A nonparametric regression cross spectrum for multivariate time series

Jan Beran
and
Mark A. Heiler
Department of Mathematics and Statistics
University of Konstanz

Abstract

We consider dependence structures in multivariate time series that are characterized by deterministic trends. Results from spectral analysis for stationary processes are extended to deterministic trend functions. A regression cross covariance and spectrum are defined. Estimation of these quantities is based on wavelet thresholding. The method is illustrated by a simulated example and a three-dimensional time series consisting of ECG, blood pressure and cardiac stroke volume measurements.

Key words: Nonparametric trend estimation, cross spectrum, wavelets, regression spectrum, phase, threshold estimator.

1 Introduction

We consider dependence structures in multivariate time series that are due to similarities in underlying deterministic trends. Suppose that we observe a multivariate time series $\mathbf{Y}(i) = (Y_1(i), ..., Y_p(i))^T$, $(i = 1, \dots, n)$. In classical spectral analysis, a time series and its autocorrelations are decomposed into sinusoidal components. Grenander and Rosenblatt (1957) extended the idea of spectral decomposition to parametric regression with deterministic explanatory variables. They show that consistency and efficiency of least squares estimators depend on the regression spectrum. For further results on the interplay between regression spectrum and spectral properties of the stochastic part see e.g. Yajima (1988).

In the present paper, spectral analysis of regression functions is extended to multivariate *nonparametric* trend functions that are estimated by wavelet thresholding. The definitions of the regression cross covariance and regression cross spectrum are adapted to this context. Asymptotic properties of estimators of these quantities are derived. Specifically, the paper is organized as follows. Basic definitions are given in section 2. Estimation of the regression spectrum, the regression cross covariances and their asymptotic distribution are considered in section 3. Algorithmic issues and a data example are discussed in section 4. Final remarks in section 5 conclude the paper. Proofs are given in the appendix.

2 Definition of the regression cross covariance and spectrum

2.1 Cross covariance and correlation

Assume that the observed time series $\mathbf{Y}(j) = (Y_1(j), ..., Y_p(j))^T$ is of the form

$$\mathbf{Y}(j) = \mathbf{f}(t_j) + \boldsymbol{\epsilon}(j), \tag{1}$$

where $t_j = j/n$ $(j = 1, \ldots, n)$, $\mathbf{f}(t) = (f_1(t), ..., f_p(t))^T \in \mathbb{C}^p$ $(t \in \mathbb{R})$ is a multivariate deterministic trend function and $\boldsymbol{\epsilon}(j)$ is zero mean stationary noise. We will assume that $f_r(t)$ is Lebesque measurable and $\int_0^1 |f_r(t)|^2 dt < \infty$, i.e. $f_r \in L^2$ where $L^2 = L^2(\mathbb{C})$ denotes the space of complex-valued functions that are square integrable on $[0, 1]$. For $f_r, f_s \in L^2$, we define

$$< f_r, f_s >= \int_0^1 f_r(t)\overline{f_s(t)}dt$$

Note that $<,>$ is a nonnegative sesquilinear form. If we restrict attention to functions that are periodic with period 1, then $<,>$ is a scalar product and the corresponding space $L^2[0, 1]$ is a Hilbert space. Also note that, since $||f_r||^2 =< f_r, f_r >< \infty$, the mean vector

$$\mathbf{m}(f) = \int_0^1 \mathbf{f}(t)dt$$

is well defined and finite. Without loss of generality (and since $\mathbf{m}(f)$ can easily be estimated and subtracted from the data), we will from now on assume $\mathbf{m}(f) = 0$. The corresponding space of functions will be denoted by $L_o^2[0,1] = \{f \in L^2[0,1] : \mathbf{m}(f) = 0\}$.

We first define the autocorrelation function of $\mathbf{f}(t) \in L_o^2$. Grenander and Rosenblatt (1957) introduced a definition of cross- and autocorrelation $\rho_{rs}(u)$ between parametric regression functions $\phi_r(t), \phi_s(t)$ with $t \in \mathbb{N}$ (also see Priestley 1989, chapter 7). In contrast, here, we consider nonparametric regression functions $f_j(t) \in L_o^2$ that are observed on an increasingly fine grid $t_1, t_2, ..., t_n$ of $t-$values in $[0, 1]$. A natural modification of the definition by Grenander and Rosenblatt definition is therefore

$$\gamma_{rs}(u) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} f_r(t_j + u)\overline{f_s(t_j)}. \ (u \in [-1, 1])$$

and

$$\rho_{rs}(u) = \frac{\gamma_{rs}(u)}{\sqrt{\gamma_r(0)\gamma_s(0)}}.$$

This leads to

Definition 1 *Let $\mathbf{f}(t) = (f_1(t), ..., f_p(t))^T \in \mathbb{C}^p$ $(t \in \mathbb{R})$ be a p-dimensional deterministic function as defined above, and such that, for each $r \in \{1, 2, ..., p\}$, $f_r \in L_o^2$ and $||f_r(t)|| > 0$. Then the regression (cross-)covariance function $\mathbf{\Gamma}(u) = [\gamma_{rs}(u)]_{r,s=1,...,p}$ and the regression (cross-)correlation function $\mathbf{R}(u) = [\rho_{rs}(u)]_{r,s=1,...,p}$ of $\mathbf{f}(t)$ are defined by*

$$\gamma_{rs}(u) =< f_r(\cdot + u), f_s >= \int_0^1 f_r(t+u)\overline{f_s(t)}dt$$

and

$$\rho_{rs}(u) = \frac{\gamma_{rs}(u)}{\sqrt{\gamma_r(0)\gamma_s(0)}}. \tag{2}$$

The Hermitian property and non-negative definiteness of $\rho_{rs}(\cdot)$ are obtained in the following

Proposition 1 *The regression cross correlation function $\rho_{rs}(u)$ defined in (2) is Hermitian and non-negative definite.*

In practice, γ_{rs} and ρ_{rs} have to be estimated, since they depend on the unobservable function $\mathbf{f}(t)$.

2.2 The spectrum

The function $\mathbf{f}(t)$ is observed for time points in the interval $[0,1]$ only. To extrapolate $\mathbf{f}$ beyond the unit interval, we will assume that $\mathbf{f}(t)$ can be decomposed into a nonperiodic "long-term" trend component $\boldsymbol{\mu}(t)$ ($t \in \mathbb{R}$) and a component $\boldsymbol{\alpha}(t)$ with $\boldsymbol{\alpha}(\text{t+T}) = \boldsymbol{\alpha}(t)$ for some $T \leq \frac{1}{2}$. To simplify presentation, it will be assumed throughout the paper that $\boldsymbol{\mu} \equiv 0$, or $\boldsymbol{\mu}$ has been estimated and removed from the data. Thus,

$$\mathbf{f}(t) = \boldsymbol{\alpha}(t)$$

with $\boldsymbol{\alpha}$ periodic with period $T \leq \frac{1}{2}$.

Remark 1 *The assumption of strict periodicity can be replaced by local periodicity, allowing the periodic shape of $\boldsymbol{\alpha}$ to change smoothly in time (see e.g. Heiler and Feng 2000). In this case, $\gamma_{rs}(u)$ can be approximated by extrapolating $\boldsymbol{\alpha}$, e.g. using local trigonometric polynomials (Heiler and Feng 2000) together with nonparametric extrapolation (see e.g. Beran and Ocker 1999).*

Consider now the characterization of $\mathbf{f}$ in the frequency domain. For $t \in [0,1]$, and $f_r \in L_o^2$ such that $\int_0^1 |f_r(t)|dt < \infty$ we may write

$$\mathbf{f}(t) = \sum_{j=-\infty}^{\infty} \mathbf{a}(j) e^{i2\pi jt}$$

where $\mathbf{a}(j) = (a_1(j), a_2(j), ..., a_p(j))^T \in \mathbb{C}^p$ are given by

$$a_r(j) = < f_r, e^{i2\pi j\cdot} > = \int_0^1 f_r(t) e^{-i2\pi jt} dt.$$

Note that

$$\mathbf{m}(f) = \mathbf{a}(0) = 0$$

and Parseval's equation yields

$$||f_r||^2 = \sum_{j=-\infty}^{\infty} |a_r(j)|^2.$$

For the covariance function we then have

$$\boldsymbol{\Gamma}(u) = \boldsymbol{\Gamma}_\alpha(u)$$

where $\mathbf{\Gamma}_\alpha = [\gamma_{\alpha;rs}]_{r,s=1,\ldots,p}$ is a $p \times p$ matrix with

$$\gamma_{\alpha;rs}(u) =< \alpha_r(\cdot + u), \alpha_s >$$

More explicitly, we have

$$\begin{aligned}\mathbf{\Gamma}_\alpha(u) &= \int_0^1 \boldsymbol{\alpha}(t+u)\overline{\boldsymbol{\alpha}^T(t)}dt \\ &= \sum_{j=-\infty}^{\infty} e^{i2\pi ju}\mathbf{a}(j)\overline{\mathbf{a}^T(j)}\end{aligned}$$

and we may introduce the following

Definition 2 *The sequence of $p \times p$ matrices $\mathbf{H}(j) = [h_{rs}(j)]_{r,s=1,\ldots,p}$ ($j \in \mathbb{Z}$) defined by*

$$\mathbf{H}(j) = \mathbf{a}(j)\overline{\mathbf{a}^T(j)}$$

is called regression spectrum of $\boldsymbol{\alpha}$.

By definition we have the following relationship between regression spectrum and covariances:

$$\mathbf{\Gamma}_\alpha(u) = \sum_{j=-\infty}^{\infty} \mathbf{H}(j)e^{i2\pi ju}$$

and

$$\mathbf{H}(j) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-i2\pi ju}\mathbf{\Gamma}_\alpha(u)du,$$

where $\int_{-1/2}^{1/2} |\mathbf{\Gamma}_\alpha(u)|du < \infty$. Writing $\mathbf{H}(j)$ in polar representation, the contribution of frequency j can also be expressed in relative terms as follows.

Definition 3 *Let $\mathbf{H}(j) = [h_{rs}(j)]_{r,s=1,\ldots,p}$ be defined as above. Then, $\tilde{\mathbf{H}} = [\tilde{h}_{rs}]_{r,s=1,\ldots,p}$ with*

$$\tilde{h}_{rs}(j) = \frac{h_{rs}(j)}{\sqrt{\gamma_{\alpha;rr}(0)\cdot\gamma_{\alpha;ss}(0)}} = \frac{|a_r(j)\overline{a_s(j)}|}{\sqrt{\gamma_{\alpha;rr}(0)\cdot\gamma_{\alpha;ss}(0)}}\exp(i\phi_{rs}(j))$$

is called the standardized regression spectrum of $\boldsymbol{\alpha}$*,*

$$\kappa_{rs}(j) = \frac{|h_{rs}(j)|}{\sqrt{\gamma_{\alpha;rr}(0) \cdot \gamma_{\alpha;ss}(0)}} = \frac{|h_{rs}(j)|}{||\alpha_r|| \cdot ||\alpha_s||}$$
$$= \frac{|a_r(j)\overline{a_s(j)}|}{\sqrt{\sum_l |a_r(l)|^2 \sum_m |a_s(m)|^2}}$$

is the relative spectral modulus and $\phi_{rs}(j)$ *the phase shift at frequency* j*.*

Remark 2 *Note that, in contrast to coherence for stochastic processes, the standardization* $\sqrt{\gamma_{\alpha;rr}(0) \cdot \gamma_{\alpha;ss}(0)}$ *is not frequency dependent. The reason is that, for a deterministic signal, no frequency dependent variances can be observed. Alternatively, one may consider the classical squared coherency function* $|h_{rs}(j)|^2/(|h_{rr}(j)||h_{ss}(j)|)$*. However, this quantity is either* 0 *or* 1*. In contrast,* $\kappa_{rs}(j)$ *defined above can assume any number between* 0 *and* 1*, thus giving a relative measure of the contribution of frequency* j *to the cross-covariance.*

Example 1 *Suppose that* f_s *is a shifted version of* f_r*, i.e.*

$$f_s(t) = c \cdot f_r(t + \Delta)$$

for some $\Delta, c \in \mathbb{R}$*. Then*

$$f_s(t) = \sum a_s(j) \exp(i2\pi jt) = c \sum a_r(j) \exp(i2\pi j(t + \Delta))$$

with

$$a_s(j) = ca_r(j) \exp(i2\pi j\Delta).$$

Hence

$$\gamma_{rs}(u) = c \sum |a_r(j)|^2 \exp(i2\pi j(u - \Delta)).$$

and

$$\tilde{h}_{rs}(j) = \frac{|a_r(j)|^2}{\sum_l |a_r(l)|^2} \cdot \exp(-i2\pi\Delta j) \;\; (j \in \mathbb{Z} \setminus \{0\})$$

Hence, for all integer frequencies $j \neq 0$

$$|h_{rs}(j)| = \frac{|a_r(j)|^2}{\sum_l |a_r(l)|^2}$$

and the phase-shift

$$\phi_{rs}(j) = -2\pi\Delta j$$

is a linear function of the shifting parameter Δ. *Thus,* $|h_{rs}(j)|$ *is equal to the relative contribution of frequency* j *to total energy* $||f_r||^2$ *of* f_r.

3 Estimation

3.1 General considerations

The definitions above suggest the following approach to analyzing an observed multivariate time series $\mathbf{Y}(j) = \mathbf{f}(t_j) + \boldsymbol{\epsilon}(j)$. In a first step, the function $\mathbf{f}$ is estimated by a suitable nonparametric method. In a second step, $\boldsymbol{\alpha}$ is estimated by eliminating the mean $\mathbf{m}$ and the nonperiodic component $\boldsymbol{\mu}(t)$. The regression spectrum of $\boldsymbol{\alpha}(t)$ can now be analyzed based on the resulting series

$$\tilde{\mathbf{Y}}(j) = \mathbf{Y}(j) - \hat{\mathbf{m}} - \hat{\boldsymbol{\mu}}(t)$$

The issue of estimating $\boldsymbol{\mu}(t)$ optimally in the given context is beyond the scope of this paper and will be considered elsewhere. Note, however, that in some applications, the size of $\boldsymbol{\mu}$ is negligible compared to $\boldsymbol{\alpha}$. For instance, for high frequency financial data, the dominating feature in the deterministic part is likely to be a (local) seasonal periodicity of one day. Similar comments apply to physiological time series such as the heart beat data considered in section 4.

In this section we consider estimation of $\mathbf{f}$ and the resulting estimation of the regression cross covariance $\boldsymbol{\Gamma}_\alpha$ and the regression spectrum $\mathbf{H}$.

3.2 Trend estimation

Since $\boldsymbol{\Gamma}_\alpha$ and the regression spectrum $\mathbf{H}$ are functionals of $\mathbf{f}$ (which we assume to be equal to $\boldsymbol{\alpha}$ - see remark at the end of the previous section), we first consider nonparametric estimation of the trend function. Methods based on wavelets are known to have attractive features, such as general applicability to L^2-functions and localization in time and frequency (see e.g. Daubechies 1999, Percival and Walden 2000). We therefore consider estimation of $\mathbf{f}$

via wavelet analysis. Let $\phi(\cdot)$ and $\psi(\cdot)$ be the father and mother wavelet respectively, i.e. $\phi(\cdot)$, $\psi(\cdot) \in L^2(\mathbb{R})$ and the set of functions

$$\{\phi_{l,k}(x),\ \ \psi_{j,k}(x),\ \ j,k,l \in \mathbb{Z},\ j \geq l\}$$

with

$$\begin{aligned}\phi_{l,k}(x) &= 2^{\frac{l}{2}}\phi(2^l x - k),\\ \psi_{j,k}(x) &= 2^{\frac{j}{2}}\psi(2^j x - k),\ \ k,j \in \mathbb{Z},\end{aligned}$$

form a basis in $L^2(\mathbb{R})$. The wavelet series expansion of a univariate function $f \in L^2(\mathbb{R})$ is then given by

$$f(x) = \sum_k \alpha_{l,k}\phi_{l,k}(x) + \sum_{j\geq l}^{\infty}\sum_k \beta_{j,k}\psi_{j,k}(x) \tag{3}$$

for almost all x, where

$$\begin{aligned}\alpha_{l,k} &= \int \phi_{l,k}(x)f(x)dx,\\ \beta_{j,k} &= \int \psi_{j,k}(x)f(x)dx.\end{aligned}$$

For the univariate model

$$Y(j) = f(t_j) + \epsilon(j)$$

with $f \in L^2[0,1]$ and $\epsilon(j)$ stationary with mean 0 and variance σ^2, Nason (1996) and Johnston and Silverman (1997) consider the estimator

$$\hat{f}(t) = \sum_k \hat{\alpha}_{l,k}\phi_{l,k}(t) + \sum_{j\geq l}^{J_n}\sum_k \hat{\beta}_{j,k}\psi_{j,k}(t), \tag{4}$$

where

$$\hat{\alpha}_{l,k} = \frac{1}{n}\sum_{u=1}^{n}\phi_{l,k}(t_u)Y(u) \tag{5}$$

and

$$\hat{\beta}_{j,k} = \frac{1}{n}\sum_{u=1}^{n}\psi_{j,k}(t_u)Y(u), \tag{6}$$

for some J_n. Under suitable regularity assumptions on f, on the wavelet basis and moment conditions on $\epsilon(j)$, the parameter estimators $\hat{\alpha}_{l,k}$ and $\hat{\beta}_{j,k}$ are asymptotically normal and unbiased (see Brillinger 1994,1996). These results carry over to the trend estimator $\hat{f}$, for suitable choices of J_n. Brillinger (1994, 1996) shows, in particular,

$$\text{var } (2^{-\frac{j}{2}}\hat{\alpha}_{j,k}) = \text{var } (2^{-\frac{j}{2}}\hat{\beta}_{j,k}) = 2\pi h_\epsilon(0)2^{-j}n^{-1} + O(n^{-2})$$
$$\text{cov } (2^{-\frac{j}{2}}\hat{\beta}_{j,k}, 2^{-\frac{j'}{2}}\hat{\beta}_{j',k'}) = O(n^{-2}),$$

for $(j,k) \neq (j',k')$. Analogously, cov $(2^{-\frac{l}{2}}\hat{\alpha}_{l,k}, 2^{-\frac{j'}{2}}\hat{\alpha}_{l',k'}) = O(n^{-2})$ for $(l,k) \neq (l',k')$ and cov $(2^{-\frac{l}{2}}\hat{\alpha}_{l,k}, 2^{-\frac{j'}{2}}\hat{\beta}_{j',k'}) = O(n^{-2})$ for all l, k, j', k'. Here $h_\epsilon(\cdot)$ denotes the spectral density of $\epsilon(j)$. The error terms are uniform in j, j', k, k', l. In wavelet thresholding, noise is removed by shrinking wavelet coefficients towards zero at a suitable rate (see e.g. Donoho and Johnstone 1994, 1995, Abramovich, Sapatinas and Silverman 1998). Here, a hard thresholding method will be applied, i.e. each $\hat{\beta}_{j,k}$ is multiplied by $\hat{w}_{j,k} := \mathbf{1}_{\{|\hat{\beta}_{j,k}| \geq \sqrt{\text{var }(\hat{\beta}_{j,k})}\lambda_j\}}$ such that (4) changes to

$$\hat{f}(t) := \sum_k \hat{\alpha}_{l,k}\phi_{l,k}(t) + \sum_{j\geq l}^{J_n}\sum_k \hat{w}_{j,k}\hat{\beta}_{j,k}\psi_{j,k}(t). \qquad (7)$$

To conclude this section, we state asymptotic results for $\hat{\mathbf{f}}$ that will be needed to derive properties of $\hat{\rho}_{rs}$ and $\hat{h}_{rs}$. The following two assumptions will be used:

(A1) The mother wavelet $\psi(\cdot)$ and the father wavelet $\phi(\cdot)$ are of bounded variation and have compact support. Dilation and translation result in an orthonormal basis for a finite interval containing $[0,1]$.

(A2) For each j, $1 \leq j \leq p$, the univariate functions $f_j(\cdot)$ $(1 \leq j \leq p)$ are bounded, of bounded variation on $[0,1]$ and vanish outside the interval. In addition, only a finite number of coefficients in the wavelet representation is non-zero.

Brillinger (1994) considers shrunken wavelet estimators for univariate $f \in L^2[0,1]$ under the assumptions (A1) and (A2) and shows that, almost everywhere in t, finite collections of $\hat{f}(t)$ are asymptotically normal with mean $f(t)$. These results can easily be carried over to multivariate trend functions:

Lemma 1 *Assume model (1) together with (A1), (A2) and assumptions on the cumulants of* $\boldsymbol{\epsilon}(i)$ *as given in Brillinger (1994). For* $1 \le r \le p$, *let*

$$\hat{\alpha}^{(r)}_{l,k} = \frac{1}{n} \sum_{u=1}^{n} \phi_{l,k}(t_u) Y_r(u), \tag{8}$$

$$\hat{\beta}^{(r)}_{j,k} = \frac{1}{n} \sum_{u=1}^{n} \psi_{j,k}(t_u) Y_r(u), \tag{9}$$

$$\hat{w}^{(r)}_{j,k} := \mathbf{1}_{\{|\hat{\beta}^{(r)}_{j,k}| \ge \sqrt{\mathrm{var}\ (\hat{\beta}^{(r)}_{j,k})} \lambda_j\}},$$

$J_n \to \infty$, $n 2^{-J_n/2} \to \infty$, λ_j *such that* $2^{\frac{j}{2}} \lambda_j = o(n^{1/2})$ $(j = l, l+1, \ldots, J_n)$ *and*

$$\sum_{j \ge l}^{J_n} 2^{\frac{j}{2}} \exp(-\lambda_j^2/(1+\eta)2) = o(1)$$

for some $\eta > 0$. *Define*

$$\hat{f}_r(t) := \sum_k \hat{\alpha}^{(r)}_{l,k} \phi_{l,k}(t) + \sum_{j \ge l}^{J_n} \sum_k \hat{w}^{(r)}_{j,k} \hat{\beta}^{(r)}_{j,k} \psi_{j,k}(t), \tag{10}$$

$1 \le r \le p$. *Also denote by* $\mathbf{h}_\epsilon = (h_{\epsilon;rs})_{1 \le r,s \le p}$ *the matrix of cross spectral densities between* $\{\epsilon_r(i), i \in \mathbb{N}\}$ *and* $\{\epsilon_s(i), i \in \mathbb{N}\}$. *Then, almost everywhere in* $t \in [0,1]$, *finite collections of* $\sqrt{n}(\hat{f}_r(t) - f_r(t))$ *are asymptotically normal with mean zero. Moreover,*

$$\gamma_f(x, y; r, s) = \mathrm{cov}\ (\hat{f}_r(x), \hat{f}_s(y)) = \frac{2\pi h_{\epsilon;rs}(0)}{n} \sum_k \phi_{l,k}(x)\phi_{l,k}(y)$$

$$+ \frac{2\pi h_{\epsilon;rs}(0)}{n} \sum_{j,k}^{J_0} w^{(r)}_{j,k} w^{(s)}_{j,k} \psi_{j,k}(x)\psi_{j,k}(y) + r(n),$$

where $w^{(i)}_{j,k} = \mathbf{1}\{\beta^{(i)}_{j,k} \ne 0\}$, $J_0 = J_0(r,s)$ *is the largest common integer such that* $w^{(r)}_{j,k} w^{(s)}_{j,k} \ne 0$ *for some* $j = J_o$, *and* $r(n) = O(2^{2J_n} n^{-2})$.

Remark 3 *One possible threshold that satisfies the assumptions of lemma 1 is given by* $\lambda_j = \sqrt{2\log(2^{-j}n)}$.

Remark 4 *Note that additional asymptotic properties for the estimators of the wavelet coefficients are also easily carried over to the multivariate case. For instance,* cov $(\hat\beta^{(r)}_{j,k}, \hat\beta^{(s)}_{j',k'}) = O(n^{-2})$ *for* $(j,k) \neq (j',k')$.

Remark 5 *If* var $(\beta^{(r)}_{j,k})$ *is unknown, then the variance has to be estimated. See Brillinger (1994) for consistent estimation of the variance and asymptotic properties.*

Remark 6 *Lemma 1 implies* var $(\hat f_r(t)) = O(n^{-1})$*, uniformly in* t*, so that* $\hat f_r(t)$ *is a asymptotically consistent estimator of* $f_r(t)$*, i.e.* $\hat f_r(t) \to f_r(t)$ *in probability for almost all* t *in* $[0,1]$.

3.3 Estimation of the regression cross covariance

3.3.1 Consistency

By assumption $\mathbf{m}(f) =< \mathbf{f}, 1 >= 0$, and $\boldsymbol{\mu}(t) = 0$ so that $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}_\alpha$. Given $\hat{\mathbf{f}}$, the cross-covariance $\boldsymbol{\Gamma}_\alpha$ can be estimated by

$$\hat{\boldsymbol{\Gamma}}_\alpha(u) = \hat{\boldsymbol{\Gamma}}(u) = \int_0^1 \hat{\mathbf{f}}(t+u)\overline{\hat{\mathbf{f}}^T(t)}dt \tag{11}$$

and the regression spectrum by

$$\hat{\mathbf{H}}(j) = \int_{-1/2}^{1/2} \exp(-i2\pi ju)\hat{\boldsymbol{\Gamma}}_\alpha(u)du.$$

Consistency of $\hat{\boldsymbol{\Gamma}}(u)$ is given by

Theorem 1 *Under the assumptions of lemma 1,* $\hat{\boldsymbol{\Gamma}}(u)$ *converges in probability to* $\boldsymbol{\Gamma}(u)$ *such that* $\hat{\boldsymbol{\Gamma}}(u) - \boldsymbol{\Gamma}(u) = O_p(n^{-1/2})$.

The proof is based on the following

Lemma 2 *Let* f_r *be estimated as in Lemma 1, and denote the estimation error by*

$$\widetilde{\epsilon}^{(r)}_n(t) = \hat f_r(t) - f_r(t).$$

Then,

$$\int_{[0,1]} \widetilde{\epsilon}_n^{(r)}(t)dt \to 0 \tag{12}$$

and

$$\int_{[0,1]} |\widetilde{\epsilon}_n^{(r)}(t)|dt \to 0 \tag{13}$$

almost surely, and both terms are of order $O_p(n^{-1/2})$*.*

3.3.2 Asymptotic normality

First, we derive the asymptotic covariance function of $\hat{\gamma}_{rs}(u)$.

Theorem 2 *For* $u, v \in [-1, 1]$,

$$\lim_{n\to\infty} n\,cov(\hat{\gamma}_{rs}(u), \hat{\gamma}_{rs}(v)) = \sigma_{rs}(u, v)$$

where σ_{rs} is finite and can be written as

$$\sigma_{rs}(u,v) = 2\pi h_{\epsilon;ss}(0)\left[\sum_k \int f_r(x+u)\phi_{l,k}(x)dx \int f_r(y+v)\phi_{l,k}(y)dy\right.$$

$$\left. + \sum_{l,k} w_{l,k}^{(s)} \int f_r(x+u)\psi_{l,k}(x)dx \int f_r(y+v)\psi_{l,k}(y)dy\right]$$

$$+ 2\pi h_{\epsilon;rs}(0)\left[\sum_k \int f_s(x)\phi_{l,k}(x+u)dx \int f_r(y+v)\phi_{l,k}(y)dy\right.$$

$$\left. + \sum_{l,k} w_{l,k}^{(r)} w_{l,k}^{(s)} \int f_s(x)\psi_{l,k}(x+u)dx \int f_r(y+v)\psi_{l,k}(y)dy\right]$$

$$+ 2\pi h_{\epsilon;sr}(0)\left[\sum_k \int f_r(x+u)\phi_{l,k}(x)dx \int f_s(y)\phi_{l,k}(y+v)dy\right.$$

$$\left. + \sum_{l,k} w_{l,k}^{(r)} w_{l,k}^{(s)} \int f_r(x+u)\psi_{l,k}(x)dx \int f_s(y)\psi_{l,k}(y+v)dy\right]$$

$$+ 2\pi h_{\epsilon;rr}(0)\left[\sum_k \int f_s(x)\phi_{l,k}(x+u)dx \int f_s(y)\phi_{l,k}(y+v)dy\right.$$

$$\left. + \sum_{l,k} w_{l,k}^{(r)} \int f_s(x)\psi_{l,k}(x+u)dx \int f_s(y)\psi_{l,k}(y+v)dy\right] + O(n^{-1/2}).$$

In order to obtain the asymptotic distribution of $\hat{\mathbf{\Gamma}}$ the following additional condition on $\boldsymbol{\epsilon}(i)$ will be used.

(A3) For $1 \le r, s \le p$, define $\epsilon_t^* := \epsilon_r(t),\ t = 1, \ldots, n$, and $\epsilon_t^* = \epsilon_s(t-n),\ t = n+1, \ldots, 2n$. Let $\{\mathcal{F}_t,\ t = 1, \ldots, n\}$ be a non-decreasing sequence of σ-fields of $\mathcal{F}$ sets and let the sequence $(\epsilon_t^*, \mathcal{F}_t, t = 1, \ldots, n)$ be a square-integrable martingale difference array with constant variance and $E(\epsilon_t^{*2}|\mathcal{F}_{t-1}) = E(\epsilon_0^{*2})$.

Theorem 3 *Let*

$$\xi_n = n^{\frac{1}{2}}[\hat{\gamma}_{rs}(u_1) - \gamma_{rs}(u_1), \hat{\gamma}_{rs}(u_2) - \gamma_{rs}(u_2), ..., \hat{\gamma}_{rs}(u_k) - \gamma_{rs}(u_k)]^T$$

Then, under (A3) and the assumptions of theorem 1,

$$\xi_n \xrightarrow{d} N(0, \Sigma)$$

where

$$\Sigma = [\sigma_{rs}(u_i, u_j)]_{r,s=1,\dots,p}$$

and "$\xrightarrow{d}$" denotes convergence in distribution.

The proof is based on the following lemmas.

Lemma 3 *Assume model (1) and the assumptions of theorem 1. Then, for each $s \in \{1, \dots, p\}$ and for almost all $x \in [0, 1]$,*

$$\begin{aligned}
\left(\hat{f}_s - f_s\right)(x) = & \frac{1}{n}\sum_{u=1}^{n} \epsilon_s(u) \sum_k \phi_{l,k}(u/n)\phi_{l,k}(x) \\
& + \frac{1}{n}\sum_{u=1}^{n} \epsilon_s(u) \sum_{j \geq l}^{J_n} \sum_k \hat{w}_{j,k}^{(s)} \psi_{j,k}(u/n)\psi_{j,k}(x) \\
& + r^{(s)}(n),
\end{aligned}$$

where $r^{(s)}(n) = O_p(n^{-1})$.

Lemma 4 *Under the assumptions of theorem 3, we have, for each pair (r, s), $1 \leq r, s \leq p$, and each $u \in [-1, 1]$*

$$\sqrt{n}(\hat{\gamma}_{rs}(u) - \gamma_{rs}(u)) \xrightarrow{d} \mathcal{N}(0, \sigma_{rs}(u, u)),$$

where $\sigma_{rs}(u, u)$ is given in theorem 2.

Theorem 3 can be extended to a functional limit theorem.

Theorem 4 *Let P_n be the probability distribution of $\sqrt{n}(\hat{\gamma}_{rs}^{(n)}(u) - \gamma_{rs}(u))$ in $C[-1, 1]$, where $C[-1, 1]$ is equipped with the uniform topology defined by the metric $d(f, g) = \sup_{-1 \leq t \leq 1} |f(t) - g(t)|$. Then P_n converges asymptotically in $C[-1, 1]$ (in the metric d) to the probability distribution P of a Gaussian process where the finite-dimensional distributions are given in theorem 3.*

3.4 Estimation of the regression cross spectrum

3.4.1 Asymptotic normality

Theorem 4 together with the continuous mapping theorem (see e.g. Pollard 1984) lead to

Theorem 5 *Under the assumptions of theorem 4, the vector*

$$\zeta_n = \sqrt{n}[\hat{h}_{rs}(j_1) - h_{rs}(j_1), ..., \hat{h}_{rs}(j_m) - h_{rs}(j_m)]^T$$

converges in distribution to an $m-$dimensional zero mean normal vector with covariance matrix

$$ncov\left(\hat{h}_{rs}(j_k), \hat{h}_{rs}(j_l)\right) = \int\int \exp\left\{-i2\pi(j_k u_1 - j_l u_2)\right\} \sigma_{rs}(u_1, u_2) du_1 du_2$$

where $\sigma_{rs}(u_1, u_2)$ is defined in theorem 2.

3.4.2 Estimation of amplitude and phase spectrum

Theorem 5 shows that finite vectors of the regression cross spectrum $\hat{h}_{rs}(j)$ converge to a complex-valued normal random variable. Let $c_{rs}(j) = \mathfrak{Re}(h_{rs}(j))$ and $q_{rs}(j) = \mathfrak{Im}(h_{rs}(j))$ so that $h_{rs}(j) = c_{rs}(j) + iq_{rs}(j)$. Then we have estimators

$$\hat{c}_{rs}(j) = \frac{1}{2}(\hat{h}_{rs}(j) + \hat{h}_{sr}(j)) \tag{14}$$

and

$$\hat{q}_{rs}(j) = \frac{1}{2i}(\hat{h}_{rs}(j) - \hat{h}_{sr}(j)). \tag{15}$$

Due to theorem 5, the vector

$$\sqrt{n}[\hat{c}_{rs}(j) - c_{rs}(j), \hat{q}_{rs}(j) - q_{rs}(j)]^T$$

converges in distribution to a bivariate normal variable with mean 0 and asymptotic covariance matrix

$$\Sigma(j) = \begin{pmatrix} \Sigma_{cc}(j) & \Sigma_{cq}(j) \\ \Sigma_{qc}(j) & \Sigma_{qq}(j) \end{pmatrix}.$$

The asymptotic distribution of the amplitude and phase spectrum then follows by straighforward calculations. Let $\kappa^*_{rs}(j)$ be the non-normalized spectral modulus. Then we have

Corollary 1 *Let* $\kappa^*_{rs}(j) = |h_{rs}(j)|$ *and*

$$\hat{\kappa}^*_{rs}(j) = (\hat{c}_{rs}(j)^2 + \hat{q}_{rs}(j)^2)^{\frac{1}{2}}.$$

Then

$$\sqrt{n}(\hat{\kappa}^*_{rs}(j) - \kappa^*_{rs}(j)) \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\kappa;rs}(j)),$$

where

$$\sigma^2_{\kappa;rs}(j) = \frac{1}{c^2_{rs}(j) + q^2_{rs}(j)} \left(c^2_{rs}(j)\Sigma_{cc}(j) + q^2_{rs}(j)\Sigma_{qq}(j) + 2c_{rs}(j)q_{rs}(j)\Sigma_{cq}(j) \right). \tag{16}$$

Corollary 2 *Let* $\phi_{rs}(j) = \arg h_{rs}(j)$ *and*

$$\hat{\phi}_{rs}(j) = \arg(\hat{c}_{rs}(j) + i\hat{q}_{rs}(j)) \in (-\pi, \pi].$$

Then

$$\sqrt{n}(\hat{\phi}_{rs}(j) - \phi_{rs}(j)) \xrightarrow{d} \mathcal{N}(0, \sigma^2_{\phi;rs}(j)),$$

where

$$\sigma^2_{\phi;rs}(j) = \frac{1}{(c^2_{rs}(j) + q^2_{rs}(j))^2} \left(q^2_{rs}(j)\Sigma_{cc}(j) + c^2_{rs}(j)\Sigma_{qq}(j) - 2c_{rs}(j)q_{rs}(j)\Sigma_{cq}(j) \right). \tag{17}$$

4 Algorithm and data examples

4.1 General considerations

Consider example 1 with two functions that are shifted versions of each other. In this case, the phase spectrum consists of a straight line modulo 2π, with the slope being proportional to the shift. Discontinuities occur at frequencies where $\phi_{rs}(j)$ crosses $-\pi$ or π. More generally, if we are given a plot of a phase spectrum between two deterministic functions the detection of any linear or piecewise linear curve may be interpreted as a constant lag over this particular range of frequencies. However, if $\mathbf{f}$ and $\mathbf{H}(j)$ have to be estimated, we have a superposition of the linear structure of the phase spectrum with the phase spectrum of the noise component. Comparing (16) with (17) we see that $\sigma^2_{\kappa;rs}(j)$ and $\sigma^2_{\phi;rs}(j)$ essentially differ by the factor $\kappa^*_{rs}(j)^{-2}$. Therefore, $\sigma^2_{\phi;rs}(j)$ will be relatively large compared to $\sigma^2_{\kappa;rs}(j)$ for all frequencies where the corresponding spectral modulus is small, and it will be relatively small where the spectrum modulus is large. Thus, in general, the phase spectrum will look more erratic than the amplitude spectrum, and estimation of common frequencies in a multivariate trend function might be easier than estimation of the phase shift. A pure visual inspection of the phase spectrum may not be sufficient to detect linear structure. In the next section, we propose a simple algorithm that takes this into account. In a first step, frequencies are identified where the amplitude spectrum is significantly larger than 0. In a second step, the phase shift is estimated using theses frequencies only.

4.2 Algorithm

A data-driven algorithm for estimating the regression cross spectrum can be defined as follows.

1. Choose a wavelet basis $\{\phi_{l,k}(x),\ \ \psi_{j,k}(x),\ \ j,k,l \in \mathbb{Z},\ j \geq l\}$ and thresholds $\lambda_j,\ j = 1,\ldots,J_n$, according to equation (10), and estimate f_r, $r = 1,\ldots,p$. This step can be carried out, for instance, using the function WAVESHRINK in the S-Plus wavelet module.

2. Apply the fast Fourier transform to obtain
$$\hat{a}_r(j) = \frac{1}{n}\sum_{t=1}^{n} \hat{f}_r(t/n)e^{-i2\pi jt/n}, \quad r = 1,\ldots,p,$$
and calculate the regression cross spectrum
$$\hat{h}_{rs}(j) = \hat{a}_r(j)\overline{\hat{a}_s(j)}$$
and estimators of the amplitude and phase spectrum (section 3.4.2).

3. Estimate the cross spectrum $h_{\epsilon,rs}(0)$ of $\boldsymbol{\epsilon}(i)$ from the estimated residuals $\hat{\epsilon} = \hat{\mathbf{Y}} - \hat{\mathbf{f}}$.

4. Use equation (16) to calculate $\sigma^2_{\kappa;rs}(j)$ and determine the set
$$J^* = \{j:\ \hat{\kappa}^*_{rs}(j) > c_{\kappa;rs} \cdot \sqrt{\sigma^2_{\kappa;rs}(j)}\}$$
for a suitably chosen $c_{\kappa;rs} \in \mathbb{R}$.

5. To estimate the phase shift, apply a local robust regression to the points $\{(j, \phi_{rs}(j)) : j \in J^*\}$, taking into account possible 2π-jumps.

4.3 Examples

The application of the asymptotic results and the practical performance of the algorithm are illustrated by two data examples.

4.3.1 Simulated example

The trend function $\mathbf{f}(x) = (f_1(x), f_2(x))^T$ $(x \in [0,1])$ is defined by

$$f_1(x) = -\sin(4\pi x) - \sin(10\pi x) - \sin(20\pi x) - \sin(30\pi x) - \sin(51\pi x),$$

$$f_2(x) = f_1(x + \Delta),$$

and

$$\Delta = .03125.$$

Moreover, $\epsilon_1(i)$, $\epsilon_2(i)$ are iid $N(0, \sigma_\epsilon^2)$ with $\sigma_\epsilon^2 = 4$ and ϵ_1, ϵ_2 independent of each other. Figure 1 shows the simulated series (figures 1a,b), the true trend functions f_1 and f_2 (figures 1c,d), and the true regression cross spectrum in terms of the amplitude (figure 1e) and the phase spectrum (figure 1f). The sample size is $n = 2048$. For better visibility, only the lowest 80 frequencies are used in the spectral plots. The estimated spectral modulus and the phase spectrum between $\hat{f}_1$ and $\hat{f}_2$, obtained by wavelet thresholding with $s12-$wavelets (see e.g. Daubechies 1999, Bruce and Gao 1996) and λ_j as in remark 3, are displayed in figures 1g and h. While the common frequencies are identified quite accurately, the linear structrue of the theoretical regression phase spectrum is lost almost entirely in its (unweighted) empirical counterpart. However, if we consider only values of $\hat{\phi}_{12}$ where $\hat{\kappa}^*_{12}(j)$ exceeds four times its standard deviation (horizontal line in figure 1g), the linear structure can be identified. In figure 1h, the five frequencies corresponding to the five highest values of $\hat{\kappa}^*_{12}(j)$ are marked by black squares. The estimate of the phase line based on these frequencies is very close to the true line.

To check the accuracy of our estimates, a small simulation study was carried out. For n =256, 512, 1024 and 2048, 500 series were simulated, and the spectral density and the lag between f_1 and f_2 were estimated. Figure 2 shows a comparison of the simulated standard deviation of $\hat{\kappa}^*_{12}(j)$ and the values of $\sigma_{\kappa;12}$ (equation (16)) calculated by plugging in the true and the estimated functions f_j respectively. The standard deviations are fairly close together for $n = 256$ and almost identical for $n = 2048$. This confirms the theoretical results.

n	256	512	1024	2048
true value	-.03125	-.03125	-.03125	-.03125
median	-0.0337069	-0.0317893	-0.0312525	-0.0312118
mean	-0.0429855	-0.0319160	-0.0312776	-0.0311385
std.dev.	0.1490462	0.01350248	0.00311929	0.00208201
mse	0.02768125	0.00417189	0.00391940	0.00389665
skewness	0.78195065	0.48571213	0.01500243	0.18100068

Table 1: Detailed summary of the lag estimation for various sample sizes, in each case running 500 simulations.

Detailed results for the lag-estimates are listed in table 1. Boxplots in figure 2 illustrate the fast convergence of $\hat{\Delta}$ to the true value $\Delta = -.03125$. Note that for small sample sizes, the true frequencies in the amplitude spectrum might fail to show up due to insufficient accuracy of the trend estimates. In particular, high deterministic frequencies may be hidden or smoothed out by the shrinkage estimate. Moreover, the true lead-lag structure may not be detected due to effects of the noise component. In the example here, a reliable estimate of Δ is obtained for $n = 512$, whereas $n = 256$ seems to be too small.

4.3.2 ECG data

We consider a trivariate time series consisting of electrocardiogram (ECG), blood pressure (BP) and the cardiac stroke volume (SV) measurements of a sleeping patient, recorded at the Beth Israel Deaconess Medical Center in Boston (Goldberger et al. 2000). The data are represented in units of 50mV (ECG), 50mmHG (BP) and 50ml (SV). The observational period consists of $n = 8192$ observations recorded at a rate of 200 observations per second. A detailed description of the data set can be found in Ichimaru and Moody (1999). For better visibility, only the first 1000 observations (5 seconds) of the raw time series are displayed in figure 3. The analysis is carried out for all 8192 observations. As expected, all time series mainly consist of a deterministic almost periodic movement, with a period of about 200 observations, indicating a heart rate around 60 beats per minute. The ECG signal starts with a small bump indicating the atrial contraction and is followed by a sharp peak representing the contraction of the ventricles.

The third peak represents the repolarization period. At the same time, contraction of the ventricles causes the stroke volume to increase up to a peak value, after which it falls back to its minimum shortly after the repolarization. The growing amount of blood pushed into the arteries also induces a rise in blood pressure. The electrocardiogram receives information directly from the heart whereas the blood pressure is measured by a catheder in the radial artery. Therefore, the delay represents the time it takes the pressure wave, initialized by a heart beat, to reach the catheder. Consider now the regression cross spectra in figure 4. All three amplitude spectra reveal dominating common frequencies around $j = 40$ indicating a common period of length around $1s$. Moreover, the phase spectrum of ECG and BP exhibits linear structures over certain ranges of frequencies. The phase spectrum between ECG and BP for the interval $j \in [34, 40]$ is drawn in the lower left panel of figure 4. We see an almost perfectly linear structure in the phase spectrum and the estimated phase-shift is about 26 time units indicating a lead of BP against ECG. Note, however, that for the corresponding dominating lag the regression cross-correlation is negative. This effect can be noticed visually when comparing e.g. local maxima of the BP signal with minima of the ECG signal after the repolarization period (local maxima of the blood pressure occur shortly before the ECG signal reaches its minimum). Due to the physiological fact that the peaks in ECG representing contractions of the ventricals are constantly around 94 data points away from the minimum after the repolarization period, we estimate a lag between the contraction of the ventricals and the maximum of the blood pressure of around $94 - 26 = 68$ time units which is equivalent to a time delay of $340ms$. A similar analysis between ECG and SV results in a lead of the ECG signal of about 11 data points or $55ms$.

5 Final remarks

We defined the regression cross covariance and cross spectrum for multivariate deterministic trend functions. This is a nonparametric multivariate extension of an analogous concept used by Grenander and Rosenblatt to obtain asymptotic results in the context of parametric regression. The usefulness of the nonparametric regression spectrum goes far beyond a purely mathematical device. It can be used as a data analytical tool to identify common frequencies and lead-lag effects in multivariate time series with strong deterministic components. The physiological series considered above is a typical example. Other examples can be found, for instance, in high frequency data with a strong daily seasonality and a large number of intra-day measurements. The greatest challenge appears to be estimation of the phase spectrum. The question in how far more accurate methods than the algorithm proposed here can be devised, will be worth pursuing in future research.

6 Acknowledgements

This research was supported in part by a grant of the German Research Foundation (DFG). We would like to thank the referees for their constructive remarks that helped to improve the presentation of the results.

7 Appendix: Proofs

Proof 1 *(Proposition 1)*
For $u \geq 0$,

$$\rho_{rs}(-u) = \frac{< f_r(\cdot - u), f_s >}{||f_r|| \cdot ||f_s||} = \frac{\overline{< f_s(\cdot + u), f_r >}}{||f_r|| \cdot ||f_s||} = \overline{\rho_{sr}(u)}.$$

Hence ρ_{rs} *is Hermitian. Consider now* ρ_{rs}, $u_1, \ldots, u_n \in [0,1]$, *and let* $\theta_1, \ldots, \theta_n$ *be arbitrary coefficients in* $\mathbb{C}$. *Then*

$$\sum_{i,j=1}^{n} \sum_{r,s=1}^{p} \theta_i \rho_{rs}(u_i - u_j) \overline{\theta_j}$$
$$= \sum_{i,j} \sum_{r,s} \theta_i \frac{\int f_r(x + u_i - u_j) \overline{f_s(x)} dx}{\sqrt{\int |f_r(x)|^2 dx \int |f_s(x)|^2 dx}} \overline{\theta_j}.$$

Since the denominator does not depend on i *and* j, *non-negative definiteness follows from*

$$\int \sum_r \sum_i \theta_i f_r(x + u_i) \sum_s \sum_j \overline{\theta_j f_s(x + u_j)} dx$$
$$= \int \left| \sum_r \sum_i \theta_i f_r(x + u_i) \right|^2 dx \geq 0.$$

Proof 2 *(Lemma 1) The proof essentially follows directly from Theorem 3 in Brillinger (1994). To see uniform convergence of $r(n)$ consider*

$$cov(\hat{f}_r(x), \hat{f}_s(y)) = \sum_{k,k'} cov(\hat{\alpha}_{l,k}, \hat{\alpha}_{l,k'})\phi_{l,k}(x)\phi_{l,k'}(y) \tag{18}$$

$$+ \sum_{j,k,j',k'} cov\,(\hat{w}^{(r)}_{j,k}\hat{\beta}^{(r)}_{j,k}, \hat{w}^{(s)}_{j',k'}\hat{\beta}^{(s)}_{j',k'})\psi_{j,k}(x)\psi_{j',k'}(y) \tag{19}$$

$$+ \sum_{k,j',k'} cov\,(\hat{\alpha}_{l,k}, \hat{w}^{(s)}_{j',k'}\hat{\beta}^{(s)}_{j',k'})\phi_{l,k}(x)\psi_{j',k'}(y) \tag{20}$$

$$+ \sum_{j,k,k'} cov\,(\hat{w}^{(r)}_{j,k}\hat{\beta}^{(r)}_{j,k}, \hat{\alpha}_{l,k'})\psi_{j,k}(x)\phi_{l,k'}(y). \tag{21}$$

The remainder $r(n)$ essentially consists of terms with $(j,k) \neq (j',k')$. Separate e.g. (19) into

$$\sum_{j,j' \leq J_0} \sum_{k,k'} cov\,(\hat{w}^{(r)}_{j,k}\hat{\beta}^{(r)}_{j,k}, \hat{w}^{(s)}_{j',k'}\hat{\beta}^{(s)}_{j',k'})\psi_{j,k}(x)\psi_{j',k'}(y)$$
$$+ \sum_{j>J_0 \vee j'>J_0} \sum_{k,k'} cov\,(\hat{w}^{(r)}_{j,k}\hat{\beta}^{(r)}_{j,k}, \hat{w}^{(s)}_{j',k'}\hat{\beta}^{(s)}_{j',k'})\psi_{j,k}(x)\psi_{j',k'}(y). \tag{22}$$

For $j, j' \leq J_0$ and suitable constants A_1 and A_2, we have

$$|\sum_{j\neq j'}^{J_0} \sum_{k \neq k'} \text{cov}\ (\hat{w}^{(r)}_{j,k}\hat{\beta}^{(r)}_{j,k}, \hat{w}^{(s)}_{j',k'}\hat{\beta}^{(s)}_{j',k'})\psi_{j,k}(x)\psi_{j',k'}(y)|$$
$$\leq \sum_{j\neq j'}^{J_0} \sum_{k \neq k'} |\text{cov}\ \ (\hat{w}^{(r)}_{j,k}\hat{\beta}^{(r)}_{j,k}, \hat{w}^{(s)}_{j',k'}\hat{\beta}^{(s)}_{j',k'})| \cdot 2^{\frac{j}{2}} 2^{\frac{j'}{2}} \sup_x\{\psi(x)^2\}$$
$$\leq A_1 \sum_{j\neq j'}^{J_0} \sum_{k \neq k'} |\text{cov}\ (\hat{\beta}^{(r)}_{j,k}, \hat{\beta}^{(s)}_{j',k'})| 2^{\frac{j}{2}} 2^{\frac{j'}{2}}$$
$$\leq A_2 \sum_{j\neq j'}^{J_0} \sum_{k \neq k'} n^{-2} 2^j 2^{j'} = O(n^{-2}).$$

Now for $j > J_0$ or $j' > J_0$, we have, recalling that $|cov\,(\hat{\beta}^{(r)}_{j,k}, \hat{\beta}^{(s)}_{j',k'})| =$

$O(2^{j/2}2^{j'/2}n^{-2})$ for $(j,k) \neq (j',k')$,

$$\begin{aligned}
&\left|\sum_{j>J_0 \vee j'>J_0}\sum_{k,k'} cov(\hat{w}^{(r)}_{j,k}\hat{\beta}^{(r)}_{j,k}, \hat{w}^{(s)}_{j',k'}\hat{\beta}^{(s)}_{j',k'})\psi_{j,k}(x)\psi_{j',k'}(y)\right| \\
&\leq \sum_{j>J_0 \vee j'>J_0}\sum_{k,k'} \left|cov(\hat{w}^{(r)}_{j,k}\hat{\beta}^{(r)}_{j,k}, \hat{w}^{(s)}_{j',k'}\hat{\beta}^{(s)}_{j',k'})\right| |\psi_{j,k}(x)\psi_{j',k'}(y)| \\
&\leq \sum_{j>J_0 \vee j'>J_0}\sum_{k,k'} \left|cov(\hat{\beta}^{(r)}_{j,k}, \hat{\beta}^{(s)}_{j',k'})\right| \cdot A_1 2^{j/2}2^{j'/2} \\
&\leq A_2 n^{-2} \sum_{j>J_0 \vee j'>J_0} 2^j 2^{j'} \sup_x\{\psi(x)^2\} \\
&= O(2^{2J_n}n^{-2}), \qquad (23)
\end{aligned}$$

uniformly in x and y where A_1, A_2 are suitable constants.

Proof 3 *(Lemma 2) As mentioned in remark 6, Brillinger shows that for each r, $1 \leq r \leq p$, the variance of the thresholding estimators, $var\,(\hat{f}_r(x))$, is of order n^{-1}. To indicate its dependence on Ω, we write $\widetilde{\epsilon}^{(r)}_n(\omega, x)$ instead of $\widetilde{\epsilon}^{(r)}_n(x)$. Define $\mathcal{X} := [0,1]$, let μ be the Lebesgue measure on $(\mathcal{X}, \mathcal{F}_\mathcal{X})$ and denote by $\mathcal{F}_\Omega$ the σ-algebra generated by the open subsets of Ω. We consider the product space $(\Omega \times \mathcal{X})$ with σ-algebra $\mathcal{F} = \mathcal{F}_\Omega \otimes \mathcal{F}_\mathcal{X}$ and corresponding measure $P \otimes \mu$. Due to the uniform convergence of $var\,[\hat{f}_r(x)]$,*

$$\int_\mathcal{X}\int_\Omega (\widetilde{\epsilon}^{(r)}_n(\omega,x) - E[\widetilde{\epsilon}^{(r)}_n(\omega,x)])^2 dPdx \to 0$$

and

$$n\int_\mathcal{X}\int_\Omega (\widetilde{\epsilon}^{(r)}_n(\omega,x) - E[\widetilde{\epsilon}^{(r)}_n(\omega,x)])^2 dPdx = O(1).$$

This implies

$$\int_\mathcal{X}\int_\Omega \widetilde{\epsilon}^{(r)}_n(\omega,x)^2 dPdx = O(n^{-1}). \qquad (24)$$

Note that $\mathcal{X}$ is a finite interval so that $(P \otimes \mu)(\Omega \times \mathcal{X}) < \infty$. Square integrability in finite measure spaces implies that

$$\int_\mathcal{X}\int_\Omega |\widetilde{\epsilon}^{(r)}_n(\omega,x)|\ dP\,dx = O(n^{-1/2}).$$

Due to the existence of both integrals we can apply Fubini's theorem so that

$$\int_\Omega \int_{\mathcal{X}} |\widetilde{\epsilon}_n^{(r)}(\omega, x)| \; dx \, dP \to 0, \;\; n \to \infty.$$

Define the random variable $z_n = \int_{\mathcal{X}} |\widetilde{\epsilon}_n^{(r)}(x)| \; dx$. *Then* $z_n \geq 0$ *for all* n *and* $E[z_n] \to 0$. *Hence,*

$$\int_{[0,1]} \widetilde{\epsilon}_n^{(r)}(\omega, x) \; dx \to 0$$

almost surely.

Proof 4 *(theorem 1) Without loss of generality it suffices to show that the theorem is true for real-valued functions* f_r *and for non-negative values of* u. *By adding and substracting terms, we split the regression covariance into four terms*

$$\begin{aligned}
(I) &= \int_0^1 \left(\hat{f}_r(x+u) - f_r(x+u)\right) \left(\hat{f}_s(x) - f_s(x)\right) dx \\
(II) &= \int_0^1 f_r(x+u) \left(\hat{f}_s(x) - f_s(x)\right) dx \\
(III) &= \int_0^1 \left(\hat{f}_r(x+u) - f_r(x+u)\right) f_s(x) dx \\
(IV) &= \int_0^1 f_r(x+u) f_s(x) dx.
\end{aligned}$$

The functions f_r *and* f_s *are bounded and of bounded variation. Then*

$$|(III)| \leq A \int \widetilde{\epsilon}_n^{(r)}(x+u) dx,$$

which tends to 0 almost surely according to lemma 2. The same holds for term (II). *Moreover, the Cauchy-Schwarz inequality implies*

$$\begin{aligned}
&\left| \int \left(\hat{f}_r(x+u) - f_r(x+u)\right) \left(\hat{f}_s(x) - f_s(x)\right) dx \right| \qquad (25) \\
&\leq \left(\int (\hat{f}_r(x+u) - f_r(x+u))^2 dx \right)^{\frac{1}{2}} \left(\int (\hat{f}_s(x) - f_s(x))^2 dx \right)^{\frac{1}{2}}.
\end{aligned}$$

The proof of lemma 2 shows that both factors converge to zero almost surely. Therefore,

$$\hat{\gamma}_{rs}(u) \xrightarrow{P} \int_0^1 f_r(x+u) f_s(x) dx = \gamma_{rs}(u).$$

Proof 5 *(theorem 2) Recall, that*

$$\hat{\gamma}_{rs}(u) - \gamma_{rs}(u) = \int [f_r(x+u)\widetilde{\epsilon}_n^{(s)}(x) + \widetilde{\epsilon}_n^{(r)}(x+u) f_s(x) + \widetilde{\epsilon}_n^{(r)}(x+u)\widetilde{\epsilon}_n^{(s)}(x)] dx.$$

Hence,

$$\begin{aligned}
&\lim_{n\to\infty} n cov\left(\hat{\gamma}_{rs}(u_i), \hat{\gamma}_{rs}(u_j)\right) \\
&= \lim_{n\to\infty} nE[\int \left(f_r(x+u_i)\widetilde{\epsilon}_n^{(s)}(x) + f_s(x)\widetilde{\epsilon}_n^{(r)}(x+u_i) + \widetilde{\epsilon}_n^{(r)}(x+u_i)\widetilde{\epsilon}_n^{(s)}(x)\right) dx \\
&\qquad \cdot \int \left(f_r(y+u_j)\widetilde{\epsilon}_n^{(s)}(y) + f_s(y)\widetilde{\epsilon}_n^{(r)}(y+u_j) + \widetilde{\epsilon}_n^{(r)}(y+u_j)\widetilde{\epsilon}_n^{(s)}(y)\right) dy] \\
&=: (I) + (II) + (III),
\end{aligned}$$

where

$$(I) = \begin{cases} \lim_{n\to\infty} & n\Big\{E[\int\int f_r(x+u_i) f_r(y+u_j)\widetilde{\epsilon}_n^{(s)}(x)\widetilde{\epsilon}_n^{(s)}(y) dx\, dy] \\ & +E[\int\int f_s(x) f_r(y+u_j)\widetilde{\epsilon}_n^{(r)}(x+u_i)\widetilde{\epsilon}_n^{(s)}(y) dx\, dy] \\ & +E[\int\int f_r(x+u_i) f_s(y)\widetilde{\epsilon}_n^{(s)}(x)\widetilde{\epsilon}_n^{(r)}(y+u_j) dx\, dy] \\ & +E[\int\int f_s(x) f_s(y)\widetilde{\epsilon}_n^{(r)}(x+u_i)\widetilde{\epsilon}_n^{(r)}(y+u_j) dx\, dy]\Big\}, \end{cases}$$

$$(II) = \begin{cases} \lim_{n\to\infty} & n\Big\{E[\int\int f_r(x+u_i)\widetilde{\epsilon}_n^{(s)}(x)\widetilde{\epsilon}_n^{(r)}(y+u_j)\widetilde{\epsilon}_n^{(s)}(y) dx\, dy] \\ & +E[\int\int f_s(x)\widetilde{\epsilon}_n^{(r)}(x+u_i)\widetilde{\epsilon}_n^{(r)}(y+u_j)\widetilde{\epsilon}_n^{(s)}(y) dx\, dy] \\ & +E[\int\int f_r(y+u_j)\widetilde{\epsilon}_n^{(s)}(y)\widetilde{\epsilon}_n^{(r)}(x+u_i)\widetilde{\epsilon}_n^{(s)}(x) dx\, dy] \\ & +E[\int\int f_s(y)\widetilde{\epsilon}_n^{(r)}(y+u_j)\widetilde{\epsilon}_n^{(r)}(x+u_i)\widetilde{\epsilon}_n^{(s)}(x)\, dxdy]\Big\}, \end{cases}$$

$$(III) = \lim_{n\to\infty} n\left\{E[\int\int \widetilde{\epsilon}_n^{(r)}(x+u_i)\widetilde{\epsilon}_n^{(s)}(x)\widetilde{\epsilon}_n^{(r)}(y+u_j)\widetilde{\epsilon}_n^{(s)}(y) dx\, dy]\right\}.$$

Consider just one part of the sum in part (I). Due to Fubini's theorem

$$\lim_{n\to\infty} nE[\int\int f_r(x+u_i) f_r(y+u_j)\widetilde{\epsilon}_n^{(s)}(x)\widetilde{\epsilon}_n^{(s)}(y)]\, dx\, dy$$

$$= \lim_{n\to\infty} n \int\int f_r(x+u_i) f_r(y+u_j) E[\widetilde{\epsilon}_n^{(s)}(x)\widetilde{\epsilon}_n^{(s)}(y)]\, dx\, dy$$

$$= \int\int f_r(x+u_i) f_r(y+u_j) \lim_{n\to\infty} nE[\widetilde{\epsilon}_n^{(s)}(x)\widetilde{\epsilon}_n^{(s)}(y)]\, dx\, dy, \tag{26}$$

where the second equation is due to Lebesgue's theorem. By applying the results of lemma 1,

$$\begin{aligned}
(26) &= \int\int f_r(x+u_i) f_r(y+u_j) \lim_{n\to\infty} n\,cov\left(\hat{f}_s(x), \hat{f}_s(y)\right)\, dx\, dy \\
&= \int\int f_r(x+u_i) f_r(y+u_j) \Bigg[2\pi h_{\epsilon^{(ss)}}(0) \sum_k \phi_{l,k}(x)\phi_{l,k}(y) \\
&+ 2\pi h_{\epsilon^{(ss)}}(0) \sum_{j\geq l,k} w_{j,k}^{(s)} \psi_{j,k}(x)\psi_{j,k}(y) + O(n^{-1})\Bigg]\, dx\, dy.
\end{aligned}$$

Rearrangig the sums and integrals results in

$$\begin{aligned}
&2\pi h_{\epsilon^{(ss)}}(0) \sum_k \int f_r(x+u_i)\phi_{l,k}(x)dx \int f_r(y+u_j)\phi_{l,k}(y)dy \\
&+ 2\pi h_{\epsilon^{(ss)}}(0) \sum_{j\geq l,k} w_{j,k}^{(s)} \int f_r(x+u_i)\psi_{j,k}(x)dx \int f_r(y+u_j)\psi_{j,k}(y)dy \\
&+ O(n^{-1}).
\end{aligned}$$

Analogous results for the other parts of (I) yield $\sigma_{rs}(u_i, u_j)$. It remains to show that the remaining parts converge to 0. By lemma 2, $\int \widetilde{\epsilon}_n^{(r)}(x)dx \to 0$ almost surely with rate $n^{-1/2}$. Then (24) together with the Cauchy-Schwarz inequality yield $\int \widetilde{\epsilon}_n^{(r)}(y+u_j)\widetilde{\epsilon}_n^{(s)}(y)dy = O_p(n^{-1})$. Hence,

$$\begin{aligned}
&\lim_{n\to\infty} n \left| E[\int\int f_r(x+u_i)\widetilde{\epsilon}_n^{(s)}(y)\widetilde{\epsilon}_n^{(r)}(y+u_j)\widetilde{\epsilon}_n^{(s)}(x)dx\, dy] \right| \\
&\leq \lim_{n\to\infty} n\cdot AE[\underbrace{\int |\widetilde{\epsilon}_n^{(s)}(x)|dx}_{=O_p(n^{-1/2})} \underbrace{\left|\int \widetilde{\epsilon}_n^{(r)}(y+u_j)\widetilde{\epsilon}_n^{(s)}(y)dy\right|}_{=O_p(n^{-1})}] \\
&= 0.
\end{aligned}$$

Analogous considerations imply that (III) converges to 0.

Proof 6 *(Lemma 3) Consider the shrinkage estimator*

$$\hat{f}_s(x) = \sum_k \hat{\alpha}_{l,k}^{(s)} \phi_{l,k}(x) + \sum_{j \geq l}^{J_n} \sum_k \hat{w}_{j,k}^{(s)} \hat{\beta}_{j,k}^{(s)} \psi_{j,k}(x). \tag{27}$$

As above, let J_0 be the largest number such that the corresponding resolution component of f_s contains at least one non-zero coefficient.
Hence,

$$\begin{aligned}
(\hat{f}_s - f_s)(x) &= \underbrace{\sum_k \hat{\alpha}_{l,k}^{(s)} \phi_{l,k}(x) - \sum_k \alpha_{l,k}^{(s)} \phi_{l,k}(x)}_{=:S(x)} \\
&\quad + \underbrace{\sum_{j \geq l}^{J_n} \sum_k \hat{w}_{j,k}^{(s)} \hat{\beta}_{j,k}^{(s)} \psi_{j,k}(x) - \sum_{j \geq l}^{J_0} \sum_k \beta_{j,k}^{(s)} \psi_{j,k}(x)}_{=:D(x)} \\
&= S(x) + D(x).
\end{aligned}$$

Inserting the empirical Fourier coefficients yields

$$\begin{aligned}
S(x) &= \sum_k [\frac{1}{n} \sum_{u=1}^n Y_s(u) \phi_{l,k}(u/n) - \alpha_{l,k}^{(s)}] \phi_{l,k}(x) \\
&= \sum_k [\frac{1}{n} \sum_{u=1}^n (Y_s(u) - E[Y_s(u)] + E[Y_s(u)]) \phi_{l,k}(u/n) - \alpha_{l,k}^{(s)}] \phi_{l,k}(x) \\
&= \sum_k [\frac{1}{n} \sum_{u=1}^n (\epsilon_s(u) \phi_{l,k}(u/n) + f_s(u/n) \phi_{l,k}(u/n)) - \alpha_{l,k}^{(s)}] \phi_{l,k}(x) \\
&= \frac{1}{n} \sum_{u=1}^n \epsilon_s(u) \sum_k \phi_{l,k}(u/n) \phi_{l,k}(x) + \sum_k [\frac{1}{n} \sum_{u=1}^n f_s(u/n) \phi_{l,k}(u/n) - \alpha_{l,k}^{(s)}] \phi_{l,k}(x).
\end{aligned}$$

Now

$$\begin{aligned}
&\left| \frac{1}{n} \sum_{u=1}^n f_s(u/n) \phi_{l,k}(u/n) - \int_0^1 f_s(x) \phi_{l,k}(x) dx \right| \\
&= \left| \frac{1}{n} \sum_{u=1}^n f_s(u/n) \phi_{l,k}(u/n) - \alpha_{l,k}^{(s)} \right| \leq \frac{V(f_s \phi_{l,k})}{n},
\end{aligned} \tag{28}$$

where $V(\cdot)$ denotes total variation (see e.g. Polya and Szegö 1964). Hence,

$$S(x) = \frac{1}{n}\sum_{u=1}^{n}\epsilon_s(u)\sum_k \phi_{l,k}(u/n)\phi_{l,k}(x) + O(n^{-1}). \tag{29}$$

For n large enough such that $J_n > J_0$:

$$\begin{aligned} D(x) &= \sum_j^{J_n}\sum_k \hat{w}_{j,k}^{(s)}\hat{\beta}_{j,k}^{(s)}\psi_{j,k}(x) - \sum_j^{J_0}\sum_k \beta_{j,k}^{(s)}\psi_{j,k}(x) \\ &= \sum_j^{J_0}\sum_k(\hat{w}_{j,k}^{(s)}\hat{\beta}_{j,k}^{(s)} - \beta_{j,k}^{(s)})\psi_{j,k}(x) + \sum_{j=J_0+1}^{J_n}\sum_k \hat{w}_{j,k}^{(s)}\hat{\beta}_{j,k}^{(s)}\psi_{j,k}(x). \end{aligned} \tag{30}$$

Now,

$$\begin{aligned} &\sum_j^{J_0}\sum_k(\hat{w}_{j,k}^{(s)}\hat{\beta}_{j,k}^{(s)} - \beta_{j,k}^{(s)})\psi_{j,k}(x) \\ &= \sum_{j,k}(\hat{w}_{j,k}^{(s)}\frac{1}{n}\sum_{u=1}^{n}\psi_{j,k}(u/n)Y_s(u) - \beta_{j,k}^{(s)})\psi_{j,k}(x) \\ &= \sum_{j,k}\{\hat{w}_{j,k}^{(s)}[\frac{1}{n}\sum_{u=1}^{n}\psi_{j,k}(u/n)(Y_s(u) - E[Y_s(u)] + E[Y_s(u)])] - \beta_{j,k}^{(s)}\}\psi_{j,k}(x) \\ &= \sum_{j,k}\{\hat{w}_{j,k}^{(s)}\frac{1}{n}\sum_{u=1}^{n}\psi_{j,k}(u/n)\epsilon_s(u) + \hat{w}_{j,k}^{(s)}\frac{1}{n}\sum_{u=1}^{n}\psi_{j,k}(u/n)f_s(u/n) - \beta_{j,k}^{(s)}\}\psi_{j,k}(x) \\ &= \sum_{j,k}\hat{w}_{j,k}^{(s)}\frac{1}{n}\sum_{u=1}^{n}\psi_{j,k}(u/n)\epsilon_s(u)\psi_{j,k}(x) \qquad (31) \\ &+ \sum_{j,k}[\hat{w}_{j,k}^{(s)}\frac{1}{n}\sum_{u=1}^{n}\psi_{j,k}(u/n)f_s(u/n) - \beta_{j,k}^{(s)}]\psi_{j,k}(x). \qquad (32) \end{aligned}$$

For the term in (32), we have

$$|\frac{1}{n}\sum_{u=1}^{n}\psi_{j,k}(u/n)f_s(u/n) - \beta_{j,k}^{(s)}| \leq \frac{V(\psi_{j,k}f_s)}{n},$$

where the total variation is $A\cdot 2^{\frac{j}{2}}$, with A a suitable constant. Consider now the binary random variable $\hat{w}_{j,k}^{(s)}$. We distinguish between the cases $\beta_{j,k}^{(s)} = 0$

and $\beta_{j,k}^{(s)} \neq 0$. *In the first case,* $n^{-1}\sum_{u=1}^{n}\psi_{j,k}(u/n)f_s(u/n) = O(2^{\frac{j}{2}}n^{-1})$. *Recall that* $j \leq J_0$. *This implies that (32) converges in probability to 0 and is of order* $O_p(n^{-1})$. *If* $\beta_{j,k}^{(s)} \neq 0$, *consider first the case* $\hat{w}_{j,k}^{(s)} = 1$. *Then, the rate of convergence in equation (32) is* n^{-1}. *Moreover,*

$$P(\hat{w}_{j,k}^{(s)} = 1) = P(|\hat{\beta}_{j,k}^{(s)}| > \lambda_j \cdot \sqrt{var\ (\hat{\beta}_{j,k}^{(s)})}) \to 1.$$

This implies that there exists a sequence of subspaces $\Omega_n \uparrow \Omega$ *such that* $P(\Omega_n) \to 1$ *and*

$$P(\hat{w}_{j,k}^{(s)} = 1 | \omega \in \Omega_n) = 1.$$

Define $\Omega_n^c = \Omega - \Omega_n$. *Then, for a suitable constant* A,

$$\begin{aligned}
&P\left(|\hat{w}_{j,k}^{(s)}\frac{1}{n}\sum_{u=1}^{n}\psi_{j,k}(u/n)f_s(u/n) - \beta_{j,k}^{(s)}| \leq \frac{V(\psi_{j,k}f_s)}{n}\right) \\
&= P\left(|\hat{w}_{j,k}^{(s)}\frac{1}{n}\sum_{u=1}^{n}\psi_{j,k}(u/n)f_s(u/n) - \beta_{j,k}^{(s)}| \leq \frac{V(\psi_{j,k}f_s)}{n} \;\middle|\; \Omega_n\right) P(\Omega_n) \\
&+ P\left(|\hat{w}_{j,k}^{(s)}\frac{1}{n}\sum_{u=1}^{n}\psi_{j,k}(u/n)f_s(u/n) - \beta_{j,k}^{(s)}| \leq \frac{V(\psi_{j,k}f_s)}{n} \;\middle|\; \Omega_n^c\right) P(\Omega_n^c) \\
&\to 1.
\end{aligned}$$

Term (32) contains a finite amount of coefficients $\beta_{j,k}^{(s)} \neq 0$ *so that convergence is uniform* ($= O_p(n^{-1})$). *The second part of* (30) *is given by*

$$\sum_{j=J_0+1}^{J_n}\sum_{k}\hat{w}_{j,k}^{(s)}\hat{\beta}_{j,k}^{(s)}\psi_{j,k}(x). \tag{33}$$

For all $j,k \in \mathbb{Z}$, $\hat{\beta}_{j,k}^{(s)}$ *is a root-n consistent estimator for* $\beta_{j,k}^{(s)}$ $(s = 1,\ldots,p)$ *and for all* $j > J_0$, $\beta_{j,k}^{(s)} = 0$. *With the preliminary remarks of section 3.2 about* $var(\hat{\beta}_{j,k}^{(s)})$, *consider the estimator* $\hat{w}_{j,k}^{(s)}$ *in the case* $w_{j,k}^{(s)} = 0$.

$$\begin{aligned}
P(\hat{w}_{j,k}^{(s)} = 1) &= P(|\hat{\beta}_{j,k}^{(s)}| > \lambda_j \cdot \sqrt{var\ (\hat{\beta}_{j,k}^{(s)})}) \\
&= P(\sqrt{n}|\hat{\beta}_{j,k}^{(s)}| > \lambda_j\sqrt{n} \cdot \sqrt{var\ (\hat{\beta}_{j,k}^{(s)})}) \\
&\leq P(\sqrt{n}|\hat{\beta}_{j,k}^{(s)}| > A\lambda_j) \\
&\leq P(\sqrt{n}|\hat{\beta}_{j,k}^{(s)}| > \eta) \to 0
\end{aligned}$$

for a suitable constant A and $\eta > 0$. Convergence is uniform in j so that (33) can be neglected asymptotically. Rearranging (31) and combining this with (29) yields the desired result.

Proof 7 *(Lemma 4) It is sufficient to show the assertion for positive values of u.*

$$\sqrt{n}(\hat{\gamma}_{rs}(u) - \gamma_{rs}(u)) = \sqrt{n}\int_0^1 [f_r(x+u)\widetilde{\epsilon}_n^{(s)}(x) + f_s(x)\widetilde{\epsilon}_n^{(r)}(x+u) + \widetilde{\epsilon}_n^{(r)}(x+u)\widetilde{\epsilon}_n^{(s)}(x)]dx.$$

Applying lemma 3, the first part leads to

$$\begin{aligned}
&\sqrt{n}\int f_r(x+u)\left[\frac{1}{n}\sum_{u=1}^{n}\epsilon_s(u)\sum_k \phi_{l,k}(u/n)\phi_{l,k}(x)\right.\\
&\left.+\frac{1}{n}\sum_{u=1}^{n}\epsilon_s(u)\sum_{j,k}\hat{w}_{j,k}^{(s)}\psi_{j,k}(u/n)\psi_{j,k}(x) + O_p(n^{-1})\right]dx\\
&= \frac{1}{\sqrt{n}}\sum_{u=1}^{n}\epsilon_s(u)\sum_k \phi_{l,k}(u/n)\int f_r(x+u)\phi_{l,k}(x)dx\\
&+\frac{1}{\sqrt{n}}\sum_{u=1}^{n}\epsilon_s(u)\sum_{j,k}\hat{w}_{j,k}^{(s)}\psi_{j,k}(u/n)\int f_r(x+u)\psi_{j,k}(x)dx + O_p(n^{-1/2})\\
&=: \sum_{u=1}^{n}\epsilon_s(u)w_{u,n}^{(s)} + O_p(n^{-1/2}),
\end{aligned} \tag{34}$$

where $w_{u,n}^{(s)}$ is a triangular array of partly deterministic weights given by

$$w_{u,n}^{(s)} := \frac{1}{\sqrt{n}}\left(\sum_k \phi_{l,k}(u/n)\int f_r(x+u)\phi_{l,k}(x)dx + \sum_{j,k}\hat{w}_{j,k}^{(s)}\psi_{j,k}(u/n)\int f_r(x+u)\psi_{j,k}(x)dx\right).$$

Therefore,

$$\sqrt{n}\int f_s(x)\widetilde{\epsilon}_n^{(r)}(x+u) = \sum_{u=1}^{n}\epsilon_r(u)w_{u,n}^{(r)} + O_p(n^{-1/2}),$$

where

$$w^{(r)}_{u,n} \approx \frac{1}{\sqrt{n}} \sum_k \phi_{l,k}(u/n) \int f_s(x)\phi_{l,k}(x+u)dx$$
$$+ \frac{1}{\sqrt{n}} \sum_{j,k} w^{(r)}_{j,k}\psi_{j,k}(u/n) \int f_s(x)\psi_{j,k}(x+u)dx.$$

As a result of the proof of theorem 1

$$n \int \widetilde{\epsilon}^{(r)}_n(x+u)\widetilde{\epsilon}^{(s)}_n(x)dx = O_p(1).$$

It follows that

$$\begin{aligned}\sqrt{n}(\hat{\gamma}_{rs}(u) - \gamma_{rs}(u)) &= \sum_{u=1}^{n}(w^{(r)}_{u,n}\epsilon_r(u) + w^{(s)}_{u,n}\epsilon_s(u)) + O_p(n^{-1/2}),\\ &= \sum_{u=1}^{2n} \epsilon^*_u w^*_{u,n} + O_p(n^{-1/2}),\end{aligned}$$

where $\epsilon^*_u = \epsilon_r(u),\ u = 1,\ldots,n,$ *and* $\epsilon^*_u = \epsilon_s(u-n),\ u = n+1,\ldots,2n,$ *with the weights* $w^*_{u,n}$ *defined accordingly. Denote by* $\hat{B}_n(u)$ *the random part of* $w^{(s)}_{u,n}$*. Then,*

$$w^{(s)}_{u,n} := A_n(u) + \hat{B}_n(u) = A_n(u) + B_n(u) + (\hat{B}_n(u) - B_n(u)),$$

with

$$A_n(u) = \frac{1}{\sqrt{n}} \sum_k \phi_{l,k}(u/n) \int f_r(x+u)\phi_{l,k}(x)dx,$$
$$\hat{B}_n(u) = \frac{1}{\sqrt{n}} \sum_{j,k} \hat{w}^{(s)}_{j,k}\psi_{j,k}(u/n) \int f_r(x+u)\psi_{j,k}(x)dx,$$
$$B_n(u) = \frac{1}{\sqrt{n}} \sum_{j,k} w^{(s)}_{j,k}\psi_{j,k}(u/n) \int f_r(x+u)\psi_{j,k}(x)dx,$$
$$\sum_{u=1}^{n} \epsilon_s(u)w^{(s)}_{u,n} = \sum_{u=1}^{n} \epsilon^{(s)}_u(A_n(u) + B_n(u)) + \sum_{u=1}^{n} \epsilon_s(u)(\hat{B}_n(u) - B_n(u)) \quad (35)$$

and

$$\sum_{u=1}^{n} \epsilon_s(u)(\hat{B}_n(u) - B_n(u)) = \frac{1}{\sqrt{n}} \sum_{u=1}^{n} [\sum_{j,k} (\hat{w}_{j,k}^{(s)} - w_{j,k}^{(s)}) \psi_{j,k} \left(\frac{u}{n}\right) b_{j,k}^{(r)}] \epsilon_s(u), \quad (36)$$

where $b_{j,k}^{(r)} := \int f_r(x+u)\psi_{j,k}(x)dx$. *Since*

$$\hat{w}_{j,k}^{(s)} - w_{j,k}^{(s)} = \mathbf{1}_{\{|\hat{\beta}_{j,k}^{(s)}| \geq \sqrt{var\,(\hat{\beta}_{j,k}^{(s)})}\lambda_j\}} - \mathbf{1}_{\{|\beta_{j,k}^{(s)}|>0\}},$$

(36) only contains non-zero elements for those indices (j,k) *where either* $|\hat{\beta}_{j,k}^{(s)}| \geq \sqrt{var\,(\hat{\beta}_{j,k}^{(s)})}\lambda_j$ *and* $\beta_{j,k}^{(s)} = 0$ *or* $|\hat{\beta}_{j,k}^{(s)}| < \sqrt{var\,(\hat{\beta}_{j,k}^{(s)})}\lambda_j$ *and* $\beta_{j,k}^{(s)} \neq 0$. *Consider first* (j,k) *with* $\beta_{j,k}^{(s)} \neq 0$. *Since* $P(\hat{w}_{j,k}^{(s)} = 1) \to 1$, $\hat{w}_{j,k}^{(s)} - w_{j,k}^{(s)} \to 0$ *in probability. Convergence is uniform in the set of* (j,k) *with* $\beta_{j,k}^{(s)} \neq 0$, *since this set is finite. For* $\beta_{j,k}^{(s)} = 0$,

$$(36) = \sum_{u=1}^{n} \hat{B}_n(u)\epsilon_s(u).$$

and $P(\hat{w}_{j,k}^{(s)} = 1) \to 0$ *uniformly in* j,k. *Consider* $\hat{B}_n(u)$ *for the case* $u = k2^{-j}$ *with* (j,k) *arbitrary. In this case there is at most a finite number of* $b_{j,k}^{(r)} \neq 0$. *For each* $\eta > 0$ *and a suitable constant* A,

$$\begin{aligned} &P(|\sqrt{n}\hat{B}_n(u)| > \eta) \\ &= P\left(\left|\sum_{j}^{J_n}\sum_{k} \hat{w}_{j,k}^{(s)}\psi_{j,k}\left(\frac{u}{n}\right) b_{j,k}^{(r)}\right| > \eta\right) \\ &\leq P\left(A\sum_{j}^{J_0}\sum_{k} |\hat{w}_{j,k}^{(s)}| > \eta\right) \to 0. \end{aligned} \quad (37)$$

Consider now $u \neq k2^{-j}$. *For sufficiently large* n *there exists an integer* k *such that* $|u - k\cdot 2^{-J_n}| < 2^{-J_n}$. *For suitable constants* A_1, A_2

$$\begin{aligned} &\left|\int f_r(x+u)\psi_{j,k}(x)dx - \int f_r(x+k2^{-J_n})\psi_{j,k}(x)dx\right| \\ &\leq A_1\cdot 2^{j/2}\int |f_r(x+u) - f_r(x+k2^{-J_n})|dx \\ &\leq A_1\cdot 2^{j/2}\sup_x |f_r(x+u) - f_r(x+k2^{-J_n})|\cdot|u-k2^{-J_n}|. \\ &\leq A_2\cdot 2^{-J_n} \to 0. \end{aligned}$$

due to assumption (A2). Therefore,

$$P\left(|\sum_{j,k}\hat{w}^{(s)}_{j,k}\psi_{j,k}\left(\frac{u}{n}\right)b^{(r)}_{j,k}| > \eta\right)$$

$$\leq P\left(\sum_{j,k}\left|\hat{w}^{(s)}_{j,k}\psi_{j,k}\left(\frac{u}{n}\right)\right|[\int|f_r(x+k2^{-j})\psi_{j,k}(x)dx| + A_2 2^{-J_n}] > \eta\right)$$

$$\leq P\left(\sum_{j,k}\left|\hat{w}^{(s)}_{j,k}\psi_{j,k}\left(\frac{u}{n}\right)\right|\int|f_r(x+k2^{-j})\psi_{j,k}(x)dx| > \frac{\eta}{2}\right) \tag{38}$$

$$+ P\left(A_2 2^{-J_n}\sum_{j,k}\left|\hat{w}^{(s)}_{j,k}\psi_{j,k}\left(\frac{u}{n}\right)\right| > \frac{\eta}{2}\right). \tag{39}$$

(38) converges to 0 in probability according to (37). Moreover,

$$P\left(\left|A_2 2^{-J_n}\sum_{j,k}^{J_n}|\hat{w}^{(s)}_{j,k}\psi_{j,k}\left(\frac{u}{n}\right)\right| > \frac{\eta}{2}\right) \leq P\left(\left|A_2 2^{-J_n}\sum_{j}^{J_n}2^{j/2}\right| > \frac{\eta}{2}\right) \to 0,$$

so that $P(\sqrt{n}\hat{B}_n(u)| > \eta) \to 0$. *Since* $\epsilon^{(s)}_u$ *form a square-integrable martingale difference array with constant variance,* $n^{-\frac{1}{2}}\sum_{u=1}^n \epsilon_s(u)$ *converges in distribution to a normal variable and* $\sum_{u=1}^n \hat{B}_n(u)\epsilon_s(u) \to 0$ *in probability. Combining all cases,* $\sum_u \epsilon_s(u)(\hat{B}_n(u) - B_n(u)) \to 0$ *in probability so that asymptotically the second part of the sum in (35) can be neglected. Hence,* $w^{(s)}_{u,n}$ *reduce to the deterministic weights*

$$\begin{aligned}\tilde{w}^{(s)}_{u,n} &= A_n(u) + B_n(u)\\ &= \frac{1}{\sqrt{n}}\sum_k \phi_{l,k}(u/n)\int f_r(x+u)\phi_{l,k}(x)dx\\ &+ \frac{1}{\sqrt{n}}\sum_{j,k} w^{(s)}_{j,k}\psi_{j,k}(u/n)\int f_r(x+u)\psi_{j,k}(x)dx.\end{aligned}$$

Similarly, $w^{(r)}_{u,n}$ *can be replaced by* $\tilde{w}^{(r)}_{u,n}$ *and hence also* $w^*_{u,n}$ *by* $\tilde{w}^*_{u,n}$.
Defining $\sigma_n^2 := var\,[\sum_{u=1}^n \tilde{w}^*_{u,n}\epsilon^*_u]$, *we have* $\sigma_n^2 \to \sigma_{rs}(u,u)$. *Since* $w^{(s)}_{j,k} = 0$,

$j > J_0$,

$$\max_{1\le u\le n} |w^{(s)}_{u,n}| = \max_{1\le u\le n} \Bigg| \frac{1}{\sqrt{n}} \sum_k \phi_{l,k}(u/n) \int f_r(x+u)\phi_{l,k}(x)dx$$
$$+ \frac{1}{\sqrt{n}} \sum_{j,k} w^{(s)}_{j,k}\psi_{j,k}(u/n) \int f_r(x+u)\psi_{j,k}(x)dx \Bigg|$$
$$\le \frac{A}{\sqrt{n}} \sum_k \int |f_r(x+u)\phi_{l,k}(x)|dx$$
$$+ \frac{A}{\sqrt{n}} \sum_{j,k}^{J_0} \int |f_r(x+u)\psi_{j,k}(x)|dx,$$
$$= O(n^{-1/2}),$$

where A is a suitable constant, we have

$$\max_{1\le u\le n} |w^*_{u,n}| = \max_{1\le u\le n} |w^{(r)}_{u,n} + w^{(s)}_{u,n}| \to 0.$$

From theorem 2 and the approximation of $w^*_{u,n}$ by $\tilde{w}^*_{u,n}$ we obtain var $[\sum_{u=1}^n \tilde{w}^*_{u,n}\epsilon^*_u] = O(1)$. Hence,

$$\max_{1\le u\le n} \frac{|w^*_{u,n}|}{\sigma_n} \to 0, \ \ n\to\infty.$$

Then theorem 4 in Beran and Feng (2001) together with (A3) imply

$$\frac{\sum_u w^*_{u,n}\epsilon^*_u}{\sigma_n} \xrightarrow{d} \mathcal{N}(0,1),$$

and hence

$$\sqrt{n}(\hat{\gamma}_{rs}(u) - \gamma_{rs}(u)) \xrightarrow{d} \mathcal{N}(0, \sigma_{rs}(u,u)).$$

Proof 8 *(Theorem 3) Applying the Cramér-Wold device, we show that for any constants θ_l, $l = 1,\dots,q,\ q\in\mathbb{N}$, not all equal to zero,*

$$n^{\frac{1}{2}} \sum_{l=1}^q \theta_l(\hat{\gamma}_{rs}(u_l) - \gamma_{rs}(u_l)) \xrightarrow{d} \mathcal{N}(0, \widetilde{\sigma}^2),$$

where $\widetilde{\sigma}^2 := \sum_{l=1}^{q}\sum_{m=1}^{q}\theta_l\theta_m\sigma_{rs}(u_l,u_m) > 0$. *Note that*

$$\begin{aligned} & n^{\frac{1}{2}}\sum_{l=1}^{q}\theta_l(\hat{\gamma}_{rs}(u_l)-\gamma_{rs}(u_l)) \\ & = n^{\frac{1}{2}}\underbrace{\int[\widetilde{\epsilon}_n^{(s)}(x)\sum_{l=1}^{q}\theta_l f_r(x+u_l)}_{=:A}+\underbrace{f_s(x)\sum_{l=1}^{q}\theta_l\widetilde{\epsilon}_n^{(r)}(x+u_l)}_{=:B} \\ & \underbrace{+\,\widetilde{\epsilon}_n^{(s)}(x)\sum_{l=1}^{q}\theta_l\widetilde{\epsilon}_n^{(r)}(x+u_l)]dx}_{=:C}\,. \end{aligned} \tag{40}$$

As shown earlier, the first part (A) *can be written approximately as*

$$\begin{aligned} \sum_{u=1}^{n}\epsilon_s(u)\Bigg[& n^{-\frac{1}{2}}\sum_k\phi_{l,k}(u/n)\sum_{l=1}^{q}\theta_l\int f_r(x+u_l)\phi_{l,k}(x)dx \\ & + n^{-\frac{1}{2}}\sum_{j,k}\hat{w}_{j,k}^{(s)}\psi_{j,k}(u/n)\sum_{l=1}^{q}\theta_l\int f_r(x+u_l)\psi_{j,k}(x)dx\Bigg], \end{aligned}$$

where the term in brackets is a weight function for the individual $\epsilon_s(u)$. *We denote the weights by* $w_{u,n}^{(s^*)}$ *and the analogous weights of term* (B) *by* $w_{u,n}^{(r^*)}$. *As above,* (C) *converges to 0 almost surely.*

Note that $\widetilde{\sigma} := \sqrt{var\left[\sum w_{u,n}^{(r^*)}\epsilon_r(u)+\sum w_{u,n}^{(s^*)}\epsilon_s(u)\right]}$. *Then a simple consequence of previous arguments is that*

$$\max_{1\le u\le n}\frac{|w_{u,n}^{(r^*)}+w_{u,n}^{(s^*)}|}{\widetilde{\sigma}}\to 0.$$

Hence, the Central Limit Theorem for the weighted sum holds such that

$$\sum w_{u,n}^{(r^*)}\epsilon_r(u)+\sum w_{u,n}^{(s^*)}\epsilon_s(u)\xrightarrow{d}N(0,\widetilde{\sigma}^2).$$

By applying the Cramér-Wold Device, this is equivalent to finite collections of $\hat{\gamma}_{rs}(\mathbf{u})$ *being asymptotically jointly normal with mean zero and covariance matrix* $(\sigma_{rs}(u_l,u_m))$, $l,m=1,\ldots,q$, *i.e.*

$$n^{\frac{1}{2}}(\hat{\gamma}_{rs}(\mathbf{u})-\gamma_{rs}(\mathbf{u}))\to\mathcal{N}(0,(\sigma_{rs}(u_l,u_m))_{1\le l,m\le q}). \tag{41}$$

Proof 9 *(theorem 4) Let P_n be the probability distribution of $\hat{\gamma}_{rs}^{(n)}(u)$ in $C[-1,1]$ with corresponding σ-algebra $\mathcal{C}$ and let $C[-1,1]$ be given the uniform topology induced by the metric $d(f,g) = \sup_{-1 \le t \le 1} |f(t) - g(t)|$. Then for any Borel subset B: $P_n(B) = P(\hat{\gamma}_{rs}^{(n)}(u) \in B)$. It remains to prove tightness of the family P_n of probability distributions of $\hat{\gamma}_{rs}^{(n)}(u)$ (Billingsley 1968). Define the modulus of continuity of an element $f \in C[-1,1]$ by*

$$w_f(\delta) = w(f,\delta) = \sup_{|s-t|<\delta} |f(s) - f(t)|, \quad 0 < \delta \le 2.$$

Then the sequence $\{P_n\}$ is tight if and only if

(i) for each positive η there exists a constant A such that

$$P_n(|\hat{\gamma}_{rs}^{(n)}(0)| > A) \le \eta, \;\; n \ge 1, \tag{42}$$

(ii) For each positive η_1 and η_2, there exists δ $(0 < \delta < 2)$ and $n_0 \in I\!N$ such that

$$P_n(w_{\hat{\gamma}_{rs}^{(n)}}(\delta) > \eta_1) \le \eta_2, \;\; n \ge n_0. \tag{43}$$

We first note that the probability distributions of $\hat{\gamma}_{rs}^{(n)}(0)$ are tight.

$$\begin{aligned}\hat{\gamma}_{rs}^{(n)}(0) = &\int f_r(x) f_s(x) dx + \int f_r(x) \widetilde{\epsilon}_n^{(s)}(x) dx \\ &+ \int f_s(x) \widetilde{\epsilon}_n^{(r)}(x) dx + \int \widetilde{\epsilon}_n^{(r)}(x) \widetilde{\epsilon}_n^{(s)}(x) dx.\end{aligned}$$

Lemma 2 shows that all parts of the sum in $\hat{\gamma}_{rs}^{(n)}(0)$ converge to 0 almost surely except $\int f_r(x) f_s(x) dx$. Tightness of $\hat{\gamma}_{rs}^{(n)}(0)$ is not influenced by $\int f_r(x) f_s(x) dx$ and each of the other parts of the sum converges to 0 in probability due to lemma 2. This implies that for each $A > 0$ and each $\eta > 0$ there exists an $n_0 \in I\!N$ with e.g. $P_n(|\int f_s(x) \widetilde{\epsilon}_n^{(r)}(x) dx| > A) < \eta$ for all $n \ge n_0$. For the first indices $1, \ldots, n_0 - 1$ the statement is trivial since finitely many random variables are always tight. For the second condition note that $w_{\hat{\gamma}_{rs}^{(n)}}(\delta) = \sup_{|u_2 - u_1| < \delta} |\hat{\gamma}_n(u_2) - \hat{\gamma}_n(u_1)|$. The continuity of the integral together with the results of lemma 2 imply that there exists an n_0 such that condition (ii) is satisfied for all $\eta_1, \eta_2 > 0$.
Therefore, P_n is tight such that $\hat{\gamma}_{rs}^{(n)}(u)$ converges to a Gaussian process where the finite dimensional distributions are given in equation (41).

Proof 10 *(theorem 5) Define* $X_n(u) := \sqrt{n}(\hat{\gamma}_{rs}(u) - \gamma_{rs}(u))$*. The preceeding discussion has shown that* X_n *is a random element in* $C[-1, 1]$ *that converges in distribution to a stochastic process whose finite dimensional distributions are asymptotically normal. Theorem 5 then follows by the continuous mapping theorem (see e.g. Pollard 1984).*

References

[1] Abramovich, F., Sapatinas, T. and Silverman, B.W., Wavelet Thresholding via a Bayesian Approach, *Journal of the Royal Statistical Society*, Vol. 60, Part 4, 725-749, 1998.

[2] Bruce, A. and Gao,. H.-Y., Applied Wavelet Analysis with S-Plus, Springer, New-York, 1996.

[3] Beran, J. and Feng, Y., Local Polynomial Estimation with a FARIMA-GARCH Error Process, *Bernoulli*, Vol. 7, No. 5, 2001.

[4] Beran, J. and Ocker, D. (1999). SEMIFAR forecasts, with applications to foreign exchange rates. *Journal of Statistical Planning and Inference, 80, 137-153.*

[5] Billingsley, P., Convergence of Probability Measures, John Wiley & Sons, Inc., New York, 1968.

[6] Brillinger, D., Uses of Cumulants in Wavelet Analysis, *Proc. SPIE, Advanced Signal Processing*, Vol. 2296, 2-18, 1994.

[7] Brillinger, D.R., Some Uses of Cumulants in Wavelet Analysis, *Nonparametric Statistics*, Vol. 6, 93-114, 1996.

[8] Daubechies, I., Ten Lectures on Wavelets, Society for Industrial and Applied Mathematics, Philadelphia, 1999.

[9] Donoho, D.L. and Johnstone, I.M., Ideal Spatial Adaption by Wavelet Shrinkage, *Biometrika*, Vol. 81, No.3, 1994.

[10] Donoho, D.L. and Johnstone, I.M., Adapting to Unknown Smoothness Via Wavelet Shrinkage, *Journal of the American Statistical Association*, Vol. 90, 1995.

[11] Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.-K., and Stanley, H.E., PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals, *Circulation*, Vol. 101(23), e215-e220, 2000.

[12] Grenander, U. and Rosenblatt, M., Statistical analysis of stationary time series, Wiley, New York, 1957.

[13] Heiler, S. und Feng, Y. (2000). Data-driven decomposition of seasonal time series. *Journal of Statistical Planning and Inference*, 91, 351 - 363.

[14] Ichimaru, Y. and Moody G.B., Development of the polysomnographic database on CD-ROM. *Psychiatry and Clinical Neurosciences*, Vol. 53, 175-177, 1999.

[15] Johnstone, I.M. and Silverman, B.W., Wavelet Threshold Estimators for Data with Correlated Noise, *Journal of the Royal Statistical Society B*, Vol. 59, No. 2, 319-351, 1997.

[16] Nason, G.P., Wavelet shrinkage using cross-validation, *Journal of the Royal Statistical Society B*, Vol. 58, 463-479, 1996.

[17] Percival, D.P. and Walden, A.T., Wavelet methods for time series analysis, Cambridge University Press, Cambridge, 2000.

[18] Pollard, D.: Convergence of Stochastic Processes, 2^{nd} Edition, Springer, New York, 1984.

[19] Polya G. and Szegö G., Aufgaben und Lehrsätze aus der Analysis, Erster Band, 3. Auflage, Springer, Berlin, 1964.

[20] Priestley, M.B., Spectral Analysis and Time Series, 6^{th} Edition, Academic Press, London, 1989.

[21] Yajima, Y., Asymptotic Properties of the LSE in a Regression Model with Long-Memory Stationary Errors, *The Annals of Statistics*, Vol. 19, No. 1, 158-177, 1991.

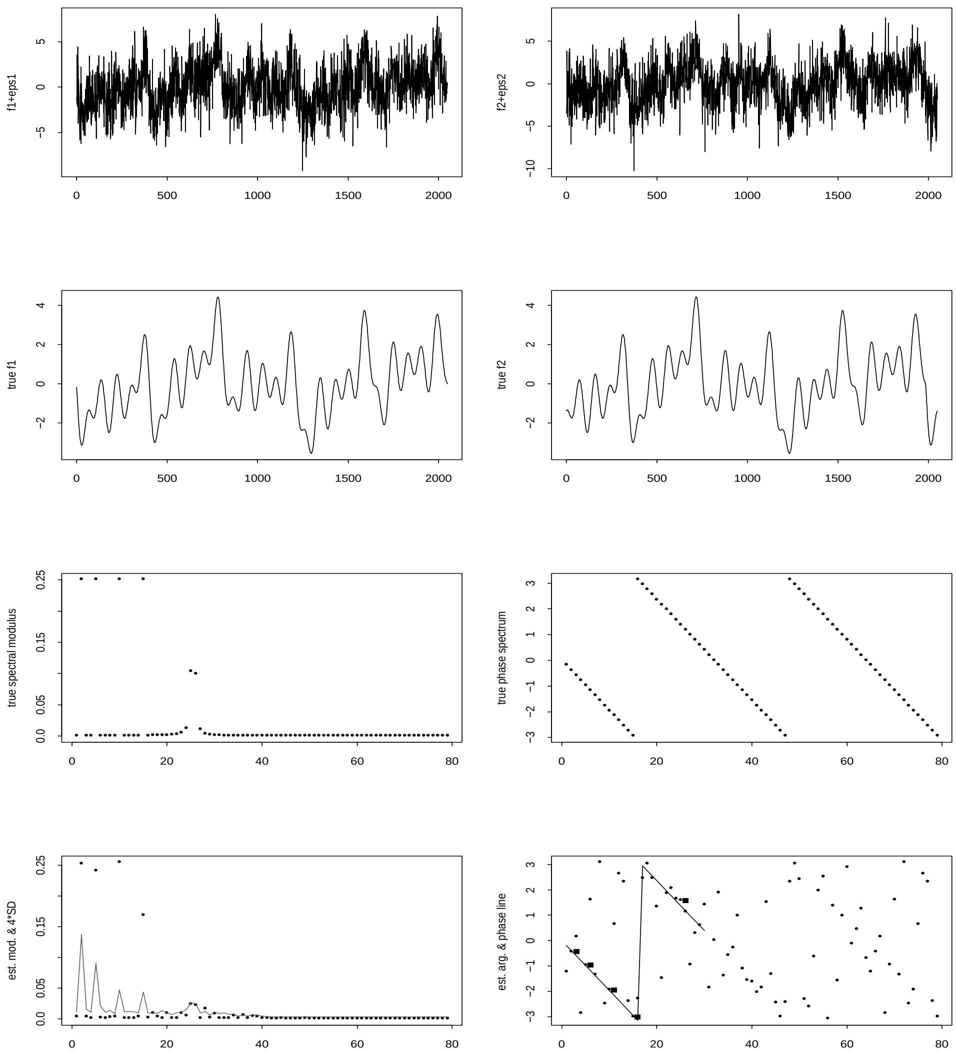

Figure 1: Simulated series $(Y_1,\, Y_2)$ of length $n = 2048$ (figures 1a,b), true trend functions $(f_1,\, f_2)$ (figures 1c,d), amplitude and phase spectrum (figures 1e,f) and estimated amplitude and phase spectrum (figures 1g,h). The horizontal line in figure 1g represents $4\sqrt{\mathrm{var}(\kappa^*_{rs}(j))}$, the solid line in figure 1h corresponds to the estimated phase line.

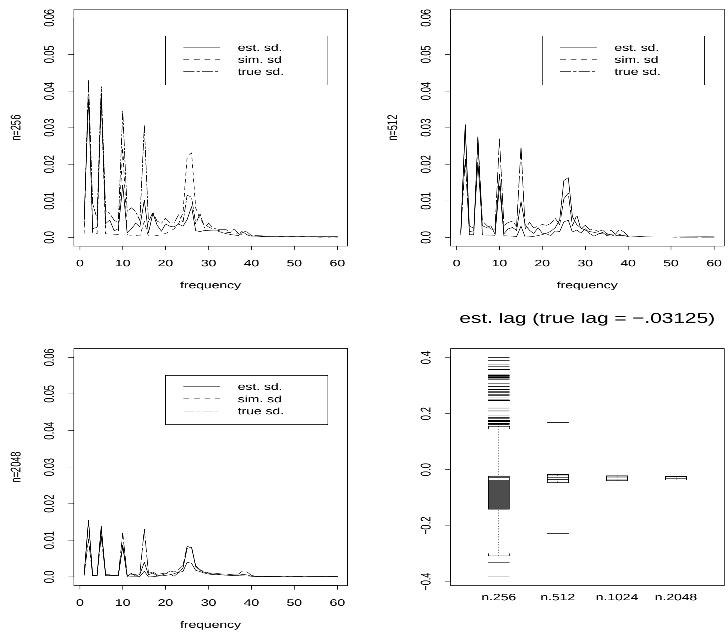

Figure 2: Simulated, estimated and true asymptotic standard deviation of the amplitude spectrum. Boxplots of the lag estimates are also given for $n = 256, 512, 1024$ and $n = 2048$.

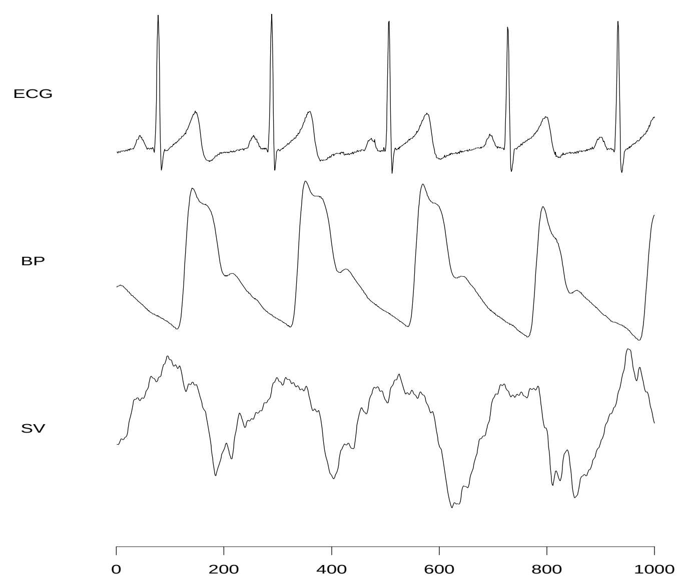

Figure 3: Electrocardiogram (ECG), blood pressure (BP) and cardiac stroke volume (SV) of a sleeping person over a period of 5 seconds (1000 observations).

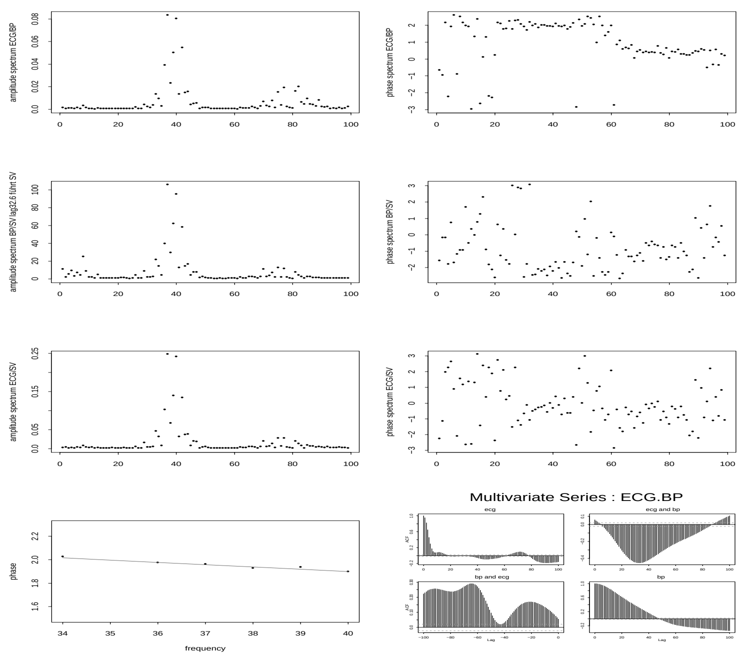

Figure 4: Amplitude and phase spectrum between the shrinkage estimates of ECG, BP and SV. The two bottom panels show the phase spectrum (lower left) between ECG and BP for $j \in [34, 40]$ and the cross covariance (lower right).