Ann. Inst. Statist. Math.
Vol. 54, No. 1, 171–200 (2002)

THE SLEX MODEL OF A NON-STATIONARY RANDOM PROCESS

HERNANDO OMBAO[1*], JONATHAN RAZ[2**],
RAINER VON SACHS[3***] AND WENSHENG GUO[2****]

[1]*Departments of Statistics and Psychiatry, University of Pittsburgh, Pittsburgh, PA 15260, U.S.A.*
[2]*Division of Biostatistics, University of Pennsylvania, Philadelphia, PA 19104, U.S.A.*
[3]*Institut de Statistique, Université Catholique de Louvain, Voie du Roman Pays 20, B-1348, Louvain-la-Neuve, Belgium*

(Received October 15, 2000; revised July 3, 2001)

Abstract. We propose a new model for non-stationary random processes to represent time series with a time-varying spectral structure. Our SLEX model can be considered as a discrete time-dependent Cramér spectral representation. It is based on the so-called Smooth Localized complex EXponential basis functions which are orthogonal and localized in both time and frequency domains. Our model delivers a finite sample size representation of a SLEX process having a SLEX spectrum which is piecewise constant over time segments. In addition, we embed it into a sequence of models with a limit spectrum, a smoothly in time varying "evolutionary" spectrum. Hence, we develop the SLEX model parallel to the Dahlhaus (1997, *Ann. Statist.*, **25**, 1–37) model of local stationarity, and we show that the two models are asymptotically mean square equivalent. Moreover, to define both the growing complexity of our model sequence and the regularity of the SLEX spectrum we use a wavelet expansion of the spectrum over time. Finally, we develop theory on how to estimate the spectral quantities, and we briefly discuss how to form inference based on resampling (bootstrapping) made possible by the special structure of the SLEX model which allows for simple synthesis of non-stationary processes.

Key words and phrases: Bootstrap, Fourier functions, Haar wavelet representation, locally stationary time series, periodograms, SLEX functions, spectral estimation, stationary time series.

1. Introduction

A zero mean stationary random process $X_t (t = 0, \pm 1, \dots)$ can be viewed as a sum of an infinite number of randomly weighted Fourier complex exponentials through the use of the Cramér representation:

$$X_t = \int_{-1/2}^{1/2} A(\omega) \exp(i2\pi\omega t) dZ(\omega) \tag{1.1}$$

*H. Ombao was supported, in part, by grants NIMH 62298 and 30915.

**J. Raz passed away on February 23, 2000. His research was supported in part by grant MH59817. —This paper would not have been possible without Jonathan Raz whose ideas form the foundation of this work. We dedicate this paper in honor of his memory.

***R. von Sachs was supported, in part, by grants NIMH 62298 and "Projet d'Actions de Recherche Concertées" No. 98/03-217 of the Belgian government.

****W. Guo was supported, in part, by grants NIMH 62298 and NCI CA84438.

where $A(\omega)$ is the transfer function and $dZ(\omega)$ is a zero mean orthonormal increment random process. Both $A(\omega)$ and $Z(\omega)$ do not change with time. The spectrum of X_t is $f_{XX}(\omega) = |A(\omega)|^2$. Spectral analysis is important in the study of stationary processes because the spectrum (i) provides information on the stochastic nature of the process and (ii) gives an immediate physical interpretation as the power frequency distribution (Priestley (1981)). For stationary processes, the distribution of power over frequency does not change with time.

The assumption of stationarity, however, is valid only as an approximation to the real situation. Furthermore, this assumption may not be reasonable in modeling data sets that are collected from a variety of fields such as geophysics, speech, economics and medicine. In neurology, for instance, electroencephalograms (EEGs) or brain waves cannot be adequately modeled as stationary processes because the waves oscillate with greater amplitude and at higher frequency during an epileptic seizure than before or after the seizure (see Ombao *et al.* (2001*a*)). Thus, there is a need to develop models whose spectral properties change with time.

In this paper, we propose a model of a zero mean non-stationary random process that has a Cramér-like representation in terms of the SLEX (Smooth Localized Complex EXponential) functions. The SLEX functions are complex-valued, orthogonal and localized in both time and frequency domains (Wickerhauser (1994), Chapter 4). Thus, they can be considered as localized versions of the Fourier complex exponential functions. We first discuss some of the models of non-stationary processes that have appeared in the literature.

Priestley (1965) first introduced the class of non-stationary random processes whose spectral properties change slowly over time. Dahlhaus (1997) refined the ideas in Priestley and introduced the class of locally stationary process. From an intuitive point of view, a random process is said to be locally stationary if one can form an interval that is approximately stationary about each time point, with a smooth change from one interval to the next. More formally, in Dahlhaus (1997), a zero mean random process $X_{t,T} (t = 0, \ldots, T-1)$ is said to be locally stationary if it admits a representation

$$X_{t,T} = \int_{-1/2}^{1/2} A^0_{t,T}(\omega) \exp(i2\pi\omega t) dZ(\omega), \qquad t = 0, \ldots, T-1, \tag{1.2}$$

where $A^0_{t,T}(\omega)$ is close to a smoothly in time varying transfer function $A(t/T, \omega)$, and where $dZ(\omega)$ is a zero mean orthonormal increment random process. This asymptotic framework of an array of processes $\{X_{t,T}\}$ guarantees that increasing the sample size T leads to more data at a local structure in time, i.e., more information is obtained about the same unknown statistical quantities $A(t/T, \omega)$ being estimated.

Our model of a discrete non-stationary random process uses, as stochastic building blocks, the SLEX vectors (discretized SLEX functions) which are (i) orthonormal; (ii) localized in time and frequency; and (iii) a generalization of the tapered Fourier vectors. The SLEX vectors are obtained by applying two specially constructed windows on the Fourier vectors. Thus, the SLEX basis vectors can be made arbitrarily close to the tapered Fourier vectors. They are localized in the time domain because the windows have compact support. Furthermore, they are localized in the frequency domain because the windows are smooth. Thus, the proposed model uses a localized basis that is appropriate for random processes whose statistical properties vary over time. Moreover, under the SLEX model, one can define the SLEX spectrum, a spectrum that is a decomposition of power over time and frequency. In doing this, the SLEX model remains in the spirit of traditional spectral analysis because it gives a spectrum that is a time-dependent analogue of the classical spectrum for stationary processes.

The Dahlhaus model, in contrast, uses the Fourier basis functions which are not localized in time. The localization is provided by the time-varying transfer function $A^0_{t,T}(\omega)$. Thus, it does not give an explicit segmentation of the time-frequency plane. In contrast, the SLEX method, for each T, does provide this explicit segmentation with a time-varying spectrum which is piecewise constant along time segments of stationarity. Simultaneously, as the Dahlhaus model, it allows, in the asymptotic limit, for a smoothly time-varying spectrum. However, the SLEX model is actually more general than the Dahlhaus model in a sense that it allows for "inhomogeneous regularity" over time, that is the degree of smoothness of the smoothly time-varying spectrum is allowed to differ across the different regions over time.

To achieve modeling of this inhomogeneous regularity in our SLEX model, we use a wavelet expansion of the SLEX spectrum over time to define both the complexity and the regularity of this spectrum as a function of time. Wavelets inherently offer the possibility of modeling functions with low regularity, in particular spatially inhomogeneous functions such as, e.g., functions of bounded variation (Daubechies (1992)). Hence, as the variation of the SLEX spectrum over time controls the degree of non-stationarity of the underlying process, this approach based on wavelets (foremost we think of using Haar wavelets) seems to be adequate for the purpose of modeling epileptic EEG datasets which typically show abrupt changes over time.

Note that our framework of using the SLEX functions, i.e. specially localized Fourier functions, as building blocks of the model, is substantially different from yet another related concept of local stationarity, Nason *et al.* (2000) introduced a new class of non-stationary random processes that have a localized representation in exactly the same spirit as the Dahlhaus model where the Fourier basis is replaced by a wavelet basis, localized both in discrete time and scale (or wavelength, and hence being reciprocal to frequency). This allows for defining a wavelet spectrum which is a measure of power of the process at a particular scale and location and, hence, gives a time-scale decomposition of power, rather than a time-frequency decomposition as treated in this paper.

We have developed in Ombao, Raz, von Sachs and Malow (2001*a*) [ORvSM] a computationally efficient methodology for analyzing bivariate non-stationary time series which we call the "Auto-SLEX method". The Auto-SLEX method is an automatic statistical procedure that simultaneously segments the non-stationary time series and estimates the time-varying spectra and coherence. The Auto-SLEX method completely parallels the Fourier-based spectral methods for stationary processes. Thus, it yields results that are easy to interpret because they are extensions of the stationary methods to the non-stationary situation. The Auto-SLEX method, similar to the Fourier methods, first computes the SLEX periodogram matrices and smooths them over frequency using a kernel smoother whose span is automatically selected by a generalized cross validation criterion. The span selection method is described in Ombao *et al.* (2001*b*). The SLEX model presented in this paper complements our methodology since we can now derive a more complete theoretical treatment of the asymptotic properties of the method. Moreover, it allows to establish the estimation theory as well as statistical inference for the method. With our new approach being based on an explicitly accessible synthesis equation, for each finite T, straightforward simulation (or resampling) of non-stationary processes is now possible. This particular interesting property is neither shared by the

Dahlhaus model nor by any other known model of local stationarity (Nason *et al.* (2000), Donoho *et al.* (1998)). For a detailed development on statistical inference based on bootstrapping we refer to our companion paper Ombao *et al.* (2000), here we only give a brief summarizing discussion.

Our paper is organized as follows. In Section 2, we state and discuss the Dahlhaus model of a locally stationary random process. As a preparation for the presentation of the SLEX model in Section 4, we introduce the SLEX vectors and transform in Section 3. In our main section, Section 4, we introduce our new SLEX model of non-stationary random processes. We first discuss the part of the model for fixed sample size T, i.e. we establish a sequence of modeled processes, including a finite sample size SLEX spectrum f_T which is piecewise constant over time. In the second part, we define a "limit" spectrum, the "evolutionary SLEX spectrum" which enjoys some smoothness properties and relates to the evolutionary spectrum of the Dahlhaus model. In particular, we establish asymptotic relationships between this smooth spectrum f and the piecewise constant spectrum f_T. We finish this section by delivering our complete model of a "locally stationary SLEX process". Before we asymptotically connect this model to the Dahlhaus model in Section 6, we discuss in Section 5 how to estimate the evolutionary SLEX spectrum. In Section 7, we discuss inference in the SLEX model that is based on the bootstrap. Finally we close up with a conclusion and discussion section which opens the view towards future applications of our new model. All proofs are deferred to the Appendix.

2. The Dahlhaus model of local stationarity

We first give the complete definition of the Dahlhaus (1997) model. As our only emphasis in this paper is on modeling the spectral structure we restrict to mean zero processes.

DEFINITION 2.1. A sequence of zero mean stochastic processes $\{X_{t,T}\}_{t=1,\dots,T}$, $T \geq 1$, is called locally stationary with transfer function A^0 if there exists a representation

$$X_{t,T} = \int_{-1/2}^{1/2} A^0_{t,T}(\omega) \exp(i2\pi\omega) dZ(\omega) \tag{2.1}$$

where

(i) $Z(\omega)$ is a stochastic process on $[-1/2, 1/2]$ with $\overline{Z(\omega)} = Z(-\omega)$ and $\text{cum}\{dZ(\omega_1), \dots, dZ(\omega_k)\} = \eta(\sum_{j=1}^{k} \omega_j) v_k(\omega_1, \dots, \omega_{k-1}) d\omega_1 \dots d\omega_k$ where $\text{cum}\{\dots\}$ denotes the cumulant of k-th order; $v_1 = 0$, $v_2(\omega) = 1$, $|v_k(\omega_1, \dots, \omega_{k-1})| \leq \text{constant}_k$ and $\eta(\omega) = \sum_{j=-\infty}^{j=\infty} \delta(\omega + 2\pi j)$ is the period 2π extension of the Dirac delta function.

(ii) There exists a constant Q and a 2π periodic function smooth $A : [0,1] \times \mathbf{R} \to \mathbf{C}$ with $A(u, -\omega) = \overline{A(u,\omega)}$ and

$$\sup_{t,\omega} |A^0_{t,T}(\omega) - A(t/T, \omega)| \leq QT^{-1} \tag{2.2}$$

for all T. $A(u,\omega)$ is assumed to be continuous in u.

Here, smoothness of A means some regularity of A, e.g. Lipschitz, both as a function of time and of frequency. We will later adopt the smoothness conditions of Assumptions 1 and 2 in Subsection 4.2.

Finally, the unique and non-negative time-varying spectrum of the Dahlhaus model at time $u \in [0,1]$ and frequency $\omega \in [-1/2, 1/2]$ is defined to be $f_{XX}^D(u,\omega) = |A(u,\omega)|^2$.

We note that the first argument of $A(t/T,\omega)$ is rescaled to live on the unit interval. Increasing the number of observations, T, does not mean looking into the future. Rather, this asymptotic framework, i.e. the concept of a doubly-indexed sequence of processes $\{X_{t,T}\}$, allows for more data to be observed at a local structure and to do asymptotic inference starting from a single realization rather than using replications of $X_{t,T}$, $(t = 0, \dots, T-1)$. This is possible by equation (2.2) which says that $A_{t,T}^0(\omega) \approx A(t/T,\omega)$, i.e., the smoothness of A in u controls the change of $A_{t,T}^0$ as a sequence in t such that it is allowed to change only slowly over time. Estimation theory then parallels the one of nonparametric regression with an asymptotically denser and denser design on $(0,1)$ (e.g. kernel estimation, as in our Section 5).

Various examples for locally stationary processes following this model are to be found in Dahlhaus (1997). Here we give only two: a) modulated processes, $X_{t,T} = \sigma(t/T) \cdot Y_t$, where $\{Y_t\}$ is a stationary process and $\sigma(u)$ some smooth function on $(0,1)$; and b) time-varying ARMA processes where the transfer function can be modeled by a rational function in the frequency ω with smoothly in time varying MA and AR coefficients, respectively. Here, for the moving average part $A_{t,T}^0(\omega) = A(t/T,\omega)$ whereas, for the autoregressive part, $A_{t,T}^0(\omega) \approx A(t/T,\omega)$. For details we again refer to Dahlhaus (1997).

We have given the Dahlhaus model as a frame of reference in developing our SLEX model of a non-stationary random process. One basic ingredient of the SLEX model is the use of the SLEX basis vectors as stochastic building blocks. In the next section, we will briefly discuss the construction of the SLEX basis vectors and the computation of the SLEX coefficients.

3. The SLEX transform

3.1 *SLEX basis vectors*

Fourier basis functions are ideal at representing stationary random processes because they are perfectly localized in frequency and the spectral properties of stationary processes do not change with time. Fourier basis functions, however, cannot adequately represent processes whose spectral properties evolve with time. In order to alleviate the time localization problem, smooth compactly supported windows have been applied to the Fourier basis functions (Daubechies (1992)). Windowed Fourier functions, however, are generally no longer orthogonal. In fact, the Balian-Low theorem says that there does not exist a smooth window such that the windowed Fourier basis vectors are both orthogonal and localized in time and frequency (Wickerhauser (1994)). The SLEX basis functions, on the other hand, are simultaneously orthogonal and localized in time and frequency. They evade the Balian-Low obstruction because they are constructed by applying a projection operator, rather than a window, on the complex exponentials. It turns out that the action of a projection operator on a periodic function is identical to applying two specially constructed smooth windows to the Fourier basis functions. It should also be noted that the SLEX basis functions furnish a complete basis of the space $L_2(\mathbb{R})$.

A SLEX basis function $\phi_\omega(u)$ has the form

$$\phi_\omega(u) = \Psi_+(u)\exp(i2\pi\omega u) + \Psi_-(u)\exp(-i2\pi\omega u) \tag{3.1}$$

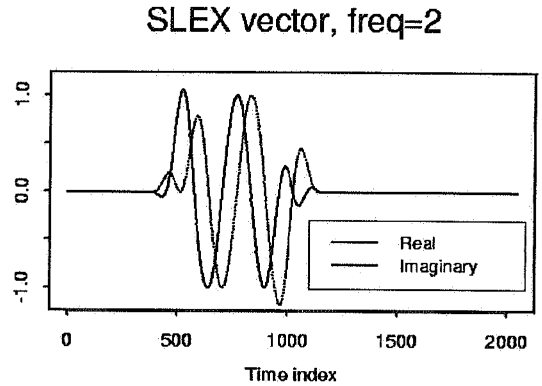

Fig. 1. A SLEX basis vector oscillating at frequency 2 and having support on the block with time points indexed by $\{512, \ldots, 1023\}$.

where $\omega \in [-1/2, 1/2]$ and $\Psi_+(u)$ and $\Psi_-(u)$ are the particularly constructed two smooth (real-valued) windows. We refer again to Wickerhauser (1994) for more details. The SLEX basis function has support on $[-\eta, 1+\eta]$, where $0 < \eta < 0.5$. Thus, SLEX functions at different dyadic blocks overlap. Though they overlap, the SLEX basis functions remain orthogonal. SLEX basis functions generalize directly to orthogonal SLEX basis vectors for representing time series. Define $\alpha_1 > \alpha_0$ to be two integer time points; $|S| = \alpha_1 - \alpha_0$; and the overlap $\epsilon = [\eta|S|]$, where $[\cdot]$ denotes the greatest integer less than or equal to its argument. The support of SLEX vectors on block S consist of time points defined on S and the overlap. We denote the support to be $\tilde{S} = \{\alpha_0 - \epsilon, \ldots, \alpha_0, \ldots, \alpha_1 - 1, \alpha_1 - 1 + \epsilon\}$. A SLEX basis vector defined on block S has elements $\{\phi_{S,\omega_k}(t)\}$, where

$$\begin{aligned} \phi_{S,\omega_k}(t) &= \phi_{\omega_k}\left((t-\alpha_0)/|S|\right) \\ &= \Psi_+\left((t-\alpha_0)/|S|\right) \exp\left\{i2\pi\omega_k(t-\alpha_0)\right\} \\ &\quad + \Psi_-\left((t-\alpha_0)/|S|\right) \exp\left\{-i2\pi\omega_k(t-\alpha_0)\right\}, \end{aligned} \tag{3.2}$$

and where $\omega_k = k/|S|$, $k = -|S|/2 + 1, \ldots, |S|/2$. In Fig. 1, we show a plot of the real and imaginary parts of a SLEX basis vector oscillating at frequency index 2 and having support at time points $\{512, \ldots, 1023\}$.

We now state a lemma on the properties of the SLEX vectors that will be used in developing estimation theory for the SLEX model. The proof can be found in Ombao (1999).

LEMMA 3.1. *Let $\phi_{S,\omega_k}(t)$ be a SLEX basis vector on block S and with frequency ω_k. The following properties are true:* (i) *the real and imaginary parts of $\phi_{S,\omega_k}(t)$ are orthogonal;* (ii) *$\phi_{S,\omega_k}(t)$ is orthogonal to its conjugate and* (iii) *for any $\omega_k \neq \omega_{k'}(k, k' = -|S|/2 + 1, \ldots, |S|/2)$, $\phi_{S,\omega_k}(t)$ and $\phi_{S,\omega_{k'}}(t)$ are orthogonal.*

In total, on each block S, the SLEX basis vectors completely parallel the discrete orthonormality properties of the classical complex exponentials over time and over frequency, and, in particular, at each time point the squared windows sum up to one, i.e. $\Psi_+^2(t) + \Psi_-^2(t) = 1$.

3.2 *Computing the SLEX transform*

The SLEX transform is a dyadic transform in the sense that it divides the time axis of the time-frequency plane in a dyadic manner. At a time resolution level j, the time series is divided into 2^j blocks. The SLEX coefficients are then computed at each of these blocks. Note that the SLEX transform shares this similar feature with the cosine packet transform. Suppose that the series $X_{t,T}$, $t = 0, \ldots, T-1$, has length $T = 2^K$ for some $K \gg 0$. The SLEX transform is computed with a pre-specified finest time scale (or deepest level) $J \leq K - 1$. At scale j ($j = 0, \ldots, J$), we define the blocks to be $S(j, b)$ where $b = 0, \ldots, 2^j - 1$. There are $M_j := |S(j,b)| = |S(j)| = T/2^j$ observations on each block at level j. When it is clear from the context, the subscripts j and b will be dropped.

In addition to being dyadic, the SLEX transform also smoothly partitions the time axis by allowing the SLEX vectors defined on adjacent blocks to have overlapping support. Despite the overlap, the SLEX vectors are still orthogonal under very mild conditions on the amount of overlap which we now discuss. Let ϵ_j be the amount of overlap between the SLEX vectors at adjacent blocks at level j. In order for the SLEX basis vectors to be orthogonal, it is sufficient to set $\epsilon_j = \epsilon_J = M_J/2$ for all $j = 0, \ldots, J-1$. When computing the SLEX coefficients on block $S(j,b)$, the range of the summation is over the "extended" block $\tilde{S} = \tilde{S}(j,b)$. The SLEX coefficients on block $S = S(j,b)$ are defined as:

$$\hat{\theta}_{S,k,T} = \frac{1}{\sqrt{M_j}} \sum_{t \in \tilde{S}} X_{t,T} \overline{\phi_{S,\omega_k}(t)}, \qquad k = -M_j/2+1, \ldots, M_j/2. \tag{3.3}$$

The SLEX periodogram, an analogue of the Fourier periodogram, is defined to be $\hat{\alpha}_{S,k,T} = |\hat{\theta}_{S,k,T}|^2$.

3.3 *The Best Basis*

The SLEX transform forms a library of orthonormal transforms (or a collection of orthonormal SLEX bases). In the SLEX model that we will present in the next section, we assume that the basis is already given. When the basis is not known, one may use the Best Basis Algorithm (BBA) of Coifman and Wickerhauser (1992) to search for the basis that is optimal according to some well defined criterion. Various criteria (cost functions) are discussed in detail in Wickerhauser ((1994), Chapter 8). Specific cost functions used in the Auto-SLEX method are discussed in ORvSM and in Ombao *et al.* (2000). Note that every basis in the SLEX library defines a particular segmentation in the SLEX library. The segmentation is implied in the support of the SLEX vectors that are included in the best basis. Suppose that the known basis is defined to be B_T. We set this important notation. Let $\bigcup_i S_i \sim B_T$ denote the collection of the blocks S_i that are defined in the segmentation or basis B_T.

Finally, we note that none of the blocks in $\bigcup S_i$ is a subset of (or is completely covered by) any other block in $\bigcup S_i$. Equivalently, none of the blocks in $\bigcup S_i$ can be expressed as a combination of the other blocks. This is a consequence of the useful restriction we have imposed on the overlap of the blocks in Subsection 3.2.

4. The SLEX model

We will now present the SLEX model of a non-stationary random process. In Subsection 4.1, we define the model for a fixed sample size T and discuss its ingredients which are the time-varying transfer function, the SLEX basis vectors and the orthogonal random process that is analogous to $Z(\omega)$ in the Cramér representation and the Dahlhaus model (Definition 2.1). We give the definition of the time-varying spectrum $f_T(u, \omega_k)$ (and its variants, namely the time-varying log spectrum and transfer function). In Subsection 4.2, we parallel the development of the Dahlhaus model by defining the asymptotic spectrum, $f(u, \omega)$, (and its variants) which is smooth as a function of time u and frequency ω. More precisely, the smoothness of $f(u, \omega)$ as a function of ω is governed by a Hölder condition that is uniform in u. Moreover, uniformly in ω, the smoothness of f as a function of time u is governed by controlling the decay of the coefficients in a wavelet expansion of $\log f(u, \omega)$. This representation is independent of T, but it determines the complexity of our sequence of models as T grows to infinity. As we primarily think about using Haar wavelets we actually have to start from formulating a set of conditions directly on f as a function of time u, which in particular allows for different regimes of Hölder regularity, including possible jumps at the transition points. As an important question of coherency of our model, we discuss how to relate the smooth function $f(u, \omega)$ (and its variants) to $f_T(u, \omega)$ for a given T. Finally, in Subsection 4.3, we close up in formulating our complete model of a locally stationary SLEX processs. The discussion of how to generalize to modeling SLEX spectra with higher regularity in time by using more regular wavelets instead of Haar is postponed to the discussion Section 8.

4.1 *The SLEX model for fixed T*

Suppose that we have a non-stationary time series $X_{t,T}, t = 0, \ldots, T-1$, and we want to model it by a representation of a collection of orthogonal SLEX basis vectors corresponding to a basis B_T (see Section 3 and, in particular, Subsection 3.3). The SLEX basis vectors have a fixed overlapping support. Moreover, they are allowed to have different length of support and hence may live on different time scales. Let J_T be the finest scale in B_T, i.e., it is the scale on which the basis vectors that have the narrowest time support live. We define $M_j := |S(j)| = T/2^j$ to be the number of points on each block on scale j where $j = 0, \ldots, J_T$. Correspondingly, the frequencies defined on this level are the grid frequencies $\omega_{k_j} = k_j/M_j$ for $k_j = -M_j/2 + 1, \ldots, M_j/2$.

DEFINITION 4.1. For a given T and basis $\bigcup_i S_i \sim B_T$, the SLEX model of a zero mean non-stationary random process is:

$$X_{t,T} = \sum_{i:\bigcup S_i \sim B_T} \frac{1}{\sqrt{M_i}} \sum_{k_i=-M_i/2+1}^{M_i/2} \theta_{S_i,k_i,T} \overline{\phi_{S_i,\omega_{k_i}}}(t) z_{S_i,k_i} \tag{4.1}$$

where $\theta_{S_i,k_i,T}$ are fixed complex-valued scalars with $\alpha_{S_i,k_i,T} = |\theta_{S_i,k_i,T}|^2 > 0$ and the random variables z_{S_i,k_i} are defined to be $dZ(k_i/M_i)$ where the increment process $dZ(\omega)$, $\omega \in [-1/2, 1/2]$, is zero mean and orthonormal.

Remarks. (1) Under the SLEX model, the transfer function $\theta_{S_i,k_i,T}$ is piecewise constant in time (in the block S_i) and is only a function of the frequency k_i. We compare the similarity with the Dahlhaus model where the transfer function $A^0_{t,T}(\omega)$ is also approximately constant in a small interval about t/T.

(2) Suppose that S_1 and S_2 are blocks in B_T. These blocks may live at different time scales, i.e., $j_1 \neq j_2$ and as a consequence, the frequency resolution at these blocks

may be different. Without loss of generality, suppose that $j_1 < j_2$. In this scenario, the frequency resolution at block S_2 is coarser than that at block S_1. This would imply that the set frequencies defined on block S_2 is a subset of the set of frequencies defined on block S_1. Thus, for some indices k_1^* and k_2^*, we may have $k_1^*/M_{j_1} = k_2^*/M_{j_2}$ and hence $z_{S_1,k_1^*} = z_{S_2,k_2^*}$. Therefore, quite naturally, some elements of the set $\{z_{S_i,k_i}\}$ may be correlated over blocks (i.e., over time).

(3) The SLEX representation at a particular time t can be expressed in terms of the SLEX basis vectors, in the given SLEX basis B_T, that have support only on the following blocks (i) the block, say S_0, on which t belongs; and (ii) the adjoining blocks on the left and right of S_0, which we denote S_L and S_R, respectively. We already know the fact that $\phi_{S,\omega_k}(t) = 0$ when $t \notin S_L \cup S_0 \cup S_R$. Thus, when $t \in S_0$, we can write

$$(4.2)\quad X_{t,T} = \frac{1}{\sqrt{|S_L|}} \sum_{k_L} \theta_{S_L,k_L,T} \overline{\phi_{S_L,\omega_{k_L}}(t)} z_{S_L,k_L} + \frac{1}{\sqrt{|S_0|}} \sum_{k_0} \theta_{S_0,k_0,T} \overline{\phi_{S_0,\omega_{k_0}}(t)} z_{S_0,k_0} + \frac{1}{\sqrt{|S_R|}} \sum_{k_R} \theta_{S_R,k_R,T} \overline{\phi_{S_R,\omega_{k_R}}(t)} z_{S_R,k_R}.$$

(4) To finally state an explicit example, we note that every piecewise stationary ARMA(p,q) process can easily be written in the form of a SLEX model, by setting the model coefficients $\theta_{S_i,k_i,T}$ to be equal to the classical ARMA transfer function on that block S_i, evaluated at frequency ω_{k_i}. We illustrate this along a simple moving average process of first order, piecewise stationary along two consecutive blocks: $X_{t,T} = \varepsilon_t + a(t/T)\varepsilon_{t-1}$, where $\text{var}(\varepsilon) = \sigma^2$ and where $a(u) = a_1 1_{[0,1/2]}(u) + a_2 1_{(1/2,1]}(u)$. For this example, we define $\theta_{S_i,k_i,T} := \sigma(1 + a_i \exp(i\omega_{k_i}))$, $i = 1, 2$, where $S_1 = [1, T/2]$, $S_2 = [T/2+1, T]$. To fully specify this process, we can, for example, impose the increment process $dZ(\omega)$ to be a Gaussian random process.

(5) More generally, smoothly time-varying ARMA-processes in the sense of the Dahlhaus model can also be embedded into the SLEX model approach. We will do this in Section 6 by relating the two model approaches which basically differ by the following three key ingredients. Continuous frequencies over $[-1/2, 1/2]$ are replaced by discrete grid frequencies. The complex exponential basis functions which correspond to no windowing, or simply to "multiplication with a rectangular" block respectively, are replaced by the sum over two smoothly windowed complex exponentials, where the difference between these two basis functions, however, shows only up along the "overlap" η discussed in Subsection 3.1. Within each block in time a smoothly varying transfer function in the Dahlhaus model is replaced by its average over the SLEX block. For more details we refer to the formulation and the proof of Theorem 6.1.

We now define the spectrum of the SLEX model for a given T and B_T. The spectrum is a function of the rescaled time u in $[0, 1]$ and is defined on a particular set of grid frequencies $\omega_k = \omega_k(u)$ in $[-\pi, \pi]$. To determine these, suppose that u is contained in a subinterval I of $[0, 1]$. There is some block S in $\bigcup_i S_i$ that is equivalent to I in the following manner. We say that $I \sim S$ if we have $t/T = u \in I \Leftrightarrow [uT] \in S$. Suppose now that $[uT] \in S_{i^*}$ for some block $S_{i^*} \in \bigcup_i S_i$. Then the corresponding grid frequencies are $\omega_{k_{i^*}} = k_{i^*}/M_{i^*}$, $k_{i^*} = -M_{i^*}/2 + 1, \ldots, M_{i^*}/2$. Next, we define the indicator on an interval I to be $2^{-\ell_I/2}\varphi^H_{\ell_I,m_I}$ which is the rescaled Haar scaling function (or Haar father wavelet), i.e., $\varphi_{\ell,m}(x) = 2^{\ell/2}\varphi(2^\ell x - m)$, where $\varphi(x) = 1_{[0,1)}(x)$ and $\ell \in \mathbf{N}_0$, $m = 0, \ldots, 2^\ell - 1$.

DEFINITION 4.2. For a given T and basis $\bigcup_i S_i \sim B_T$, the **spectrum** of a SLEX model at rescaled time u (as given above) is defined to be:

$$f_T(u, \omega_{k_{i^*}}) = \sum_{i:\bigcup S_i} \alpha_{S_i,k_i,T} 1_{S_i}([uT]) \tag{4.3}$$

$$= \sum_{I_i \sim S_i} \alpha_{S_i,k_i,T} 2^{-\ell_{I_i}/2} \varphi_{\ell_{I_i},m_{I_i}}(u) \tag{4.4}$$

where $k_{i^*} = -|S_{i^*}|/2+1, \ldots, |S_{i^*}|/2$, $\alpha_{S_i,k_i,T} = |\theta_{S_i,k_i,T}|^2$.

Remarks. (1) The scale ℓ_I and translate m_I are equal to the scale and translate of the SLEX segment $S \sim I$. In other words, $I = [2^{-\ell_I} m_I, 2^{-\ell_I}(m_I+1))$. The function $f_T(u, \omega_k)$ is piecewise constant as a function of continuous rescaled time u.

(2) Note that instead of using the Haar scaling functions themselves which would imply a different weighting of the spectral power along intervals of different length, we rather use the indicator functions to define this piecewise constant spectrum over time.

(3) We now show the connection between the definition of the spectrum and Equation (4.2). Note that for a fixed $u^* = t^*/T \in I^*$ so that $[u^*T] \in S^*$ and for a fixed ω_k^*,

$$f_T(u^*, \omega_k^*) = E|\frac{1}{\sqrt{|S^*|}} \sum_t X_t \phi_{S^*,\omega_{k^*}}(t)|^2 = |\theta_{S^*,k^*,T}|^2 = \alpha_{S^*,k^*,T}.$$

We have demonstrated a natural connection: the SLEX spectrum arises as the square of the SLEX transfer function. The above follows directly from the (i) orthogonality of the SLEX basis vectors: $< \phi_{S,k}, \phi_{S',k'} >= |S|\delta(S-S')\delta(k-k')$ and (ii) from the definition of the random process $z_{S,k} : E|z_{S,k}|^2 = 1$.

There are other ways of defining the spectrum in terms of $\alpha_{S,k,T}$ and the Haar father wavelets φ_{ℓ_I,m_I}, $I \sim S$. Note that since each value of u corresponds to exactly one $\alpha_{S,k,T} > 0$, $S \sim I$, and since $2^{-\ell_I/2}\varphi_{\ell_I,m_I}(u)$ is either 1 or 0, we have $f_T(u, \omega_k) > 0$. This allows a formulation of the spectrum in terms of any (even non-linear) function of f_T being still equivalent to equation (4.4). Interesting candidate functions for us are the log-spectrum and, as 'square-root' of the spectrum, the transfer function. In the following definitions, we assume, as in Definition 4.2, that $[uT] \in S_{i^*}$ for some $S_{i^*} \in \bigcup_i S_i$.

DEFINITION 4.3. The **log-spectrum** of the SLEX model for a given T and basis B_T is defined to be

$$\log f_T(u, \omega_{k_i^*}) = \sum_{I_i \sim S_i} \log(\alpha_{S_i,k_i,T}) 2^{-\ell_{I_i}/2} \varphi_{\ell_{I_i},m_{I_i}}(u), \text{ where} \tag{4.5}$$

$$k_{i^*} = -|S_{i^*}|/2+1, \ldots, |S_{i^*}|/2.$$

Note that this definition of the log-spectrum has the advantage that its companion expansion (4.8), below, on the log-scale ensures that the spectrum itself still remains non-negative.

Finally, we give a formal definition of the time-varying transfer funtion of the SLEX model. This definition is already implied in the definition of the SLEX model.

DEFINITION 4.4. The **transfer function** of the SLEX model for a given T is defined to be

$$A_T(u, \omega_{k_i^*}) = \sum_{I \sim S} \theta_{S,k,T} 2^{-\ell_I/2} \varphi_{\ell_I, m_I}(u), \qquad k_{i^*} = -|S_{i^*}|/2 + 1, \dots, |S_{i^*}|/2. \tag{4.6}$$

Note that although this last definition is more flexible than the first one, we will not use it in this work to model and estimate *univariate* SLEX processes. Definition 4.4 will prove useful only for the proof of Theorem 6.1 further below. Its full strength will be used in a multivariate treatment which aims at developing a theory of consistent estimation of cross-spectra and coherencies but which we do not want to include in this paper on presenting the SLEX-wavelet model in its own right. (We also refer to the motivation and preliminary analysis of ORvSM in the bivariate case.)

4.2 *The SLEX model with a structure for asymptotic theory*

As $T \to \infty$, we allow the SLEX model to become more complex, that is, for it to have more segments and for J_T to increase although not as fast as $K = \log_2 T$ increases, i.e. $2^{J_T}/T \to 0$ as $T \to \infty$. To develop an asymptotic theory, we need a well-defined structure for the change in complexity of the model. We also need a well-defined estimand (the so-called "evolutionary" SLEX spectrum) which does not depend on T. This goes along with imposing regularity conditions on the evolutionary SLEX spectrum as a function of (rescaled) time.

We first define the evolutionary SLEX spectrum $f(u, \omega)$ which does not depend on T. Our idea is basically to define it (actually rather its logarithm, $\log f$, to guarantee positivity of the spectrum) via a wavelet expansion, and then to show how to relate this to $\log f_T$ (and hence to f_T), for a given T. However, as we will use Haar wavelets for this wavelet representation, we first need to impose some regularity assumptions on $f(u, \omega)$, which will imply the convergence of the wavelet series pointwise in each continuous $u \in (0, 1)$.

ASSUMPTION 1. For $f(u, \omega)$ as a function of $\omega \in [-1/2, 1/2]$, uniformly in $u \in [0, 1]$ we assume a Hölder condition of order $\mu \in (0, 1]$ with constant $L > 0$

$$|f(u, \omega) - f(u, \omega^*)| \leq L|\omega - \omega^*|^\mu. \tag{4.7}$$

ASSUMPTION 2. We assume that there exists a hierarchical collection $\mathcal{I}$ of dyadic subintervals of $[0, 1]$

$$\mathcal{I} = \{I_{\ell m} = [2^{-\ell}m, 2^{-\ell}(m + 1)) : \ell = 0, 1, \dots, m = 0, \dots, 2^\ell - 1\}$$

and a subset of intervals I_ν, $I_\nu = [u_\nu, u_{\nu+1}) \in \mathcal{I}$ such that $\bigcup_\nu I_\nu = [0, 1]$ and that $f(u, \omega)$, as a function of u, is Hölder of order $0 < s_\nu < 1$ on I_ν, for all $\omega \in [-1/2, 1/2]$. Moreover, in the transition points u_ν between $I_{\nu-1}$ and I_ν we allow for a finite number of possible jumps of finite height.

Now we can state our announced representation of $\log f(u, \omega)$, as a function of $u \in [0, 1]$, with respect to the Haar wavelet basis $\{\varphi_{0,0}\} \cup \{\psi_{\ell,m}\}_{\ell \geq 0, m \geq 0}$ of $L_2([0, 1])$.

DEFINITION 4.5. The **evolutionary SLEX spectrum** $f = f^S$ at a frequency $\omega \in [-1/2, 1/2]$ is defined by the set of wavelet coefficients (i.e. the scaling coefficient $\beta_{-1,0}(\omega)$ and mother wavelet coefficients $\beta_{\ell,m}(\omega)$, $\ell = 0, 1, \ldots; m = 0, \ldots, 2^\ell - 1$) of its logarithmic wavelet expansion:

$$\log f(u,\omega) = \beta_{-1,0}(\omega)\varphi_{0,0}(u) + \sum_{\ell=0}^{\infty}\sum_{m=0}^{2^\ell-1} \beta_{\ell,m}(\omega)\psi_{\ell,m}(u), \tag{4.8}$$

where

$$\beta_{-1,0}(\omega) = \int_0^1 \log f(u,\omega)\varphi_{0,0}(u)du, \qquad \beta_{\ell,m}(\omega) = \int_0^1 \log f(u,\omega)\psi_{\ell,m}(u)du.$$

Note that, for ease of notation, we drop the superscript "S" from the SLEX spectrum $f^S(u,\omega)$. To prevent confusion with the definition of a Dahlhaus evolutionary spectrum $f^D(u,\omega)$, we keep that superscript "D" in the sequel, however.

Remarks. (1) Note that by Assumption 2 we impose a possibly very inhomogeneous regularity of $f(u,\omega)$ over $[0,1]$ (which will lead to an inhomogeneous segmentation basis B_T for the sequence of fixed T model spectra f_T; cf. our comment to Theorem 4.1 below). Moreover, it implies that the wavelet expansion (4.8) is well-defined, both pointwise in u for all points of continuity and, more generally, in $L_2[0,1]$. Approximation rates, i.e. rates of convergence of this (finite) wavelet series to its limit, will be derived by Theorem 4.1 below.

(2) If $f(u,\omega)$ is continuous everywhere in u, then Assumption 2 implies also that for the wavelet coefficients $\beta_{\ell m} = \beta_{\ell m}(\omega)$ of Definition 4.5, ω fixed, there exists a constant $C < \infty$ such that, with $s := \inf_\nu s_\nu$,

$$\sup_m |\beta_{\ell m}| \le C 2^{-(s+1/2)\ell}. \tag{4.9}$$

We would have liked to give our conditions directly by a decay (i.e. a summability) condition on the wavelet coefficients $\beta_{\ell,m}$. However, for Haar wavelets no necessary and sufficient condition exists to characterize the smoothness of the spectrum in u by the decay of its wavelet coefficients; we refer to Härdle *et al.* (1998), and to our discussion of the more regular case in Section 8.

(3) In Assumption 2 we allow our union of intervals I_ν, $I_\nu \in \mathcal{I}$ such that $\bigcup_\nu I_\nu = [0,1]$, to be a finite or an infinite one. In view of the wavelet expansion (4.8) we observe, in the finite case, the possibility of a very sparse modeling, e.g. to distinguish regions of piecewise constancy of $f(u,\omega)$ in time from those where the evolutionary spectrum is only known to be (Hölder-) smooth. Moreover, should the spectrum be discontinuous, only few wavelet coefficients will be affected by those jumps at dyadic positions, and "pointwise convergence" can still be considered to hold in these points of one-sided continuity. The case of a (locally) infinite union amounts to defining a point u of local Hölder continuity of order $s = \inf_\nu s_\nu$ which can be positive or zero (i.e. just continuity).

(4) Representing the spectrum f on the log-scale does not only guarantee that it is always positive. It also gives us the possibility to directly relate this expansion to our work ORvSM on the Auto-SLEX method where we used a well-defined quantitative criterion for the search of the best basis B_T which is actually based on the log-scale. The

use of the log-scale in criteria for selecting the best segmentation of a non-stationary time series was developed in Donoho *et al.* (1998), Section 14. We envision to further develop the theory of this selection algorithm by possibly using the model developed in our work here.

We now show that, for a given f, T, J_T (which define the underlying basis B_T), $\log f_T$ arises as the wavelet expansion of $\log f$ in time truncated to scale J_T. For this, let, in the sequel, denote $g_T(u,\omega) := \log f_T(u,\omega)$ and denote $g(u,\omega) := \log f(u,\omega)$. The following proposition would as well hold for $g_T(u,\omega) := f_T(u,\omega)$ and $g(u,\omega) := f(u,\omega)$ and for $g_T(u,\omega) := A_T(u,\omega)$ and $g(u,\omega) := A(u,\omega)$ (i.e. for the real and imaginary part of these complex functions). Note that due to the use of the Haar wavelets, both the log-spectrum and the transfer function actually imply the same definition of each $f_T(u,\omega)$, for given T and J_T.

PROPOSITION 4.1. *The wavelet expansion of* $g(u,\omega) = \log f(u,\omega)$ *truncated to scale* J_T *leads to Definition* 4.3, *i.e., for each fixed* T, J_T *and* $\omega_k = k/M_{J_T}$,

$$g_T(u,\omega_k) = \beta_{-1,0}(\omega_k)\varphi_{0,0}(u) + \sum_{\ell=0}^{J_T-1}\sum_{m=0}^{2^\ell-1}\beta_{\ell,m}(\omega_k)\psi_{\ell,m}(u) \tag{4.10}$$

$$= \sum_n \gamma_{J_T,n}(\omega_k)\varphi_{J_T,n}(u), \tag{4.11}$$

where

$$\beta_{\ell,m}(\omega) = \int_0^1 g(u,\omega)\psi_{\ell,m}(u)du, \tag{4.12}$$

and

$$\gamma_{J_T,n}(\omega) = \int_0^1 g(u,\omega)\varphi_{J_T,n}(u)du. \tag{4.13}$$

We show this proposition by the following remark giving an explicit intuitive explanation.

Remark. We show now the one-to-one relationship between the coefficients $\{\beta_{\ell,m}(\omega)\}$ and the set of $\tilde{\alpha}_{S,k,T} := \log(\alpha_{S,k,T})$ as appearing in Definition 4.3 of $\log f_T$. In fact, the set of $\tilde{\alpha}_{S,k,T}$ is defined by the following algorithm. (Note that in equation (4.10), a non-zero coefficient $\beta_{J-1,m}$ defines a non-zero SLEX coefficient, i.e. an $\tilde{\alpha}_{J,n} = \tilde{\alpha}_{S(J)}$ on the finest scale $J = J_T$):

Define, for each k, $\omega_k^j = k/M_j$ where $k = -M_j/2+1,\dots,M_j/2$ and for each $0 \le j \le J_T$,

$$\tilde{\alpha}^*_{j,n} := \tilde{\alpha}^*_{j,n}(\omega_k^j) = \int_0^1 g_T(u,\omega_k^j)2^{-j/2}\varphi_{j,n}(u)du. \tag{4.14}$$

Arrange the $\{\tilde{\alpha}^*_{j,n}\}_{j=0,\dots,J;n=0,\dots,2^j-1}$ in a usual dyadic tree (top-down, left-right), the coarsest resolution $j=0$ on top. Then proceed bottom-up as follows:

For $j = J : 1$

For $n = 0, 2, \dots, 2^j - 2$

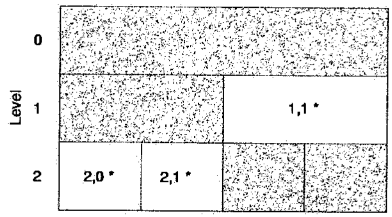

Fig. 2. Illustration of the algorithm. In this example, suppose that $J_T = 2$ and that $\tilde{\alpha}^*_{2,2} = \tilde{\alpha}^*_{2,3}$. The algorithm chooses $\tilde{\alpha}^*_{2,0}, \tilde{\alpha}^*_{2,1}$ and $\tilde{\alpha}^*_{1,1}$ to be the uniquely determined $\tilde{\alpha}$'s.

If $\tilde{\alpha}^*_{j,n} = \tilde{\alpha}^*_{j,n+1}$ then delete $\tilde{\alpha}^*_{j,n}$ and $\tilde{\alpha}^*_{j,n+1}$, else mark these $\tilde{\alpha}^*_{j,n}$ and $\tilde{\alpha}^*_{j,n+1}$.
END
END
Mark $\tilde{\alpha}^*_{0,0}$.
Define, for $j = 0, \ldots, J$, $n = 0, \ldots, 2^j - 1$, $\tilde{\alpha}_{j,n} = \tilde{\alpha}^*_{j,n}$ for those coefficients that are in terminal nodes of the tree, i.e. coefficients that are marked and whose descendants (at finer scales), if any, are unmarked. In other words, in each branch take the marked coefficients at the finest resolution. We refer to Fig. 2 for an illustrating example.

By this we get a unique covering of $[0, 1]$ by these specifically chosen $\tilde{\alpha}_{j,n}$'s which, for each of the fixed k, gives the set of $\tilde{\alpha}_{S,k,T}$ occuring in the definition of $g_T(u, \omega_k)$. In other words, for a given set of β's (and given T and J_T), this algorithm yields $\tilde{\alpha}$'s that are uniquely determined. Note that the transition from the β's to the γ's does only work for each finite T and J_T separately (and differently).

Conversely, for a given set $\{\tilde{\alpha}_{j,n} : \bigcup I_{j,n} \text{covers}[0, 1]\}$, i.e., for a given specific B_T, we define the following:

$$g_T(u, \omega_k) = \sum_{j,n} 2^{-j/2} \varphi_{j,n}(u) \tilde{\alpha}_{j,n}(\omega_k) \tag{4.15}$$

$$\beta_{\ell,m} = \int_0^1 g_T(u, \omega_k) \psi_{\ell,m}(u) du. \tag{4.16}$$

Thus, there is a one-to-one relationship between the set of $\tilde{\alpha}_{S,k,T}$ for the given T and the set of $\beta_{\ell,m}(\omega_k)$ for $\ell < J_T$, or the $\gamma_{J_T,n}(\omega_k)$, respectively. Under this model, the $\tilde{\alpha}_{S,k,T}$ evolve as $T \to \infty$ (so do the $\gamma_{J_T,n}$), but the $\beta_{\ell,m}$ are fixed.

It remains to discuss in more detail in which way our asymptotics of growing complexity of our model are to be understood, i.e. how f_T actually tends to f. In order that our model allows for asymptotic estimation theory (to be treated in Subsection 5.2) we assume that the number of points in each segment increases as T increases. But, of course, also the number of segments should be allowed to asymptotically increase if we want to model an evolutionary spectrum which, as a function of time, is not a finite linear combination of wavelets. This needs in particular to allow J_T to increase as T increases, but at a slower rate than $K_T = \log_2(T)$.

Assumption 3. (a) We assume that, as $K = K_T \to \infty$, either J_T is fixed or $J_T \to \infty$ such that $K_T - J_T \to \infty$ (i.e. $2^{J_T}/T \to 0$). Furthermore, $J_T \le J_{2T}$.

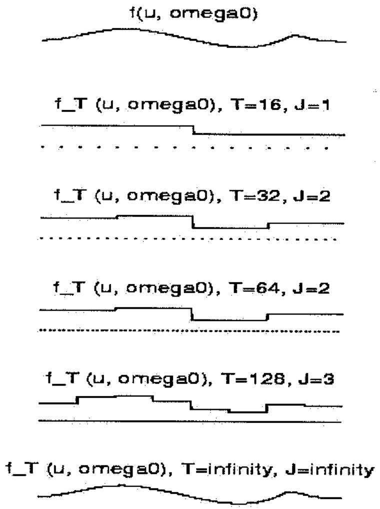

Fig. 3. Illustration of the asymptotic structure of the SLEX model. The top curve is a hypothetical true time-varying spectrum $f(u,\omega)$ plotted against rescaled time u at a fixed frequency ω. The next four curves show the piecewise constant approximation $f_T(u,\omega)$ for dyadic values of T and non-decreasing values of the finest level J_T. The dots indicate the location of the time points. Note that in going from $T=32$ to $T=64$, we do not increase J_T. This allows T to increase faster than 2^{J_T}, so that the number of time points M_J in a block at the finest level tends to infinity as T tends to infinity.

(b) For the length M_j of each segment $S(j)$, $j=0,\dots,J_T$, we assume that $M_{J_T}/M_j = c_j$ where $0 < c_j \leq 1$.

Remark. Assumption 3 implies that when $K \to K+1$ (that is, $T \to 2T$), either one of two events must happen: (A) $J_T \to J_{2T} = J_T + 1$ or (B) $J_T \to J_{2T} = J_T$. Event A means that each segment at the finest scale in the SLEX transform splits into two segments. Event B means that the finest scale in the SLEX transform does not split, so $f_{2T} = f_T$, but the number of points in each block at the finest scale doubles. Event A means that f_T is a linear combination of the elements of $\{\psi_{\ell,m} : \ell \leq J_T - 1\}$, while f_{2T} is the same linear combination of the elements of $\{\psi_{\ell,m} : \ell \leq J_T - 1\}$ **plus** a linear combination of the elements of $\{\psi_{\ell,m} : \ell = J_T\}$. Event B means that f_T and f_{2T} are the same linear combination of the elements of $\{\psi_{\ell,m} : \ell \leq J_T - 1\}$. If in Assumption 3, $J_T \to \infty$ such that $2^{J_T}/T \to 0$, events A and B each must occur an infinite number of times as $K_T \to \infty$. We refer to Fig. 3 for an illustrating example.

The following theorem justifies the asymptotic part of our SLEX model, i.e. it confirms that Definitions 4.3 and 4.5 are consistent with each other. Moreover, it gives rates of convergence of the finite sample spectrum to the evolutionary spectrum, for an optimally chosen J_T. This implies finally how to choose, for a given f and T, the "optimal" f_T (and hence the underlying basis B_T). Again we concentrate ourselves to treat the log-spectrum, only.

THEOREM 4.1. *Given the SLEX model as in Definition* 4.1, *with its SLEX spectrum f_T defined via Definition* 4.3 *and its evolutionary SLEX spectrum f as in Definition* 4.5 *fulfilling Assumptions* 1, 2 (*including the given $\mathcal{I}$*) *and* 3 *and a sequence of* *frequencies* $\omega_{k,T} \to \omega$, *we have the following convergences:*

(i) *Let* $u \in I_\nu = [u_\nu, u_{\nu+1}) \in \mathcal{I}$(*with* $\bigcup_\nu I_\nu = [0,1]$), *let* $2^{J_T} \sim T^{\mu/(\mu+s_\nu)}$ *as* $T \to \infty$. *Then*

$$|\log f_T(u, \omega_{k,T}) - \log f(u,\omega)| = O(T^{-s_\nu \mu/(\mu+s_\nu)}), \tag{4.17}$$

uniformly in $u \in I_\nu$.

(ii) *Let* $s := \inf_\nu s_\nu$ *and* $2^{J_T} \sim T^{\mu/(\mu+s)}$. *Then, as* $T \to \infty$,

$$\left| \int_0^1 du (\log f_T(u, \omega_{k,T}) - \log f(u,\omega))^2 \right| = O(T^{-(2s\mu)/(\mu+s)}). \tag{4.18}$$

As practical implication of this theorem, in view of Proposition 4.1 we observe in particular that, for each fixed T, our optimal segmentation B_T $(= \bigcup_\nu \bigcup_i S_i^\nu)$ will consists of segments $S_i^\nu \sim I_\nu$ living on the finest scale J_T determined by $2^{J_T} \sim T^{\mu/(\mu+s_\nu)}$. As different I_ν can have different Hölder exponents ν, this allows for a possibly very inhomogeneous segmentation. In addition the given optimal choice of J_T shows that the smoother the evolutionary spectrum is in time u (i.e. the faster its wavelet coefficients decay), the smaller has J_T to be chosen. On the contrary the smoother the spectrum in frequency the larger J_T should be chosen to benefit from a finer frequency resolution. Moreover, we add that the pointwise convergence of part (i) remains valid even in points of jumps as those are modeled to occur only at the dyadic change-points u_ν.

We finish this subsection by adding an example which relates back to our first example given in Remark (4) following Definition 4.1. We first identify the smoothness parameters in frequency and time for this moving average process with piecewise constant time behavior. The latter one gives rise to $s_1 = s_2 = 1$. As moving average specta are smooth in frequency, we need to set $\mu = 1$, i.e. again the maximal value in our considered Hölder class. Hence, the rate of approximation of Theorem 4.1 is, for each $u \in [0,1]$ of order $O(T^{-1/2})$ in equation (4.17). To add a second moving average example which is smoothly varing over time we choose $b(u)$ not to be piecewise constant but to be a Hölder function of order $s_1 < 1$ on the first block and $0 < s_2 < s_1$ on the second block. One possibility is to set, on the first half, $b(u) =$ const. $(u - 1/4)^{s_1}$ (which is Hölder s_1 in $u = 1/4$ (and differentiable elsewhere)) and on the second half we take $b(u) =$ const. $(u - 3/4)^{s_2}$. Then we observe that the best segmentation on the first half is given by the condition $2^{J_1} \sim T^{1/(1+s_1)}$ whereas on the second half by $2^{J_2} \sim T^{1/(1+s_2)}$. This gives a finer segmentation on the second half as the function $b(u)$ is not as smooth in time as it is on the first half.

4.3 *The complete model of a locally stationary SLEX process*

We complete this central section on building our model of a locally stationary SLEX process by a summarizing definition which is completely parallel to the one of a Dahlhaus locally stationary process as in Definition 2.1.

DEFINITION 4.6. A sequence of zero mean stochastic processes $\{X_{t,T}\}_{t=1,\dots,T}$ is called locally stationary SLEX process with evolutionary spectrum if for each $T \geq 1$ there exists a representation as in (4.1), i.e.,

$$X_{t,T} = \sum_{i:\bigcup S_i \sim B_T} \frac{1}{\sqrt{M_i}} \sum_{k_i=-M_i/2+1}^{M_i/2} \theta_{S_i,k_i,T} \overline{\phi_{S_i,\omega_{k_i}}(t)} z_{S_i,k_i} \tag{4.19}$$

with (finite-sample) SLEX spectrum as in (4.4), i.e., for $k_i = -M_i/2+1, \dots, M_i/2$,

$$f_T(u, \omega_{k_i}) = |\theta_{S_i,k_i,T}|^2 2^{-\ell_{I_i}/2} \varphi_{\ell_{I_i}, m_{I_i}}(u), \quad \text{for} \quad [uT] \in S_i, \tag{4.20}$$

and evolutionary SLEX spectrum $f(u,\omega)$ given by (4.8), i.e.,

$$\log f(u,\omega) = \lim_{T\to\infty} \log f_T(u,\omega) \tag{4.21}$$

$$= \lim_{T\to\infty} \left(\beta_{-1,0}(\omega)\varphi_{0,0}(u) + \sum_{\ell=0}^{J_T-1} \sum_{m=0}^{2^\ell-1} \beta_{\ell,m}(\omega)\psi_{\ell,m}(u) \right). \tag{4.22}$$

$f(u,\omega)$ is assumed to fulfill the Hölder regularity as given by Assumptions 1 and 2, and the limit $J_T \to \infty$ as $T \to \infty$ is given by Assumption 3 and Theorem 4.1.

5. Estimation theory

5.1 *Finite sample results*

We first state the results that hold for finite sample size T, i.e. under the SLEX model (4.1) for a given T and B_T. The results for the properties of the SLEX coefficients and SLEX periodogram parallel those for the classical Fourier coefficients and periodograms for the stationary time series.

THEOREM 5.1. *Given the SLEX model* (4.1) *of Definition* 4.1, *the SLEX coefficients $\hat\theta_{S,k,T}$, defined in* (3.3), *and the SLEX periodograms $\hat\alpha_{S,k,T} = |\hat\theta_{S,k,T}|^2$ enjoy the following finite sample properties*:

$$E(\hat\theta_{S,k,T}) = 0 \tag{5.1}$$

$$\text{Var}(\hat\theta_{S,k,T}) = E(\hat\alpha_{S,k,T}) = \alpha_{S,k,T} = f_T(u,\omega_k), \qquad [uT] \in S. \tag{5.2}$$

Theorem 5.1 shows that the SLEX "periodogram" is unbiased for the SLEX spectrum $f_T(u,\omega_k)$. Note the parallel that the classical (Fourier) periodogram is asymptotically unbiased for the spectrum for a stationary random process.

The following theorem uses Lemma 3.1 and assumes a Gaussian SLEX process.

THEOREM 5.2. *Under the SLEX model* (4.1) *with Gaussian increments $\{z_{S,k}\}$, the distribution of the SLEX coefficients $\hat\theta_{S,k,T}$ is complex normal, and the SLEX coefficients are independent.*

COROLLARY 5.1. *The periodograms $\hat\alpha_{S,k,T}$ of a Gaussian SLEX process are independent and distributed as*

$$\hat\alpha_{S,k,T} \sim \begin{cases} f_T(t/T,\omega_k)\chi_2^2/2, & \textit{if} \quad \omega_k \neq 0, 1/2 \\ f_T(t/T,\omega_k)\chi_1^2, & \textit{if} \quad \omega_k = 0, 1/2. \end{cases}$$

Note that, analogously to the classical periodogram, these properties remain valid asymptotically if the assumption of Gaussianity of the SLEX process is given up.

5.2 *Asymptotic theory*

As an immediate corollary to the approximation Theorem 4.1 and to Corollary 5.1 above, we get the *asymptotic* distribution of the SLEX periodograms for the general, i.e. *non-Gaussian* case.

COROLLARY 5.2. *The asymptotic distribution of* $\hat{\alpha}_{S,k,T}$ *is* $f(u,\omega_k)\chi_2^2/2$, *where* $[uT] = t_T \in S_T$ *for a sequence of segments* $S = S_T$.

This corollary implies that the SLEX "periodogram" is an inconsistent estimator of the SLEX "spectrum," as in the stationary case where the classical Fourier periodogram is inconsistent for the stationary spectrum. Furthermore, it suggests smoothing the periodogram across frequency to obtain a consistent estimator. One possibility to do so is to use a conventional kernel smoother in the frequency domain, i.e. for each $u \in (0,1)$, i.e. each time point $[uT] \in S_T = S_{u,T}$, we consider a kernel smoother of the SLEX periodograms $I_S(u,\omega_k) = \hat{\alpha}_{S,k,T}$ in this segment $S = S_T$ (not necessarily on the finest resolution level J_T) of length $M := |S_T| = T/2^{J_T}$:

$$\hat{f}(u,\omega) = (Mb_f)^{-1} \sum_{k=-M/2+1}^{M/2} K_f((\omega - \lambda_k)/b_f) I_S(u,\lambda_k), \tag{5.3}$$

where $\lambda_k = k/M$; $K_f(\vartheta)$ is the kernel function with compact support $[-1/2,1/2]$; $K_f(\vartheta) = K_f(-\vartheta)$ and $\int K_f(\vartheta)d\vartheta = 1$. Its bandwidth b_f is assumed to fulfill $b_f \to 0$ and $Mb_f \to \infty$ as $T \to \infty$. In this paper, we have used I also to mean an interval on $[0,1]$. We believe that the notation will be clear from the context of the discussion.

In fact this is the estimator that we considered in ORvSM, Theorem 5.1, where we proved consistency of this estimator for a locally stationary process following the Dahlhaus model. With the following theorem we complete our asymptotic theory which actually simplifies under the SLEX model. We note that this approach completely parallels using smoothed Fourier periodograms as an estimator of the spectrum of a stationary processes.

It is interesting to observe that in the resulting kernel estimator in Equation (5.3) we automatically include some sort of "smoothing" in rescaled time. We observe that in rescaled time u the shrinking length of the segment $|S_T|/T$ is proportional to $b_t = M_T/T = 2^{-J_T}$ which thus plays the role of a "bandwidth" in time.

For the following results on the asymptotics of this kernel estimator, the evolutionary SLEX spectrum is supposed to be smooth, i.e. Hölder of order s, on the whole interval $[0,1]$. Otherwise, a modification of the following result is straightforward in regarding the collection of intervals $\{I_\nu\}$ on which the spectrum is piecewise smooth.

THEOREM 5.3. *Suppose* $X_{t,T}$ *is a locally stationary SLEX process* (*Definition* 4.6) *with evolutionary SLEX spectrum fulfilling Assumptions* 1 *and* 2. *Then, uniformly in* u *and in* $k \neq 0$ *or* $\omega \neq 0$, *respectively, as* $J_T \to \infty$ *as in Assumption* 3 *the following properties hold*:

(i) $E\{I_S(u,\omega_k)\} = f(u,\omega_k) + O(b_t^s)$,

(ii) $E\{\hat{f}(u,\omega)\} = f(u,\omega) + O(b_t^s) + O(b_f^\mu) + O((Mb_f)^{-\mu})$,

(iii) $\text{var}\{\hat{f}(u,\omega)\} = (b_f b_t T)^{-1} f^2(u,\omega) \int K^2(\vartheta)d\vartheta + o((b_f b_t T)^{-1})$.

(iv) *For an optimal choice of* b_f *and* b_t (*i.e.* J_T), *the mean relative squared error of the kernel estimator* (5.3) *fulfills*

$$E(\hat{f}(u,\omega)/f(u,\omega)-1)^2 = O(T^{-2\mu s/(\mu+s+2\mu s)}).$$

Remarks. (1) The rate of consistency is $(b_f b_t T)^{-1/2} = (b_f M_T)^{-1/2}$ which is similar to the rate for smoothing a periodogram over M_T frequencies in a neighborhood of bandwidth b_f.

(2) The kernel estimator (5.3) has the same structure as a segmented periodogram-based statistic in the central limit theorem of Dahlhaus ((1997), Theorem A.2). In fact, the asymptotic considerations here are simpler, as due to our construnction of the SLEX model with spectra $f_T(u,\omega_k)$ being piecewise constant along blocks, the asymptotically vanishing bias due to non-stationarity is completely described by the behaviour of J_T, i.e. $b_t = 2^{-J_T}$, respectively. Hence with the same techniques applied on the higher cumulants of $\hat{f}(u,\omega)$, asymptotic normality of our kernel estimator can be shown.

(3) In fact, the rate of convergence of the mean relative squared error is the optimal one for nonparametric estimation of bivariate functions with different ("anisotropic") degree of regularity in its two arguments (time and frequency, here). This rate was, in more generality, developed in Neumann and von Sachs (1997), Theorem 2.1. As can be seen from the proof of part (iv) here, this optimal rate is achieved for the estimator (5.3) by choosing the bias in each direction to be equal, i.e. $b_t^s = b_f^\mu =: b$ and, as usual, balancing the leading parts of squared bias and variance of the estimator. The optimal bandwidth rates are thus derived by choosing $b \sim T^{-\mu s/(\mu+s+2\mu s)}$.

6. Relating the SLEX and Dahlhaus models of local stationarity

In this section, we establish an asymptotic mean square equivalence between our proposed SLEX model and the Dahlhaus (1997) model of a locally stationary process. The equivalence implies that the SLEX model is capable of modeling locally stationary processes that have smoothly time-varying spectrum despite being a sequence of piecewise stationary processes with spectra being piecewise constant in time.

We now give discuss our motivations for establishing the asymptotic equivalence between the SLEX model and the Dahlhaus model. First, our model allows for a simple and systematic method of estimation of the time-dependent second-order structure of the process. Thus, if one believes the observed data as a realization from the Dahlhaus model, one can use the SLEX model to fit the data and still obtain valid estimates of the time-varying spectrum, transfer function, or related quantities. Secondly, this equivalence relates our SLEX model to the popular autoregressive models with time-varying coefficients, which are a sub-class of the Dahlhaus model of locally stationary processes. This result is given in Section 4 of Dahlhaus (1997). Finally, and this is perhaps one of the biggest achievements of this new model, it allows for a straightforward synthesis of a locally stationary model. In addition, realizations from a Dahlhaus model can be easily generated by the SLEX model. Thus, in addition to inference based on asymptotic theory, the SLEX model allows for inference and model validation based on resampling. Compare our remarks on this capacity in Section 7. This particular interesting property is neither shared by the Dahlhaus model nor by any other known model of local stationarity (Nason *et al.* (2000), Donoho *et al.* (1998)).

As a preparation in showing the relationship we first establish a (spectral) representation of the variance of a SLEX process $\{X_{t,T}\}$ in terms of the evolutionary SLEX spectrum which, for this particular aim, is supposed to be smooth on the whole interval $[0,1]$ (otherwise, a modification of the following result is straightforward in regarding the collection of intervals on which the spectrum is piecewise smooth). For the following proposition to hold we need to make an additional assumption on the overlap η of the SLEX-transform. We assume that for each block S, $\epsilon = [\eta|S|] = O(1)$ as $|S|$ tends to infinity. Then we have the following asymptotic behavior of $\mathrm{var}(X_{t,T}) = |S|^{-1} \sum_k |\theta_{S,k,T}|^2 |\phi_{S,k}(t)|^2, t \in S$.

PROPOSITION 6.1. *For a locally stationary SLEX process $\{X_{t,T}\}$ with defining sequence $\{B_T\}_T$ (i.e. $\{f_T\}_T$) and evolutionary spectrum $f(u,\omega)$ fulfilling Assumptions* 1 *and* 2, *we have the following spectral representation of its variance*:

$$T^{-1} \sum_{t=1}^{T} \mathrm{var}(X_{t,T}) = \int_0^1 \int_{-1/2}^{1/2} f(u,\omega) d\omega\, du + O((2^{J_T}/T)^{\mu}) + O(2^{-sJ_T}).$$

6.1 *Discrete approximation of a Dahlhaus process*

A second result which will be used below is on the approximation of a frequency-continuous Dahlhaus process by its discrete approximation in frequency.

PROPOSITION 6.2. *To each sequence $\{X_{t,T}\}_{T \geq 1}$ of Dahlhaus processes with representation as in equation* (1.2), *with $A^0_{t,T}(\omega)$ being of bounded variation in ω over* $[-1/2, 1/2]$,

$$X_{t,T} = \int_{-1/2}^{1/2} A^0_{t,T}(\omega) \exp(i2\pi\omega t) dZ(\omega), \qquad t = 0, \ldots, T-1, \tag{6.1}$$

there exists, for a given set of discrete Fourier frequencies $\{\omega_k = k/N\}$, with an even integer $0 < N = N_T \leq T$, the following approximating process

$$X^{DD}_{t,T} = \sum_{k=-N/2}^{N/2} A^0_{t,T}(\omega_k) \exp(i2\pi\omega_k t) \xi^D_k, \qquad t = 0, \ldots, T-1, \tag{6.2}$$

where ξ^D_k are mean zero random variables with $\mathrm{Cov}(\xi^D_k, \xi^D_l) = N^{-1}\delta_{kl}$ *and* $\mathrm{Cov}(dZ(\omega), \xi^D_k) = \frac{\delta(\omega-\omega_k)}{N} d\omega$, *such that, as $N_T \to \infty$ with $T \to \infty$,*

$$E|X_{t,T} - X^{DD}_{t,T}|^2 = O(N^{-1}).$$

6.2 *Relationship of the SLEX model to the Dahlhaus model*

We formulate now our theorem on the (asymptotic) mean-square equivalence between a locally stationary SLEX process (Definition 4.6) and a Dahlhaus process (Definition 2.1). We restrict ourselves to treat the case of homogeneous Hölder regularity of order s on any given interval, w.l.o.g. assumed to be the complete $[0,1]$. Otherwise, for each subinterval in the partitioning of $[0,1]$ a possibly different Dahlhaus model (i.e. with different regularity) can be found similarly to the case treated here. In case of discontinuous transitions between the different SLEX spectra, the complete approximation would then be a piecewise smooth concatenation of Dahlhaus processes.

Note that here, by equation (6.4) below, we need to use an analogue of equation (4.22) to control the convergence of $A_T(u,\omega)$, defined by Definition 4.4, to the "evolutionary transfer function" $A(u,\omega)$. Compare our note preceeding Proposition 4.1 that the latter one does indeed imply the former one (i.e. the definition on the level of the (log) spectrum).

In the following theorem and discussion, we used S to denote both the blocks defined by the basis B_T and also the quantities that refer to the "SLEX" such as the SLEX transfer function ($\theta_{S,k,T}$) and the orthonormal increment process ($z_{S,k}$). We believe that the distinction between the two should be clear from the context.

THEOREM 6.1. (1) *Given a Dahlhaus process* $\{X_{t,T}^D\}_{T\geq 1}$ *as in Definition* 2.1, *with* $A_{t,T}^0(\omega)$ *being of bounded variation in* ω *over* $[-1/2,1/2]$, *and evolutionary spectrum* $f^D(u,\omega)=|A^D(u,\omega)|^2$ *fulfilling the regularity Assumptions* 1 *and* 2, *there exists a SLEX process* $\{X_{t,T}^S\}_{T\geq 1}$

$$X_{t,T}^S=\sum_{i:\bigcup S_i\sim B_T}\frac{1}{\sqrt{M_i}}\sum_{k_i=-M_i/2+1}^{M_i/2}\theta_{S_i,k_i,T}\overline{\phi_{S_i,\omega_{k_i}}}(t)z_{S_i,k_i} \tag{6.3}$$

with evolutionary transfer function

$$A^S(u,\omega)=\beta_{-1,0}(\omega)\varphi_{0,0}(u)+\sum_{\ell=0}^{\infty}\sum_{m=0}^{2^\ell-1}\beta_{\ell,m}(\omega)\psi_{\ell,m}(u), \tag{6.4}$$

given by $A^D(u,\omega)$, *where for each* T *and* J_T *the sequence* $\theta_{S,k,T}$ *is uniquely given by the set of coefficients* $\{\beta_{\ell m}\}_{j<J_T}$ *(analogously to the algorithm in the remark to Proposition* 4.1), *such that, with* $2^{J_T}\sim T^{s\mu/(\mu+s)}$, *as* $T\to\infty$,

$$T^{-1}\sum_{t=1}^{T}E|X_{t,T}^S-X_{t,T}^D|^2=O((2^{J_T}/T)^\mu)+O(2^{-sJ_T})=O(T^{-s\mu/(\mu+s)}).$$

The increment processes $dZ^D(\omega)$ *and* $z_{S,k}=dZ^S(k/|S|)$ *are such that*

$$\mathrm{Cov}(dZ^D(\omega),dZ^S(\lambda))=\delta(\omega-\lambda)d\omega.$$

(2) *Conversely, given a SLEX process* $\{X_{t,T}^S\}_{T\geq 1}$ *as in equation* (6.3) *with evolutionary transfer function* $A^S(u,\omega)$ *determined by* (6.4), *there exists a Dahlhaus process* $\{X_{t,T}^D\}_{T\geq 1}$ *with evolutionary transfer function* $A^D(u,\omega)$ *given by* $A^S(u,\omega)$ *and spectrum* $f^D(u,\omega)=|A^S(u,\omega)|^2$ *such that as* $T\to\infty$,

$$\sup_{t,\omega}|A_{t,T}^0-A^D(t/T,\omega)|=O(T^{-1}),$$

and, if $2^{J_T}\sim T^{s\mu/(\mu+s)}$, *as* $T\to\infty$,

$$T^{-1}\sum_{t=1}^{T}E|X_{t,T}^S-X_{t,T}^D|^2=O((2^{J_T}/T)^\mu)+O(2^{-sJ_T})=O(T^{-s\mu/(\mu+s)}).$$

The increment processes $dZ^D(\omega)$ *and* $z_{S,k}=dZ^S(k/|S|)$ *are related to each other as above in part* (1) *of the theorem.*

To give an example we like to relate back to our discussion after Definition 4.1 in Remarks (4) and (5) and our second example we gave at the end of Subsection 4.2, following Theorem 4.1. There we mentioned a sequence of transfer functions being piecewise constant in time and, as in the second example, converging to a (piecewise) smooth function in time. The proof of the above theorem shows that the two models differ only by the fact that, firstly, SLEX basis functions correspond to the complex exponentials of the Dahlhaus model, and secondly, that the SLEX coefficients $\theta_{S,k,T}$ of frequency ω_k in block S correspond to the average over time of the Dahlhaus "finite sample transfer function" $A^0_{t,T}(\omega_k)$ over this block S. Asymptotically both converge to the same smooth object in time, i.e. $A^D(u,\omega) = A^S(u,\omega) = \sigma(1 + a(u)\exp(i\omega))$, i.e. the transfer function of a time-varying moving average process of order 1.

7. Inference in the SLEX model based on the bootstrap

The SLEX model inherently includes an elegant and conceptually simple way of deriving statistical inference on the SLEX spectrum based on *bootstrapping*. Unlike other models on local stationarity, the SLEX model includes the notion of a spectrum $f_T(u,\omega_k)$ for each fixed T, and implicitly (by its transfer function $\theta_{S,k,T}$) this shows up as building part of the model-generating process. This automatically opens a powerful possibility to resample from our fixed sample size model (4.1) i.e., with slight abuse of notation, given estimators $\hat{B}_T = \bigcup_i \widehat{S}_i$ of $B_T = \bigcup_i S_i$ and $\{\hat{\theta}_{\widehat{S}_i,\widehat{k}_i,T}\}$ of the $\{\theta_{S_i,k_i,T}\}$,

$$X^*_{t,T} = \sum_{i:\bigcup \widehat{S}_i \sim \hat{B}_T} \frac{1}{\sqrt{|\widehat{S}_i|}} \sum_{\widehat{k}_i} \hat{\theta}_{\widehat{S}_i,\widehat{k}_i,T} \overline{\phi_{\widehat{S}_i,\omega_{\widehat{k}_i}}(t)} z^*_{\widehat{S}_i,\widehat{k}_i}, \tag{7.1}$$

by drawing independent copies of the $\{z^*_{\widehat{S}_i,\widehat{k}_i}\}$ and generating bootstrap realizations X^*_t.

In Ombao, von Sachs and Guo [OvSG] (2000), we describe a SLEX bootstrap procedure which is based on Franke and Härdle [FH] (1992). In FH, they defined a "residual" in the frequency domain as the ratio of the raw Fourier periodograms to the smoothed periodograms. In our approach, we generalize the FH approach by defining residuals in each stationary block. We define a residual in each block to be the ratio of the raw empirical SLEX coefficients (rather than periodograms as in FH) to the smoothed SLEX periodograms. The residuals are then studentized to obtain sample analogues of the random z's in equation (4.1). We sample from these residuals to obtain the z^*'s that will be used in equation (7.1) to synthezise a bootstrap time series. As the segmentation of the time series is estimated from the data, we must account for the uncertainty due to the segmentation in addition to the usual sampling variability. In order to account for this extra variability, we apply the Auto-SLEX method to the bootstrap time series $X^*_{t,T}$ to obtain a segmentation for $X^*_{t,T}$.

In addition to the nonparametric approach, we also developed a parametric alternative where we draw the independent z^*'s from a Gaussian distribution. This procedure is justified based on the following argument. The residuals are studentized versions of the empirical SLEX coefficients. The SLEX coefficients, just like the Fourier coefficients, are asymptotically Gaussian. We do not provide a proof for the asymptotic normality of the SLEX coefficients but the proof will follow that of the Fourier coefficients in Brillinger (1981). Our parametric bootstrap procedure is parallel to FH where they draw the bootstrap residuals from a chi-square distribution. The FH residuals are based on the Fourier periodograms which have asymptotic chi-squared distributions.

It is natural that in forming a confidence interval (CI) for the SLEX spectrum $f_T(u,\omega)$ (which inherently determines the underlying model segmentation B_T), to replace CI's based on the asymptotic normality of our kernel spectral estimator $\hat{f}_T(u,\omega)$ as in equation (5.3), by those from a bootstrap procedure. The bootstrap based CI's are more accurate, particularly for small sample sizes, than the CI's based on asymptotic normality (as a consequence of the higher accuracy of the bootstrap distribution shown in a more general set-up by Dahlhaus and Janas (1996)). Note that in practice one would use $|\hat{\theta}_{\widehat{S}_i,\widehat{k}_i,T}| = \{\hat{f}_T(u,\omega_{\widehat{k}_i})\}^{1/2}$, where $[uT] \in \widehat{S}_i$, as estimates of the time-varying transfer function. Moreover, similarly to FH, we only treat pointwise CI's since it is known from stationary spectral analysis that simultaneous CI's (i.e. confidence bands) are generally too wide so they are not useful for practical purposes.

8. Discussion and conclusion

In this paper, we proposed the SLEX model of a non-stationary random process. The SLEX model uses the SLEX basis vectors, which are simultaneously orthogonal and localized in time and frequency, as stochastic building blocks. We defined the SLEX model for a fixed sample size T and its associated spectrum f_T which can be estimated by the smoothed SLEX periodograms. We also discussed inference on the SLEX model which is based on the bootstrap. Resampling is possible under the SLEX model because the special structure of the SLEX model for each T allows for an accessible synthesis equation. This particular property is unique to the SLEX model as it is not shared by the other models of local stationarity.

To allow for a growing complexity of the model as T tends to infinity, we defined the "limit" of f_T which is the smooth evolutionary spectrum f. This f, as a function of time, has a representation in terms of Haar wavelets which allows us to develop rates of convergence of the piecewise constant f_T to the smooth f. Consistency of smoothed SLEX periodograms has been shown by reference to this uniquely defined estimand f and (optimal) rates of convergence have been developed which depend on the smoothness of f. As a further implication, an asymptotic mean square equivalence between the Dahlhaus model and the SLEX model arises. As a consequence, the SLEX model is also equivalent to the class of time-varying AR processes which is a sub-class of the Dahlhaus model. Moreover, one can generate a realization from the Dahlhaus model in a straightforward and efficient manner by using the synthesis equation of the SLEX model.

We mention the possibility to enlarge our considered model class by including more regular evolutionary SLEX spectra. For this we need to use a representation of the spectrum in time u with respect to more regular wavelets, e.g. members of the more regular Daubechies wavelets with compact support. In this case the regularity of the SLEX spectrum can directly be characterized by the decay of the wavelet coefficients; we refer to the general theory of function classes such as Sobolev or Besov spaces (see Härdle *et al.* (1998)). This would on the other hand also allow to include functions with jumps at arbitrary locations, for which instead of a pointwise treatment (in time u) we needed to pass to a treatment in $L_2([0,1])$ (cf. our result (4.18) in Theorem 4.1). We decided, however, not to include these generalizations in our paper to prevent the presentation of our main ideas from being obscured by unnecessary technical discussions. In particular we would need to stick to the log-spectrum definition (4.5) (which would no more be equivalent to (4.4)) to continue to ensure that the spectrum itself still remains positive. Although conceptually not a real problem, this would cause various extra constructions to derive the same set of assertions.

As for future work we plan to extend the model presented in this paper to the multivariate situation. The SLEX transform is complex-valued, and hence the SLEX model will allow in a natural way to model the time-lag and inter-relationships between components of multivariate time series. Moreover, in the multivariate model, we can define the time-dependent spectrum of each component and evolutionary cross spectrum and coherence between any pairs of time series.

In addition, we are also working on formulating a solution to the problem of over-all consistency of the Auto-SLEX method (ORvSM), using the SLEX model. To comment briefly, the Auto-SLEX method is based on the SLEX periodograms calculated over all possible segmentations, and it uses the Best Basis Algorithm of Coifman and Wickerhauser (1992), a data-driven procedure, to select the best adapted segmentation of the given time series. Thus, it is necessary to study the bias that the Auto-SLEX method suffers when an incorrect segmentation is selected. We believe that consistency might be more natural to show in the context of the SLEX model rather than in, e.g., the Dahlhaus model. This is due, in part, to the fact that both the SLEX model and the Auto-SLEX method use the same basic elements, i.e. the SLEX basis vectors and (squared) coefficients (i.e. periodograms) for which a coherent asymptotic theory has now been established by this paper.

Appendix: Proofs

PROOF OF THEOREM 4.1. Let $J = J_T$. W.l.o.g. we only treat the case of $I_\nu = I_1 = [0,1]$, with $s := s_1$. We actually need to consider a sequence of grid frequencies $\omega_{k,T} = \omega_{k(T)} = k(T)/|S(J)|$ which tends to the fixed ω as $T \to \infty$, such that $|\omega - \omega_{k,T}| = O(|S(J)|^{-1}) = O(2^J/T)$. (Compare also Brillinger (1981), Corollary 5.4.1, p. 135.)

Part (i): As our Assumption 2 implies continuity of the (log-) spectrum as a function in u, hence we have pointwise convergence. More explicitly, with $g_T(u,\omega) := \log f_T(u,\omega)$ and $g(u,\omega) := \log f(u,\omega)$,

$$|g_T(u,\omega_{k,T}) - g(u,\omega)| \le |g_T(u,\omega_{k,T}) - g(u,\omega_{k,T})| + |g(u,\omega_{k,T}) - g(u,\omega)|.$$

For the first term, we use the decay of the wavelet coefficents, i.e. we use the Hölder continuity of Assumption 2 and a result of Härdle *et al.* ((1998), Theorem 9.4) which gives us a rate of convergence of order $O(2^{-sJ_T})$. For the second term, we use the Lipschitz-condition (4.7) to bound

$$|g(u,\omega_{k,T}) - g(u,\omega)| \le L|\omega_{k,T} - \omega|^\mu$$

to be of order $O((2^{J_T}/T)^\mu)$, uniformly in u, as $|\omega - \omega_{k,T}| = O(2^J/T)$. To get a final order for the pointwise convergence we need to balance the two terms which gives us that $2^{J_T} \sim T^{\mu/(\mu+s)}$. This leads to an overall rate of $O(T^{-s\mu/(\mu+s)})$.

Part (ii): For the convergence in $L_2[0,1]$ we treat the second term, using the Lipschitz-condition (4.7), similarly to the proof of part (i), i.e.

$$\int_0^1 du(g(u,\omega_{k,T}) - g(u,\omega))^2 = O((2^{J_T}/T)^{2\mu}).$$

For the first part, we use condition (4.9), in the following Parseval relation, for fixed frequency ω.

$$\int_0^1 du(g_T(u,\omega_{k,T}) - g(u,\omega_{k,T}))^2 = \sum_{\ell \geq J}\sum_m |\beta_{\ell,m}(\omega_{k,T})|^2 = O(2^{-2sJ}),$$

as $\sup_m |\beta_{\ell,m}(\omega_{k,T})|^2 = O(2^{-(2s+1)\ell})$. Balancing the two rates gives the same condition for the optimal choice of J_T as before as $2^{J_T} \sim T^{2\mu/(2\mu+2s)}$ which leads to an overall rate of convergence in $L_2[0,1]$ of order $O(T^{-2s\mu/(\mu+s)})$ being quite naturally the square of the pointwise order above.

Proof of Theorem 5.1. By the linearity of the SLEX transform,

$$E(\hat{\theta}_{S,k,T}) = \frac{1}{\sqrt{|S|}}\sum_{t=0}^{T-1}\phi_{S,k,T}(t)E(X_t) = 0.$$

Furthermore,

$$\mathrm{Var}(\hat{\theta}_{I,k,T}) = E|\hat{\theta}_{I,k,T}|^2 = |\theta_{I,k,T}|^2.$$

This result follows directly from

$$\sum_t \phi_{S,k,T}(t)\overline{\phi_{S',k',T}(t)} = |S|\delta(S-S')\delta(k-k')$$

and from $E[dZ(\omega)\overline{dZ(\omega')}] = \delta(\omega-\omega')$.

Proof of Theorem 5.2. We need to show that the real and imaginary parts of $\hat{\theta}_{S,k,T}$ are (i) each normally distributed; (ii) each independent; and (iii) have the same variance which is equal to half the variance of $\hat{\theta}_{S,k,T}$.

The real and imaginary parts of $\hat{\theta}_{S,k,T}$ are, $\mathrm{Re}(\hat{\theta}_{S,k,T}) = (\hat{\theta}_{S,k,T} + \overline{\hat{\theta}_{S,k,T}})/2$ and $\mathrm{Im}(\hat{\theta}_{S,k,T}) = (\hat{\theta}_{S,k,T} - \overline{\hat{\theta}_{S,k,T}})/(2i)$, respectively. Their normality follows directly from the linearity of the SLEX transform. Lemma 3.1 can be used to prove that the variance of the real and imaginary parts are each $(1/2)|\theta_{S,k,T}|^2$. This is a straightforward consequence of Theorem 5.1 and of the fact that $E(\hat{\theta}_{S,k,T}\hat{\theta}_{S,k,T}) = 0$ as well as $E(\overline{\hat{\theta}_{S,k,T}\hat{\theta}_{S,k,T}}) = 0$.

To show the independence of the real and imaginary parts, it is sufficient to show that they are uncorrelated because they are jointly normally distributed. Note that

$$\begin{aligned} E(\mathrm{Re}(\hat{\theta}_{S,k,T})\mathrm{Im}(\hat{\theta}_{S,k,T})) &= \frac{-1}{4i}E([\hat{\theta}_{S,k,T}\overline{\hat{\theta}_{S,k,T}} - \overline{\hat{\theta}_{S,k,T}}\hat{\theta}_{S,k,T}] \\ &\quad + [\overline{\hat{\theta}_{S,k,T}\hat{\theta}_{S,k,T}} - \hat{\theta}_{S,k,T}\hat{\theta}_{S,k,T}]) = 0. \end{aligned}$$

The independence of the SLEX coefficients follows directly from the orthogonality of the SLEX basis vectors in Lemma 3.1.

Proof of Corollary 5.1. As a consequence of Theorem 5.2 where the SLEX coefficients are shown to be independent, it follows directly that the SLEX periodograms are also independent. On the distribution of the SLEX periodograms: By the independence and normality of the real and imaginary parts of the SLEX coefficients, it follows that $\hat{\alpha}_{S,k,T}/f(u,\omega_k)$ is distributed as independent $\chi^2_2/2$ random variable for

$k = 1, \ldots, M_j/2 - 1$, where $u = t/T$ and $[uT] = t_T \in S_T$ for a sequence of segments S_T. When $k = 0, M_j/2$, the SLEX periodogram $\hat{\alpha}_{S,k,T}/f(u, \omega_k)$ is distributed as a χ_1^2 random variable.

PROOF OF THEOREM 5.3. Part (i) is just a preparation for the followings parts and follows immediately from equation (5.2) in Theorem 5.1 and by the Hölder continuity of order s of the SLEX spectrum in u. Compare also Theorem 4.1.

For part (ii) we use this result in observing that

$$\begin{aligned} E\{\hat{f}(u,\omega)\} &= (Mb_f)^{-1} \sum_{k=-M/2+1}^{M/2} K_f((\omega - \lambda_k)/b_f) f(u, \lambda_k) + O(b_t^s) \\ &= b_f^{-1} \int K((\omega - \lambda)/b_f) f(u, \lambda) d\lambda + O((Mb_f)^{-\mu}) + O(b_t^s) \\ &= f(u, \omega) + O(b_f^\mu) + O((Mb_f)^{-\mu}) + O(b_t^s), \end{aligned}$$

due to the Hölder continuity of order μ of the SLEX spectrum in frequency.

We deliver the proof of part (iii) only for Gaussian SLEX processes, to keep this proof simple. Using the independence of the SLEX periodograms (shown in Corollary 5.1)

$$\begin{aligned} \mathrm{var}\{\hat{f}(u,\omega)\} &= (b_f b_t T)^{-2} \sum_{k=-M/2+1}^{M/2} K_f^2((\omega - \lambda_k)/b_f) f_T^2(u, \lambda_k) \\ &= (b_f b_t T)^{-1} f^2(u,\omega) \int K^2(\vartheta) d\vartheta + O((b_f M)^{-1}(b_t^s + b_f^\mu + (b_f M)^{-\mu})). \end{aligned}$$

For the last part (iv) we set $b_t^s = b_f^\mu =: b$ and, as usual, balance the leading parts of squared bias and variance of the estimator, i.e. setting $(Tb^{1/s}b^{1/\mu})^{-1} = b^2$. This leads to choosing $b \sim T^{-\mu s/(\mu+s+2\mu s)}$, which results into a final rate of convergence of the mean relative squared error of $T^{-2\mu s/(\mu+s+2\mu s)}$. Note that this is the best possible rate for anisotropic functions in a bivariate Hölder class of different degrees of regularity μ and s, as was shown by Theorem 2.1 of Neumann and von Sachs (1997) (for Sobolev classes, strictly speaking) and by Neumann (2000) for more general Besov classes, including Hölder. An estimator which achieves this rate is actually the one with $b_t^s = b_f^\mu$.

PROOF OF PROPOSITION 6.1.

$$\begin{aligned} T^{-1} \sum_{t=1}^{T} \mathrm{var}(X_{t,T}) &= T^{-1} \sum_{t=1}^{T} \sum_{i:\cup_i S_i = B_T} |S_i|^{-1} \sum_{k_i} |\theta_{S_i,k_i,T}|^2 |\phi_{S_i,k_i}(t)|^2 \\ &= T^{-1} \sum_{i:\cup_i S_i = B_T} |S_i|^{-1} \sum_{k_i} f_T(u_i, k_i) \sum_{t \in S_i} |\phi_{S_i,k_i}(t)|^2, \end{aligned}$$

where the u_i are such that $[u_i T] = t_i \in S_i$.

As $\epsilon_i = [\eta |S_i|] = O(1)$ for all i, $\sum_{t \in S_I} |\phi_{S_i,k_i}(t)|^2 = \sum_{t \in \tilde{S}_i} |\phi_{S_i,k_i}(t)|^2 + O(\epsilon_i)$, where $\tilde{S}_i$ denotes the support of the SLEX basis function $\phi_{S_i,k_i}(t)$, over which it is orthonormal: $\sum_{t \in \tilde{S}_i} |\phi_{S_i,k_i}(t)|^2 = |S_i|$. Note that $S_i \subset \tilde{S}_i$ and that 2ϵ is the length of $\tilde{S}_i \setminus S_i$. Hence, the above is equal to

$$= T^{-1} \sum_{i:\cup_i S_i = B_T} |S_i|^{-1} \sum_{k_i} f_T(u_i, \omega_{k_i})(|S_i| + O(1)).$$

We continue with the leading term, noting that as $|S_i|^{-1}\sum_{k_i} f_T(u_i,\omega_{k_i}) = O(1)$ and there are at most 2^{J_T} terms in the first sum over i such that this first remainder is at most of order $O(2^{J_T}/T)$.

Using part of the results of Theorem 4.1, we do the following replacement

$$f_T(u_i,\omega_{k_i}) = f(u_i,\omega_{k_i}) + O(2^{-sJ_T}),$$

getting an additional remainder of this order $O(2^{-sJ_T})$, uniformly in ω and u. Furthermore, using that

$$|S_i|^{-1}\sum_k f(u_i,k_i/|S_i|) = \int_{-1/2}^{1/2} f(u_i,\omega)d\omega + O(|S_i|^{-\mu}),$$

as f is Hölder of order μ in ω, we get an additional remainder of at most order $O((2^{J_T}/T)^{\mu})$. The now leading term is

$$\sum_{i:\cup_i S_i = B_T} |S_i|/T \int_{-1/2}^{1/2} f(u_i,\omega)d\omega.$$

Now, replacing the sum over i by the integral over $u \in [0,1]$, this equals

$$\int_0^1 \int_{-1/2}^{1/2} f(u,\omega)d\omega\, du + O\left(\left(\sup_i |S_i|/T\right)^s\right),$$

as $f(u,\omega)$ is Hölder of order s in u. We finish this proof by noting that this last remainder is of order $O(2^{-sJ_T})$, as by Assumption 3 we imposed that the length of each segment asymptotically grows as $T/2^{J_T}$.

Proof of Proposition 6.2. By construction, the processes $X_{t,T}$ and $X_{t,T}^{DD}$ each have mean zero. Thus,

$$E|X_{t,T} - X_{t,T}^{DD}|^2 = \operatorname{var}(X_{t,T}) + \operatorname{var}(X_{t,T}^{DD}) - 2\operatorname{Cov}(X_{t,T}, X_{t,T}^{DD}).$$

We have, by approximation of the sum over Fourier frequencies by the corresponding integral for the function $A_{t,T}^0(\omega)$ of bounded variation in ω, that

$$\operatorname{var}(X_{t,T}) = \int |A_{t,T}^0(\omega)|^2 d\omega, \tag{9.1}$$

$$\operatorname{var}(X_{t,T}^{DD}) = N^{-1}\sum_k |A_{t,T}^0(\omega_k)|^2 = \int |A_{t,T}^0(\omega)|^2 d\omega + O(N^{-1}), \tag{9.2}$$

$$\operatorname{Cov}(X_{t,T}, X_{t,T}^{DD}) = \operatorname{var}(X_{t,T}^{DD}), \tag{9.3}$$

by construction of the sequence ξ_k^D.

Proof of Theorem 6.1. We begin by summarizing the proof of part (2): Starting from a SLEX process $\{X_{t,T}^S\}_{T\geq 1}$ with given $A^S(u,\omega)$ we construct our approximating Dahlhaus process as follows: Define the sequence $A_{t,T}^0(\omega) := A^S(t/T,\omega)$; and, for $t \in S_i$, define $\tilde{A}^i_{t,T}(\omega_{k_i}) := \theta_{S_i,k_i,T}$ (hence being constant in time over the block S_i). As our SLEX process model is based on Haar wavelets this is actually equal to saying that $\tilde{A}^i_{t,T}(\omega_{k_i}) = |S_i|^{-1} \sum_{t\in S_i} A^0_{t,T}(\omega_{k_i})$. Note that $|A^0_{t,T}(\omega_{k_i}) - \tilde{A}^i_{t,T}(\omega_{k_i})| = O(|S_i|^{-s})$ uniformly over k_i and $t \in S_i$.

As for part (1), starting from a Dahlhaus process $\{X^D_{t,T}\}_{T\geq 1}$ with $A^D(u,\omega)$ we retrieve, via its wavelet expansion (as in equation (4.8), but now for the real and imaginary part of $A^D(u,\omega)$, assuming of course the same regularity for both), for given T and with the particular choice of $2^{J_T} \sim T^{s\mu/(\mu+s)}$, the $\theta_{S_i,k_i,T}$ for the blocks $S_i \sim B_T$, i.e. $\cup_i S_i = B_T$. In fact, as above, $\theta_{S_i,k_i,T} = \tilde{A}^i_{t,T}(\omega_{k_i})$, for $t \in S_i$.

Then, for both parts, the main idea is, as in the proof of Proposition 6.1, to decompose

$$E|X^S_{t,T} - X^D_{t,T}|^2 = \mathrm{var}(X^S_{t,T}) + \mathrm{var}(X^D_{t,T}) - 2\,\mathrm{Cov}(X^S_{t,T}, X^D_{t,T}),$$

and to show that, summing up over t, each term of the r.h.s. tends to the same quantity as arising as the limit of Proposition 6.1. I.e., we use Proposition 6.1 to derive the asymptotic expression for $\mathrm{var}(X^S_{t,T})$ which will be the same as those for $\mathrm{var}(X^D_{t,T})$. The same principle holds for the covariance using that $\theta_{S_i,k_i,T}\overline{A^0_{t,T}}(\omega_{k_i})$ will also tend to $f^D(u_i, \omega_{k_i})$ with $t_i = [Tu_i] \in S_i$. For the last we will quite naturally need, as intermediate, a set of discretized Dahlhaus processes with a frequency resolution of $|S_i|^{-1}$ on block S_i.

We deliver now the details of this proof only for part (2). In virtue of Proposition 6.2, it is sufficient to study the difference between the SLEX and a discretized Dahlhaus process $\{X^{DD}_{t,T}\}$. Hereby we use for the different blocks S_i arising in the segmentation B_T of the given SLEX process $X^S_{t,T}$ different discretized Dahlhaus processes $\{X^i_{t,T}\}$, with frequency discretization $\{\omega_{k_i} = k_i/|S_i|\}_{k_i=-|S_i|/2,\ldots,|S_i|/2}$. Note however that again $|S_i| = |S_{J_T}| = T/2^{J_T}$ for all i.

$$T^{-1}\sum_{t=1}^{T} E|X^S_{t,T} - X^D_{t,T}|^2 \leq T^{-1} \sum_{i:\cup_i S_i = B_T}\ \sum_{t\in S_i} 2\left(E|X^S_{t,T} - X^i_{t,T}|^2 + E|X^i_{t,T} - X^D_{t,T}|^2\right).$$

By Proposition 6.2 each term in the second sum is of order $O(2^{J_T}/T)$. As the sums over i and t count T terms altogether, the second term is again bounded by $2^{J_T}/T$.

For the first terms we use again that, for each i,

$$E|X^S_{t,T} - X^i_{t,T}|^2 = \mathrm{var}(X^S_{t,T}) + \mathrm{var}(X^i_{t,T}) - 2\,\mathrm{Cov}(X^S_{t,T}, X^i_{t,T}).$$

Mimicking the idea of the proof of Proposition 6.1 to treat $T^{-1}\sum_{t=1}^T \mathrm{var}(X^i_{t,T})$ we get that the two first terms tend to $\int_0^1\int_{-1/2}^{1/2} f(u,\omega)d\omega\, du$, where quite naturally, $f(u,\omega) = |A^S(u,\omega)|^2$. The remainders are, for the first term, of order $O((2^{J_T}/T)^\mu) + O(2^{-sJ_T})$ (as in Proposition 6.1) and, for the second term, of order $O(2^{J_T}/T) + O(T^{-s})$ which can be seen as follows. (Note that for this term we can re-combine the splitted sum over i and t into one sum over $t = 1,\ldots,T$.) Starting from equation (9.2), with $N = |S_i| = T/2^{J_T}$, i.e.

$$\mathrm{var}(X^i_{t,T}) = N^{-1}\sum_k |A^0_{t,T}(\omega_{k_i})|^2 = \int |A^0_{t,T}(\omega)|^2 d\omega + O(N^{-1}),$$

we continue approximating

$$\int |A^0_{t,T}(\omega)|^2 d\omega = \int |A(t/T,\omega)|^2 d\omega + O(T^{-1}),$$

and

$$T^{-1}\sum_{t=1}^{T}\int|A(t/T,\omega)|^2 d\omega=\int_0^1\int_{-1/2}^{1/2}|A(u,\omega)|^2 d\omega du+O(T^{-s}),$$

as $f(u,\omega)$ is Hölder of order s in time u.

It remains to show that $\mathrm{Cov}(X_{t,T}^S, X_{t,T}^{DD})$ tends to the same expression as the variance terms (with at least the order of the first variance term). Using again the principle of the proof of Proposition 6.1, the summed covariance is equal to

$$\sum_{i:\cup_i S_i=B_T}\sum_{t\in S_i}|S_i|^{-1}\sum_{k_i=-|S_i|/2}^{|S_i|/2}\theta_{S_i,k_i,T}\phi_{S_i,k_i}(t)\overline{A_{t,T}^0}(\omega_{k_i})\exp(-i\omega_{k_i}t),$$

having used that

$$\mathrm{Cov}(|S_i|^{-1/2}dZ^S(\ell/|S_i|),\xi_{k_i}^i)=\delta_{\ell k_i}|S_i|^{-1}.$$

As $|A_{t,T}^0(\omega_{k_i})-\tilde{A}_{t,T}^i(\omega_{k_i})|=O(|S_i|^{-s})$ uniformly over k_i and $t\in S_i$, we finish the proof of the covariance part analagously to the one of the variance term of $X_{t,T}^S$, with a leading term

$$\sum_{i:\cup_i S_i=B_T}|S_i|^{-1}\sum_{t\in S_i}\sum_{k_i=-|S_i|/2}^{|S_i|/2}\theta_{S_i,k_i,T}\overline{\tilde{A}_{t,T}^i}(\omega_{k_i})\phi_{S_i,k_i}(t)\exp(-i\omega_{k_i}t).$$

As $\tilde{A}_{t,T}^i(\omega_{k_i})=\theta_{S_i,k_i,T}$, for, $t\in S_i$, this leading term converges to

$$\sum_{i:\cup_i S_i=B_T}\sum_{k_i=-|S_i|/2}^{|S_i|/2}\theta_{S_i,k_i,T}\overline{\theta_{S_i,k_i,T}},$$

analogously to the proof of Proposition 6.1. For this last step we have used that

$$\left|\left|S_i\right|^{-1}\sum_{t\in S_i}\phi_{S_i,k_i}(t)\exp(-i\omega_{k_i}t)-1\right|\le 4\epsilon_i/\left|S_i\right|,$$

as by construction of the SLEX vectors (see Subsection 3.2) those differ from the complex exponentials only along a region of $\pm\epsilon$ at each of the two end points of the block S.

References

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*, McGraw-Hill, New York.

Coifman, R. and Wickerhauser, M. (1992). Entropy based algorithms for best basis selection, *IEEE Trans. Inform. Theory*, **32**, 712–718

Dahlhaus, R. (1997). Fitting time series models to nonstationary processes, *Ann. Statist.*, **25**, 1–37.

Dahlhaus, R. and Janas, D. (1996). A frequency domain bootstrap for ratio statistics in time series analysis, *Ann. Statist.*, **24**, 1934–1963.

Daubechies, I. (1992). *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

Donoho, D., Mallat, S. and von Sachs, R. (1998). Estimating covariances of locally stationary processes: Rates of convergence of best basis methods, Tech. Report, No. 517, Statistics Department, Stanford University, California.

Franke, J. and Härdle, W. (1992). On bootstrapping kernel spectral estimates, *Ann. Statist.*, **20**, 121–145.

Härdle, W., Kerkyacharian, G., Picard, D. and Tsybakov, A. (1998). *Wavelets, Approximation, and Statistical Applications*, Lecture Notes in Statist., No. 129, Springer, New York.

Nason, G., von Sachs, R. and Kroisandt, G. (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum, *J. Roy. Statist. Soc. Ser. B*, **62**, 271–292.

Neumann, M. H. (2000). Multivariate wavelet thresholding in anisotropic function spaces, *Statist. Sinica*, **10**, 399–432.

Neumann, M. H. and von Sachs, R. (1997). Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra, *Ann. Statist.*, **25**, 38–76.

Ombao, H. (1999). Statistical analysis of non-stationary time series, Ph.D. dissertation, Department of Biostatistics, University of Michigan.

Ombao, H., von Sachs, R. and Guo, W. (2000). Estimation and inference for time-varying spectra of locally stationary SLEX processes, *Discussion Paper, No. 00/27, Institut de Statistique, UCL, Louvain-la-Neuve* and *Proceedings of the 2nd International Symposium on Frontiers of Time Series Modeling, Nara, Japan, December 14–17, 2000.*

Ombao, H., Raz, J., von Sachs, R. and Malow, B. (2001*a*). Automatic statistical analysis of bivariate nonstationary time series, *J. Amer. Statist. Assoc.*, **96**, 543–560.

Ombao, H., Raz, J., Strawderman, R. and von Sachs, R. (2001*b*). A simple generalised cross validation method of span selection for periodogram smoothing, *Biometrika*, **88**(4) (to appear).

Priestley, M. (1965). Evolutionary spectra and non-stationary processes, *J. Roy. Statist. Soc. Ser. B*, **28**, 228–240.

Priestley, M. (1981). *Spectral Analysis and Time Series*, Academic Press, London.

Wickerhauser, M. (1994). *Adapted Wavelet Analysis from Theory to Software*, IEEE Press, Wellesley, Massachusetts.