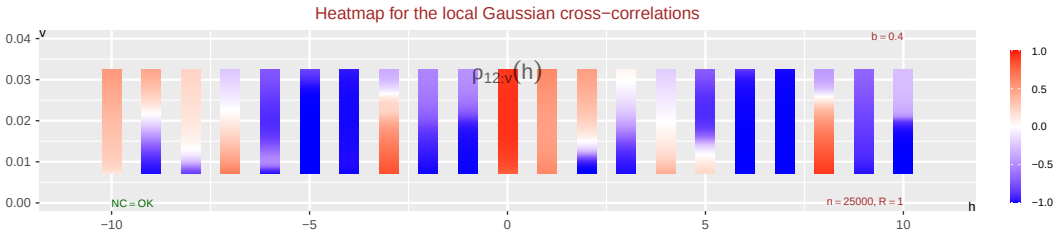

Figure S6.6. Heatmap plots for the local Gaussian cross-correlations $\rho_{k\ell:v}(h)$, for diagonal points v in the lower tail. Values close to $+1$ and -1 can indicate that the estimates have started to degenerate.

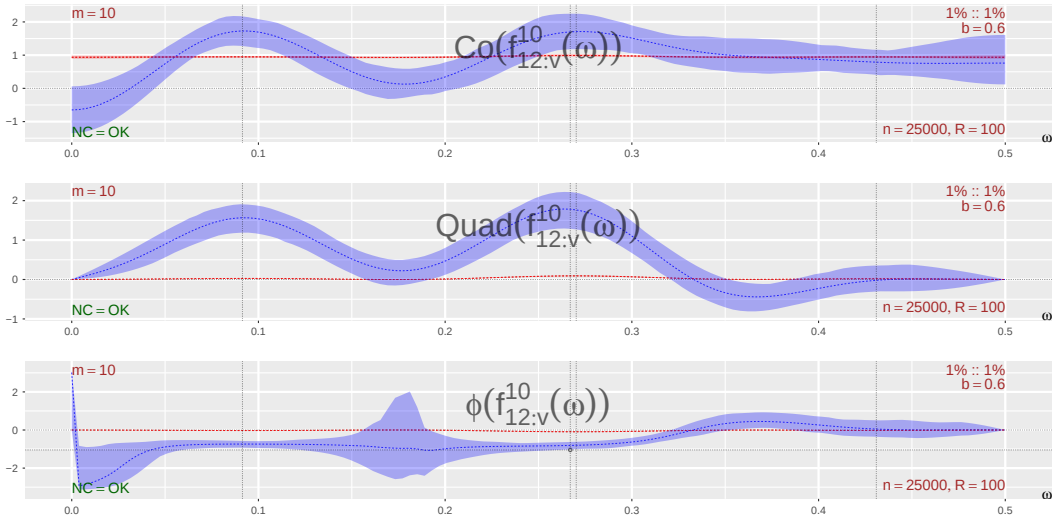

Figure S6.7. The detection of the $\alpha_1 = 0.267$ frequency in the lower tail of the *local trigonometric* example requires a large sample and an investigation far out in the lower tail. In this case the result ended out being contaminated by the $\alpha_2 = 0.091$ frequency.

necessary to go down to at least the 1% percentile in order to detect the local structures due to the $i = 1$ component (frequency $\alpha_1 = 0.267$). Note that Figures S6.4 and S6.5 are based on only 1 single realisation, and other realisations might look slightly different.

The content of Figures S6.4 and S6.5 are based on estimates $\hat{\rho}_{k\ell:v}(h)$ of the local Gaussian cross-correlations $\rho_{k\ell:v}(h)$. Note that these local Gaussian estimates focus on points in the periphery of the observations, and the kernel function in the estimation algorithm might then, as explained in the sensitivity analysis in Section S3.3, give estimates that degenerate towards $+1$ or -1. The presence/absence of this degeneration-problem can be investigated using the heatmaps in Figure S6.6, and the conclusion in this case is that it seems to be some (but not all) estimates that have degenerated in this manner.

Figure S6.7 shows the situation when $R = 100$ replicates are used to estimate the local Gaussian cross-spectrum $f^m_{k\ell:v}(\omega)$ at the diagonal point v that corresponds to the 1% percentile. The conclusion in JT22, when a similar investigation was performed for the univariate *local trigonometric* example, see in particular JT22 [Figure G.3], was that the local Gaussian estimates resulted in a plot with the expected "cosine-shape" and a peak near the frequency $\alpha_1 = 0.267$ of the $i = 1$ component.

The situation for the bivariate case investigated in this section, see Figure S6.7, diverges from the one seen for the univariate case. There is also for this case a peak near $\alpha_1 = 0.267$ due to the $i = 1$ component, but there is also a similar clear peak near the frequency $\alpha_2 = 0.091$ due to the $i = 2$ component. This indicates that the bandwidth $\boldsymbol{b} = 0.4$ is too large in this case, since it is a clear contamination from the $i = 2$ component.

The $i = 1$ component $\left(C_{1,1}(t), C_{2,1}(t)\right)$ was included in the *local bivariate trigonometric* example in order to emphasise that extra care is needed when investigating points v in the extreme tails of a sample — and the present discussion reveals that it even for a known

model can be quite tricky to detect local interdependency-structures, in particular if they do not occur frequently in the observations. The size of the sample is pivotal for the level of details that can be extracted, but the selection of the bandwidth must also take into account the point $\boldsymbol{v}$ under investigation.

Keep in mind, as explained in Section S3.3, that a too small bandwidth $\boldsymbol{b}$ can trigger degenerate estimates close to $+1$ and -1, whereas a too big bandwidth returns estimates that can be contaminated by neighbouring structures. A local Gaussian investigation should thus investigate a range of bandwidths, in order to see what kind of behaviour that occurs. The heatmap and distance plots can be useful for this task.

Author Contributions: Conceptualization, LAJ and DT; methodology, LAJ and DT; software, LAJ; validation, LAJ; formal analysis, LAJ and DT; investigation, LAJ and DT; writing—original draft preparation, LAJ; writing—review and editing, LAJ and DT; visualization, LAJ and DT; supervision, DT. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The `EuStockMarkets`-dataset is a part of the `datasets`-package of `R`, R Core Team (2020). The scripts needed for the reproduction of the examples in this paper is contained in the `R`-package `localgaussSpec`, cf. Section S6.1 for further details.

Acknowledgments: The authors are most grateful for the valuable comments and suggestions from the referees and the associate editor.

Conflicts of Interest: The authors declare no conflict of interest.

References

Birr, Stefan, Tobias Kley, and Stanislav Volgushev. 2019. Model assessment for time series dynamics using copula spectral densities: A graphical tool. *Journal of Multivariate Analysis* 172: 122–46. Dependence Models, https://doi.org/10.1016/j.jmva.2019.03.003.

Bollerslev, Tim, and Eric Ghysels. 1996. Periodic Autoregressive Conditional Heteroscedasticity. *Journal of Business & Economic Statistics* 14: 139–51. https://doi.org/10.1080/07350015.1996.10524640.

Brockwell, Peter J., and Richard A Davis. 1986. *Time Series: Theory and Methods*. New York: Springer-Verlag New York, Inc.

Ghalanos, Alexios. 2022. *rmgarch: Multivariate GARCH Models.* R Package Version 1.3-9. Retrieved from https://cran.r-project.org/package=rmgarch (accessed December 10, 2022).

Ghalanos, Alexios. 2022. *rugarch: Univariate GARCH Models*.5 R Package Version 1.4-9. Retrieved from https://cran.r-project.org/package=rugarch (accessed December 10, 2022).

Hjort, Nils Lied., and M. C. Jones. 1996. Locally parametric nonparametric density estimation. *Annals of Statistics* 24: 1619–47. https://doi.org/10.1214/aos/1032298288.

Jordanger, Lars Arne, and Dag Tjøstheim. 2022. Nonlinear Spectral Analysis: A Local Gaussian Approach. *Journal of the American Statistical Association* 117: 1010–27. https://doi.org/10.1080/01621459.2020.1840991.

Klimko, Lawrence A., and Paul I. Nelson. 1978. On Conditional Least Squares Estimation for Stochastic Processes. *Annals of Statistics* 6: 629–42. https://doi.org/10.1214/aos/1176344207.

Künsch, Hans R. 1989. The Jackknife and the Bootstrap for General Stationary Observations. *The Annals of Statistics* 17: 1217–41. https://doi.org/10.1214/aos/1176347265.

Politis, Dimitris N., and Joseph P. Romano. 1992. A General Resampling Scheme for Triangular Arrays of α-Mixing Random Variables with Application to the Problem of Spectral Density Estimation. *The Annals of Statistics* 20: 1985–2007. https://doi.org/10.1214/aos/1176348899.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Tjøstheim, Dag, and Karl Ove Hufthammer. 2013. Local Gaussian correlation: A new measure of dependence. *Journal of Econometrics* 172: 33–48. https://doi.org/10.1016/j.jeconom.2012.08.001.