SPATIO-SPECTRAL ANALYSIS OF BRAIN SIGNALS

Hernando Ombao[1], Xiaofeng Shao[2], Elena Rykhlevskaia[3],
Monica Fabiani[2] and Gabriele Gratton[2]

[1]*Brown University,* [2]*University of Illinois at Urbana-Champaign*
and [3]*Stanford University*

Abstract: Numerous studies in neuroscience have demonstrated links between brain spectral activity and cognition, sleep, disease diagnosis and treatment outcomes. In this paper, we study the variation of spectral activity across the brain cortical surface. We rigorously develop the concept of a location-dependent temporal spectrum for a wide class of spatio-temporal processes. Under the proposed asymptotic framework, the location-dependent spectrum can be estimated consistently. A non-parametric smoothing method is proposed and, under regularity conditions, is shown to be asymptotically normal and mean-square consistent. The paper concludes with an analysis of an event-related optical signal (EROS)data recorded from a single subject during a spatial-verbal Stroop task.

Key words and phrases: Event-related optical signal (EROS), kernel-smoothing, Spatio-temporal random process, spectrum.

1. Introduction

Neuroscientists are able to observe different facets of brain activity underlying cognitive function through non-invasive modalities such as electroencephalography, magnetoencephalography, functional magnetic resonance imaging and optical imaging. A substantial amount of research has focused on spectral activity, such as rhythms observed in the electroencephalogram (EEG). Some of them have investigated how these oscillatory or spectral phenomena vary over space (e.g., Gratton, Villa, Fabiani, Colombis, Palin, Bolcioni and Fiori (1992)). By and large, these studies have treated the time series observed at individual locations as separate measures, and then either compared the spectral properties of individual locations (thus generating amplitude or power maps observed at different locations), or analyzed the coherence between selected pairs of locations. Here we develop a rigorous statistical framework for the spectral analysis of data obtained from different locations using the Cramér representation of spatio-temporal processes.

We further consider an application of this approach to a data set obtained from optical imaging, which is a new technology that measures changes in the

light absorption and scattering properties of brain tissue following local event-related alterations in neuronal activity. As discussed in Cohen (1972), Stepnoski, Porta, Raccuia-Behling, Blonder, Slusher, and Kleinfeld (1991) and Gratton and Fabiani (1998), the phenomenon can plausibly be due to movement of ions across the neuronal membrane and the reorientation of the membrane proteins during neuronal activity. The event-related optical signal (EROS) possesses both good temporal and spatial resolution and thus is suited for studying the time course of activity in localized cortical areas. Gratton, Goodman-Wood and Fabiani (2001) demonstrated that EROS has good temporal correspondence with event-related brain potentials (ERPs). Moreover, Gratton, Fabiani, Corballis, Hood, Goodman-Wood, Hirsh, Kim, Friedman and Gratton (1997) demonstrated that EROS has good spatial correspondence with fMRI. EROS has been used in several neuroscience investigations, including cognitive aging (Fabiani, Low, Wee, Sable and Gratton (2006)) and active and passive oddball tasks (Low, Leaver, Kramer, Fabiani and Gratton (2006)). Recently, Rykhlevskaia, Fabiani and Gratton (2006) developed a lagged covariance model that uses EROS to study functional connectivity.

EROS, due to its good spatial (sub-centimeter) and temporal (millisecond scale) localization properties, may provide ideal data to investigate changes in the spectral behavior of the signal across extended cortical areas. The standard approach to quantifying spectral activity is through Fourier analysis (see Muthuswamy and Thakor (1998) and Dummermuth and Molinari (1987) among others). Most of these conventional approaches do not consider the spatial variation of the temporal spectrum. However, methods that incorporate spatial and topographic information in the spectral analysis of brain signals are being developed - see, for example, Koenig, Marti-Lopez and Valdes-Sosa (2001) and Hoogenboom, Schoffelen, Oostenveld, Parkes and Fries (2006). In this paper, our goal is to develop a general representation of spatio-temporal processes. Using the Cramér representation, we rigorously define the location-dependent (spatially-varying) temporal spectrum that is the primary quantity of interest. We develop a nonparametric method for estimating the spatially-varying temporal spectrum and an asymptotic framework under which we can derive asymptotic mean-square consistency and normality of the estimator.

The remainder of this paper is organized as follows. In Section 2, we discuss the Cramér representation of a spatio-temporal process and introduce the location-dependent temporal spectrum. In Section 3, we present our estimation procedure and establish the mean-square consistency and asymptotic normality. In Section 4, we analyze an EROS data set. We conclude with a discussion of our contribution, its limitations, and future work.

2. Spatio-Temporal Process

To study the oscillatory properties of brain signals, we first consider the Cramér representation of a stationary *univariate* temporal process $X(t) = \int_{-\pi}^{\pi} A(\lambda)$ $\exp(i\lambda t)dZ(\lambda)$, where $A(\lambda)$ is the complex-valued Hermitian transfer function and $Z(\lambda)$ is a stochastic process with zero mean uncorrelated increments, that is, $\mathbf{E}[dZ(\lambda)] = 0$ and $\mathrm{cov}[dZ(\lambda), dZ(\omega)] = \delta(\omega - \lambda)d\lambda$. Thus, a stationary temporal process may be viewed as a linear combination of infinitely many sinusoids (the Fourier waveforms $\exp(i\lambda t)$) having random coefficients $A(\lambda)dZ(\lambda)$. The variance of the random coefficient $A(\lambda)dZ(\lambda)$ is the **spectrum** of the process at frequency λ, $f(\lambda) = \mathbf{E}|A(\lambda)dZ(\lambda)|^2 = |A(\lambda)|^2 d\lambda$. The oscillatory content of signals is characterized by the spectrum which, in fact, is related to the variance decomposition of $X(t)$. Due to the orthonormality of the increments $dZ(\lambda)$, the variance of the stochastic process has the decomposition $\mathrm{var}[X(t)] = \int_{-\pi}^{\pi} f(\lambda)d\lambda$. Thus, one may interpret the temporal spectrum at some frequency λ_0 to be approximately the variation in $X(t)$ that is "explained" by the Fourier waveform $\exp(i\lambda_0 t)$ that oscillates at frequency λ_0.

We extend the Cramér representation to random processes that are observed across both *space and time.* In particular, we consider the setting where time series are observed across several locations in space. Let $X_t(s)$ be the spatio-temporal process defined on location $s = (s_1, s_2) \in \mathbb{R}^2$ and time point $t \in \mathbb{Z}$. Suppose that the spatio-temporal data at hand are observed at locations $s \in \{1, \ldots, n_1\} \times \{1, \ldots, n_2\}$ and time points $t \in \{1, \ldots, T\}$. There are a total of $n = n_1 n_2$ locations in space and a total of T time points. For this particular study, it is assumed that, for each location, the process is stationary over time; over space, the process need not be stationary.

The Cramér representation of the spatio-temporal process $X_t(s)$ is

$$X_t(s) = \int_{-\pi}^{\pi} A(u_{s,n}, \lambda) \exp(i\lambda t) dZ_s(\lambda), \tag{2.1}$$

where $s = (s_1, s_2)$ is the spatial index in the observation space and $u_s = u_{s,n} = (s_1/n_1, s_2/n_2)$ is the corresponding location in the re-scaled space. The elements of the above model are as follows.

1. $A(u, \lambda)$ is the location-dependent complex-valued transfer function. It is defined on a fixed re-scaled spatial domain $u = (u_1, u_2) \in [0, 1]^2$ and frequency domain $\lambda \in [-\pi, \pi]$. Moreover, it is Hermitian, that is, $A(.; -\lambda) = A^*(.; \lambda)$.
2. $Z_s(\lambda)$ is a complex-valued stochastic process that satisfies

$\mathbf{E}[dZ_s(\lambda)] = 0$ and $\mathrm{cov}[dZ_s(\lambda), dZ_{s'}(\lambda')] = \rho(s, s')\, \delta(\lambda - \lambda')d\lambda$ with $\rho(s, s) = 1$.

The primary quantity of interest in the above model is the location-dependent temporal spectrum

$$f(u, \lambda) = |A(u, \lambda)|^2 \tag{2.2}$$

defined for $u \in [0, 1]^2$ and $\lambda \in [-\pi, \pi]$. One may interpret $f(u_s, \lambda)$ as the variance of the time series at corresponding location s that is explained by the oscillation at frequency λ. For a univariate time series, as T grows to infinity, we have more and more information around the frequency λ, thus a consistent estimator can be constructed assuming certain smoothness conditions on the spectrum. Under our framework, as n_1 and n_2 grow to infinity, more and more data is being observed around any given location u, so a consistent estimator of $f(u, \lambda)$ is possible under smoothness conditions with respect to u. In other words, we require that $f(u, \lambda)$ be smooth across both space and frequency. More rigorous conditions are given in the next section. The adopted asymptotic framework is similar to that used by Dahlhaus (1997, 2000) for locally stationary temporal processes.

Remark 2.1. The temporal process at location s can be viewed as a linear combination of sinusoids with random coefficients $A(u_{s,n}, \lambda)dZ_s(\lambda)$ whose variance is the location-specific temporal spectrum $f(u_{s,n}, \lambda)$. The univariate temporal process is a special case of the above model. The classical Cramér spectral representation at a particular location s has increments $dZ_s(\lambda)$ that are uncorrelated across frequencies λ.

Remark 2.2. The spatio-temporal covariance between observations $X_t(s)$ and $X_{t'}(s')$ is

$$\text{cov}[X_t(s), X_{t'}(s')] = \rho(s, s') \int_{-\pi}^{\pi} A(u_s, \lambda) A^*(u_{s'}, \lambda) \exp(i\lambda(t - t')) d\lambda.$$

For a fixed location (i.e., $s = s'$), the process is stationary over time. Denote the temporal lag to be $h = t - t'$. Then the temporal covariance is

$$\text{cov}[X_t(s), X_{t'}(s)] = \int_{-\pi}^{\pi} f(u_s, \lambda) \exp(i\lambda h) d\lambda.$$

By setting the temporal lag $h = 0$, we obtain the location-specific variance decomposition $\text{var}[X_t(s)] = \int_{-\pi}^{\pi} f(u_s, \lambda) d\lambda$. For a fixed time t, the spatial covariance is

$$\text{cov}[X_t(s), X_t(s')] = \rho(s, s') \int_{-\pi}^{\pi} A(u_s, \lambda) A^*(u_{s'}, \lambda) d\lambda.$$

Thus, non-stationarity in space is allowed even if $\rho(s, s')$ depends only on the difference $s - s'$.

Remark 2.3. An example of the above process is the first order temporal auto-regressive (AR(1)) process with location-dependent coefficient: $X_t(s) = \phi(u_s)X_{t-1}(s) + \varepsilon_t(s)$ where $\phi(u)$ is the AR(1) coefficient at location u, $\{\varepsilon_t(s)\}$ is a spatial process that is independent over time with $\mathbf{E}\varepsilon_t(s) = 0$ and $\text{var}(\varepsilon_t(s)) = 1$. The location-dependent transfer function is $A(u, \lambda) = (\sqrt{2\pi})^{-1}\sum_{j=0}^{\infty}[\phi(u)\exp(-i\lambda)]^j$ that gives rise to the equivalent Cramér representation $X_t(s) = \int_{-\pi}^{\pi}(\sqrt{2\pi})^{-1}[1 - \phi(u_s)e^{-i\lambda}]^{-1}\exp(i\lambda t)dZ_s(\lambda)$. The corresponding location-dependent temporal spectrum is $f(u, \lambda) = |1 - \phi(u)\exp(-i\lambda)|^{-2}/(2\pi)$.

Remark 2.4. The spatio-temporal model and estimation method that we develop in this paper differs from that by Fuentes (2002). This work is aimed at modelling the spatial variation of the *temporal spectrum*, while Fuentes (2002) deals with the modelling and estimation of the *spatial spectrum*.

Remark 2.5. The primary contribution of the paper is the generalization of spatio-temporal processes to the situation where the temporal spectrum is allowed to vary across space according to some smoothness conditions. Our framework has the transfer function $A(\cdot, \cdot)$ defined on a domain where space is rescaled to a unit square (or cube) and it allows for a consistent spectral estimator at any location in the fixed spatial domain. This framework also suggests a natural estimator whose asymptotic properties are developed in the next section.

3. Estimation Theory

For univariate temporal stationary processes, the spectrum can be consistently estimated by noting its relationship with the auto-covariance. Let X_t be a zero mean stationary process with an auto-covariance sequence $\gamma(k) = \mathbf{E}[X_{t+k}X_t]$ that satisfies $\sum_k |\gamma(k)| < \infty$. The spectrum of X_t is

$$f(\lambda) = \frac{1}{2\pi}\sum_k \gamma(k)\exp(-i\lambda k).$$

Write the sample auto-covariance as $\widehat{\gamma}(k) = T^{-1}\sum_{t=1}^{T-|k|}X_t X_{t+|k|}$. The lag window estimator (Brockwell and Davis (1991)) for $f(\lambda)$ is defined to be $\widehat{f}(\lambda) = (2\pi)^{-1}\sum_{k=-B_T}^{B_T}a(kb_T)\widehat{\gamma}(k)\exp(-i\lambda k)$, where $a(\cdot)$ is a window function, $B_T = b_T^{-1}$ is the bandwidth, $B_T \to \infty$, and $B_T = o(T)$ as $T \to \infty$.

For a fixed spatial location u and frequency λ, the location-dependent temporal spectrum $f(u, \lambda)$ can be estimated as follows. We first compute the lag window estimator at each location s_j, $j = 1, \ldots, n = n_1 n_2$, denoted by

$$\widehat{f}(u_{s_j}, \lambda) = \frac{1}{2\pi}\sum_{k=-B_T}^{B_T}\widehat{\gamma}(s_j; k)a(kb_T)e^{-ik\lambda},$$

where $\widehat{\gamma}(s_j; k) = T^{-1} \sum_{t=1}^{T-|k|} X_t(s_j) X_{t+|k|}(s_j)$. We then form an estimator of $f(u, \lambda)$ by smoothing the lag window estimates within a spatial neighborhood of u, that is,

$$\widetilde{f}(u, \lambda) = \sum_{j=1}^{n} w_{jn}(u) \widehat{f}(u_{s_j}, \lambda),$$

where the weights $w_{jn}(u)$ are nonnegative, and $\sum_{j=1}^{n} w_{jn}(u) = 1$.

To facilitate the investigation of asymptotic distributional properties of $\widetilde{f}(u, \lambda)$, we consider a smoothly spatially-varying temporal **linear** process as a special case of the spatio-temporal model in (2.1). Let

$$X_t(s) = \sum_{j=0}^{\infty} b_j(u_{s,n}) \varepsilon_{t-j}(s), \tag{3.1}$$

where $b_j(\cdot)$, $j = 0, 1, \cdots$ are all smooth functions of $u \in [0, 1]^2$. We assume the innovation processes $\{\varepsilon_t(s)\}$ are spatially stationary processes that are independent over time with $\mathbf{E}\varepsilon_t(s) = 0$ and $\text{var}(\varepsilon_t(s)) = 1$. Under (3.1), the location dependent temporal spectrum is

$$f(u, \lambda) = |A(u, \lambda)|^2, \quad \text{where} \quad A(u, \lambda) = \frac{1}{\sqrt{2\pi}} \sum_{j=0}^{\infty} b_j(u) e^{-ij\lambda}.$$

The theoretical argument is an extension of Anderson (1971) from the estimation of the spectrum for linear processes to the estimation of spatially-varying temporal spectrum. The spatially-smoothed temporal spectrum is a natural estimator, although, as we demonstrate, the derivation of its asymptotic properties is not trivial.

3.1. Assumptions

Assumption 3.1. $\mathbf{E}\varepsilon_0^4(s) < \infty$,

$$\Omega := \sup_{u \in [0,1]^2} \sum_{j=0}^{\infty} |b_j(u)| < \infty, \quad \text{and} \quad \lim_{m \to \infty} \sup_{u \in [0,1]^2} \sum_{k=m+1}^{\infty} |b_k(u)| = 0.$$

Assumption 3.2. There exists a constant C and a positive integer m_0 such that

$$\sup_{u \in [0,1]^2} \left| \sum_{j=0}^{m} \frac{\partial b_j(u)}{\partial u} e^{ij\lambda} \right| \leq C \text{ for } m \geq m_0.$$

The assumptions impose smoothness conditions on $f(u, \lambda)$ in space and frequency. The first of them implies that for each location s, $X_t(s)$ is a short memory time series, while the second is slightly stronger than the condition $\sup_{u \in [0,1]^2} |\partial A(u, \lambda)/\partial u| \leq C$, which implies the continuous differentibility of $A(u, \lambda)$ with respect to u. We note that the temporal auto-regressive process with location-dependent coefficients satisfies the assumptions.

Assumption 3.3. The window function $a(\cdot)$ is an even, Lipschitz continuous function with support on $[-1, 1]$, and $a(0) = 1$.

Most commonly-used window functions satisfy Assumption 3.3; see Priestley (1981). Write $C(j_1, j_2, \ldots, j_k) = \mathrm{cum}(\varepsilon_0(s_{j_1}), \varepsilon_0(s_{j_2}), \ldots, \varepsilon_0(s_{j_k}))$ for $k = 3, 4$ and $\rho(s_{j_1} - s_{j_2}) = \mathrm{cov}(\varepsilon_0(s_{j_1}), \varepsilon_0(s_{j_2}))$. Hereafter we shall use $n \to \infty$ to mean $\min(n_1, n_2) \to \infty$, for convenience.

Assumption 3.4. There exists a sequence $H_n(u) \to \infty$ as $n \to \infty$, such that

$$\lim_{n \to \infty} H_n(u) \sum_{j,j'=1}^{n} w_{jn}(u) w_{j'n}(u) \rho^2(s_j - s_{j'}) = 1 \quad \text{and} \tag{3.2}$$

$$\limsup_{n \to \infty} H_n(u) \sum_{j,j'=1}^{n} w_{jn}(u) w_{j'n}(u) |C(j, j, j', j')| < \infty. \tag{3.3}$$

Further $w_{jn}(u) = 0$ if $|u_j - u| > G_n$, where $G_n \to 0$ when $n \to \infty$.

Remark 3.1. The weights $w_{jn}(u)$ take the form $w_{jn}(u) = K_n(s - s_j)$, where $s = (\lfloor u_1 n_1 \rfloor, \lfloor u_2 n_2 \rfloor)$. Here $\lfloor a \rfloor$ is the integer part of a. In our application, we take the kernel function K_n to be the tensor product of two one-dimensional kernels: $K_n(s) = K_{H_{1n}}^{(1)}(s_1) K_{H_{2n}}^{(2)}(s_2)$, where H_{1n} and H_{2n} are two positive integers that satisfy $H_{jn} = o(n_j)$, $j = 1, 2$, as $n \to \infty$. For example,

$$K_{H_{jn}}^{(j)}(s_j) = \begin{cases} \frac{1}{H_{jn}} & -\lfloor \frac{(H_{jn}-1)}{2} \rfloor \leq s_j \leq \lfloor \frac{H_{jn}}{2} \rfloor \\ 0 & \text{otherwise} \end{cases} \quad j = 1, 2. \tag{3.4}$$

Assumption 3.5. For any three permutations of $(1, 2, 3, 4)$, (i_1, i_2, i_3, i_4), (k_1, k_2, k_3, k_4) and (g_1, g_2, g_3, g_4), we have

$$\sum_{j_1, j_2, j_3, j_4=1}^{n} w_{j_1 n}(u) w_{j_2 n}(u) w_{j_3 n}(u) w_{j_4 n}(u) \{ |C(j_{i_1}, j_{i_2}, j_{i_3})| |C(j_{k_1}, j_{k_2}, j_{k_3})| |\rho(s_{j_{g_1}} - s_{j_{g_2}})|$$

$$+ |C(j_1, j_2, j_3, j_4)|^2 + |C(j_1, j_2, j_3, j_4)| |\rho(s_{j_{g_1}} - s_{j_{g_2}}) \rho(s_{j_{g_3}} - s_{j_{g_4}})|$$

$$+ |\rho(s_{j_{g_1}} - s_{j_{g_2}}) \rho(s_{j_{g_3}} - s_{j_{g_4}}) \rho(s_{j_{k_1}} - s_{j_{k_2}}) \rho(s_{j_{k_3}} - s_{j_{k_4}})| \} = O(H_n(u)^{-2}). \tag{3.5}$$

Remark 3.2. Note that if the process $\{\varepsilon_0(s), s \in \mathbb{Z}^2\}$ is Gaussian, then (3.3) holds trivially, and the first three terms in the curly bracket of (3.5) vanish. Let $\Pi^2 = [-\pi, \pi]^2$. Define $g(\lambda) = (2\pi)^{-2} \sum_{h \in \mathbb{Z}^2} \rho^2(h) e^{ih'\lambda}$. Under (3.4), if $g(0) \in (0, \infty)$, then

$$\sum_{j,j'=1}^{n} w_{jn}(u) w_{j'n}(u) \rho^2(s_j - s_{j'})$$

$$= \int_{\Pi^2} \Big| \sum_{j=1}^{n} w_{jn}(u) e^{i(s_j - u)'\lambda} \Big|^2 g(\lambda) d\lambda \sim [H_{1n} H_{2n}]^{-1} g(0).$$

Here $a_n \sim b_n$ means $\lim_{n \to \infty} a_n / b_n = 1$. So we can take $H_n(u) = H_{1n} H_{2n} g(0)^{-1}$. If $g(0) = \infty$, then $H_n(u)$ also depends on the decay rate of $g(\lambda)$ as $\lambda \downarrow 0$. In this case, the process $\{\varepsilon_0(s), s \in \mathbb{Z}^2\}$ possesses long memory. Regarding the last term in the curly bracket of (3.5), the result follows from the Cauchy-Schwarz inequality in view of (3.2) and the fact that $w_{jn}(u) \in [0, 1]$.

The following assumptions are made to obtain the explicit form of the bias; see Theorem 3.2.

Assumption 3.6.

$$\sup_{u \in [0,1]^2} \sum_{j=0}^{\infty} j^2 |b_j(u)| < \infty, \quad \text{and} \quad \lim_{x \to 0} x^{-2}(1 - a(x)) = c_2 \in (0, \infty). \quad (3.6)$$

Assumption (3.6) implies that $f(u, \lambda)$ is twice continuous differentiable with respect to λ. We write $f_{\lambda\lambda}(u, \lambda) = \partial^2 f(u, \lambda)/\partial \lambda^2$.

Assumption 3.7. $f(u, \lambda)$ and $f_{\lambda\lambda}(u, \lambda)$ are twice continuously differentiable with respect to u.

For $u = (u_1, u_2)'$, write $f_j(u, \lambda) = \partial f(u, \lambda)/\partial u_j$, $f_{jk}(u, \lambda) = \partial^2 f(u, \lambda)/\partial u_j \partial u_k$, $f_{\lambda\lambda j}(u, \lambda) = \partial f_{\lambda\lambda}(u, \lambda)/\partial u_j$, $f_{\lambda\lambda jk}(u, \lambda) = \partial^2 f_{\lambda\lambda}(u, \lambda)/\partial u_j \partial u_k$, for $j, k = 1, 2$.

Assumption 3.8. Suppose $w_{jn}(u) = 0$ if $|u_{j1} - u_1| > G_{1n}$ or $|u_{j2} - u_2| > G_{2n}$, where $G_{1n}, G_{2n} \to 0$ as $n \to \infty$. Assume the weight $w_{jn}(u)$ is a tensor product of of two symmetric one-dimensional kernels (see Remark 3.1) in the sense that

$$\sum_{j=1}^{n} w_{jn}(u)(u_j - u) = 0, \quad \sum_{j=1}^{n} w_{jn}(u)(u_{j1} - u_1)(u_{j2} - u_2) = 0$$

There exists two sequences $\{L_{1n}(u)\}$ and $\{L_{2n}(u)\}$, such that

$$\lim_{n \to \infty} L_{kn}(u) = 0, \quad \lim_{n \to \infty} L_{kn}^{-1}(u) \sum_{j=1}^{n} w_{jn}(u)(u_{jk} - u_k)^2 = 1, \; k = 1, 2$$

3.2. Main results

Theorem 3.1. *Under Assumptions 3.1−3.5, we have*

$$\sqrt{Tb_T}\sqrt{H_n(u)}(\widetilde{f}(u,\lambda) - \mathbf{E}[\widetilde{f}(u,\lambda)]) \to_D N(0,\sigma^2(u,\lambda)), \qquad (3.7)$$

where $\sigma^2(u,\lambda) := \int_{-1}^1 a^2(u)du f^2(u,\lambda)(1 + \eta(\lambda))$, $\eta(\lambda) = 1$ if $\lambda = k\pi$ and 0 otherwise.

Remark 3.3. For univariate spectrum estimation, asymptotic normality has been established by Brillinger (1969), Anderson (1971), Rosenblatt (1985) and, more recently, Shao and Wu (2007) in a variety of settings. For fixed $u_j = u_{s_j}$, it is true that (see Anderson (1971))

$$\sqrt{Tb_T}(\widehat{f}(u_j,\lambda) - \mathbf{E}\widehat{f}(u_j,\lambda)) \to_D N(0,\sigma^2(u_j,\lambda)).$$

Due to the smoothing in space, we have a further reduction of variance by a factor of $H_n(u)$. Note that this comes at the cost of at the cost of a higher bias or a reduced spatial resolution.

Remark 3.4. There is an interesting interplay between smoothing in frequency and in space. If there is a higher priority placed on frequency resolution than spatial resolution, one should construct a lag window $a(\cdot)$ that tapers quickly and a smoothing window that is broad. If a spatial resolution cannot be compromised, then one must constrict smoothing in space but allow the time lag window $a(\cdot)$ to decay slowly.

Remark 3.5. To perform statistical inference (compute confidence intervals or conduct tests of hypothesis), one would need to estimate the unknown factor in $H_n(u)$ that characterizes the dependence of the underlying innovation process $\{\varepsilon_0(s), s \in \mathbb{Z}^2\}$; see Remark 3.2. A possible solution is to approximate the sampling distribution of $\widetilde{f}(u,\lambda)$ by performing a spatial block bootstrap within a small neighborhood of u, i.e. $\{u_j : w_{jn}(u) \neq 0\}$. However, due to a small number of locations in the EROS dataset, we do not pursue this possibility here.

Next, we consider the joint convergence of $\{\sqrt{Tb_T H_n(u^{(j)})}(\widetilde{f}(u^{(j)},\lambda_k) - \mathbf{E}\widetilde{f}(u^{(j)},\lambda_k))\}_{j,k=1,2}$, where $u^{(1)} \neq u^{(2)}$ and $\lambda_1 \neq \lambda_2 \in [0,\pi]$. We introduce the following extra assumption.

Assumption 3.9. For $u^{(1)} \neq u^{(2)}$, we have

$$\sum_{j,j'=1}^n w_{jn}(u^{(1)})w_{j'n}(u^{(2)})[\rho^2(s_j - s_{j'}) + |C(j,j,j',j')|]$$
$$= o(H_n(u^{(1)})^{-\frac{1}{2}}H_n(u^{(2)})^{-\frac{1}{2}}).$$

Remark 3.6. For $k = (k_1, k_2)$, let $\kappa(k) = \sup_{|h| \geq |k|} |\rho(h)|$. Since $w_{jn}(u) = 0$ if $|u_j - u| > G_n$, we have

$$\sum_{j,j'=1}^{n} w_{jn}(u^{(1)}) w_{j'n}(u^{(2)}) \rho^2(s_j - s_{j'})$$

$$\leq \kappa \left(\frac{1}{2} \sqrt{\lfloor n_1(u_1^{(1)} - u_1^{(2)}) \rfloor^2 + \lfloor n_2(u_2^{(1)} - u_2^{(2)}) \rfloor^2} \right).$$

Thus Assumption 3.9 is satisfied if the right side of the above equation is $o(H_n(u^{(1)})^{-1/2} H_n(u^{(2)})^{-1/2})$ and the innovation process $\{\epsilon_0(s), s \in \mathbb{Z}^2\}$ is Gaussian.

Corollary 3.1. *Under Assumptions* 3.1−3.5 *and* 3.9, $\{\sqrt{Tb_T H_n(u^{(j)})}(\widetilde{f}(u^{(j)}, \lambda_k) - \mathbf{E}\widetilde{f}(u^{(j)}, \lambda_k))\}$ *are jointly asymptotically independent* $N(0, \sigma^2(u^{(j)}, \lambda_k))$, $j, k = 1, 2$.

Next we state the explicit form of the mean square error of $\widetilde{f}(u, \lambda)$.

Theorem 3.2. *Under Assumptions* 3.1−3.8 *we have*

$$\mathbf{E}|\widetilde{f}(u, \lambda) - f(u, \lambda)|^2$$
$$= \sigma^2(u, \lambda)(Tb_T H_n(u))^{-1} + \frac{1}{2}[L_{1n}(u)f_{11}(u, \lambda) + L_{2n}(u)f_{22}(u, \lambda)]$$
$$+ o((Tb_T H_n(u))^{-1}) + o(b_T^2) + o(L_{1n}(u) + L_{2n}(u)). \tag{3.8}$$

Proof of Theorem 3.2. In view of Theorem 3.1 and its proof, it suffices to find the explicit form of the bias of $\widetilde{f}(u, \lambda)$. Note that

$$\mathbf{E}\widetilde{f}(u, \lambda) - f(u, \lambda)$$
$$= \sum_{j=1}^{n} w_{jn}(u)[\mathbf{E}\widehat{f}(u_j, \lambda) - f(u, \lambda)]$$
$$= \sum_{j=1}^{n} w_{jn}(u)[\mathbf{E}\widehat{f}(u_j, \lambda) - f(u_j, \lambda)] + \sum_{j=1}^{n} w_{jn}(u)[f(u_j, \lambda) - f(u, \lambda)]$$
$$= I_1 + I_2.$$

By Taylor's expansion, we get, in view of Assumption 3.8,

$$I_2 = \frac{1}{2} \sum_{j=1}^{n} w_{jn}(u) \Big\{ (u_{j1} - u_1)^2 f_{11}(u, \lambda) + 2(u_{j1} - u_1)(u_{j2} - u_2)f_{12}(u, \lambda)$$
$$+ (u_{j2} - u_2)^2 f_{22}(u, \lambda) \Big\} + o(L_{1n}(u) + L_{2n}(u))$$
$$= \frac{1}{2}[L_{1n}(u)f_{11}(u, \lambda) + L_{2n}(u)f_{22}(u, \lambda)] + o(L_{1n}(u) + L_{2n}(u)).$$
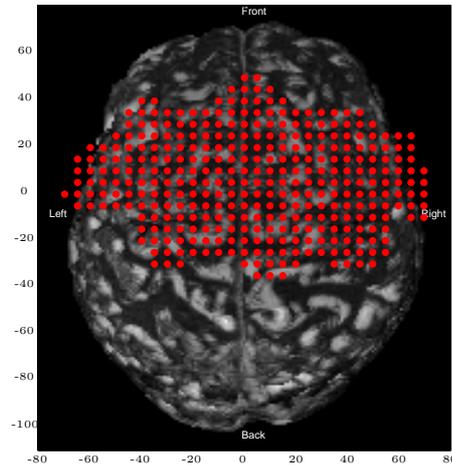
Figure 1. Locations of recording channels on the cortical surface.

By Anderson (1971, Thm 9.4.3) or Priestley (1981, p.459), we have $B_T^2(\mathbf{E}\widehat{f}(u_j, \lambda)$ $-f(u_j, \lambda)) \to c_2 f_{\lambda\lambda}(u_j, \lambda)$. So

$$I_1 = B_T^{-2} \sum_{j=1}^{n} w_{jn}(u) c_2 f_{\lambda\lambda}(u_j, \lambda) + o(B_T^{-2})$$

$$= c_2 B_T^{-2} \frac{1}{2}[L_{1n}(u) f_{\lambda\lambda11}(u, \lambda) + L_{2n}(u) f_{\lambda\lambda22}(u, \lambda)] + o((L_{1n}(u) + L_{2n}(u))b_T^2)$$

$$+ o(b_T^2)$$

$$= o(b_T^2),$$

where the last equality follows from Assumption 3.8. Therefore (3.8) follows.

Remark 3.7. In Theorem 3.2, the exact form of the dominant term depends on the form of the kernel function K_n. If we use the simple kernel (3.4), then it is

$$\sigma^2(u, \lambda)(Tb_T H_n(u))^{-1} + \frac{H_{1n}^2}{24n_1^2} f_{11}(u, \lambda) + \frac{H_{2n}^2}{24n_2^2} f_{22}(u, \lambda).$$

4. Data Analysis

The EROS data set was recorded from a single participant in a spatial-verbal Stroop task experiment conducted at the Cognitive Neuroimaging Laboratory at the University of Illinois at Urbana-Champaign. This experiment was designed to explore brain activity induced by the cognitive effort due to preparatory processes
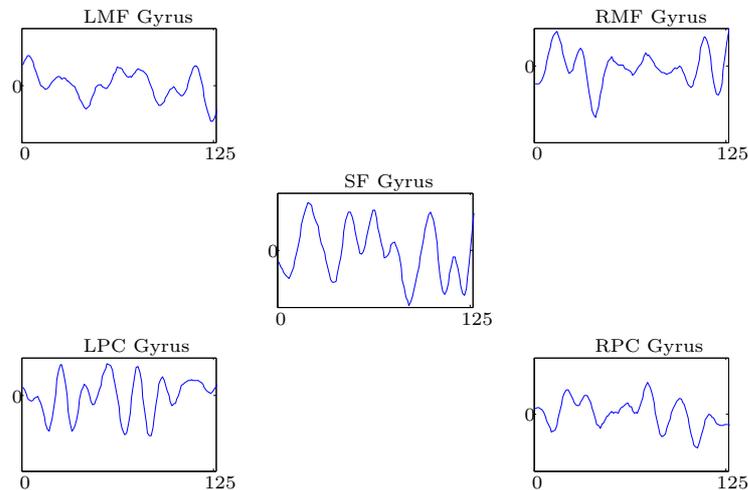
Figure 2. Event-related optical signals at different cortical locations. Each time series has length $T = 125$. We first computed the average across trials that were time-locked to the position (P) cue preceded by meaning (M) cue; we then computed the average across trials that were time locked to the position (P) cue preceded by the position (P) cue. The plotted time series is the difference between the switched cue and the no-switch cue at the ampling rate: one time point per 16 milliseconds. The specific Talairach locations on the cortex where the time series where extracted were top left: left medio-frontal gyrus ($x = -40$ mm, $y = 20$ mm); top right: right medio-frontal gyrus ($x = 40$ $y = 20$); center: superior-frontal gyruns ($x = 0$, $y = -12$); bottom Left: left posterior-central gyrus ($x = -50$, $y = -28$); bottom Right: right posterior-central gyrus ($x = 50$, $y = -28$).

associated with task switching. The task switching in this paradigm is defined as switching between processing spatial or verbal features of stimuli. Each trial included the visual presentation of the words "Above" or "Below" that were located either above or below a central fixation cross, sometimes, the word "Above" would actually appear below the cross. The subject was instructed to press one button for above and another for below. Across the different trials, the subject was instructed to pay attention to either the word position or the word meaning (denoted by a "P" or "M" cue presented two seconds prior to the actual stimulus). Therefore, during the two seconds following the cue, the participant had to prepare to respond to the spatial (position) or verbal (meaning) features of the stimulus. In the current paper, we consider data obtained during this two second period.

    The EROS data, specifically the phase delay data, was recorded with a multi-

channel frequency-domain optical instrument, with a modulation frequency of 220 Hertz and an effective sampling rate of 16 milliseconds, over 160 source-detector pairs located on both the left and right frontal cortex. A temporal low pass filter was applied with a cutoff of 8 Hertz. Separate averages, time-locked to the cues, were computed for each type of cue (P or M), switch (from P to M or from M to P) or no-switch trial types and recording channel. Difference time series between the switch and no-switch trial types were then computed to isolate brain activity related to task switching. The difference between switch and non-switch conditions is interpreted as the change in brain activation from control condition to switching. It is assumed that the non-switch condition provides a baseline brain activation, whereas the switch condition is the more demanding activity that causes the brain to engage additional resources and activate additional areas. Analogous to baseline conditions in fMRI studies, this experiment used the non-switch condition as that which accounts for brain activity due to visual stimulation and other processes activated by the experimental situation in general, but not specific to task switching. Finally, we note that averaging signals has long been a tradition in ERP studies with the goal of improving the signal to noise ratio. The EROS analysis software currently available provides only averages rather than single trial recordings as a data output. In this analysis, we treat these averaged time series as though they were an ensemble realization from a spatio-temporal process; this is a limitation of our application.

The source and detector locations were digitized and then mapped onto a surface image of the brain, thus allowing for surface image reconstructions of the optical activity for each data point in Talairach space. This yielded a three-dimensional data matrix (x and y surface locations, and time). Each time series was associated with a particular location, with 289 locations sampled on a grid with a 5 millimeter step size. For this particular analysis, we only used 21×11 time series located on the (x, y) grid $[-50, 50] \times [-30, 24]$ square millimeter surface that was common to all 32 subjects in the experiment (although data from only one subject are presented for the current illustration).

Location-dependent normalized temporal spectra were computed using a frequency bandwidth of $B_T = 5$ for the two tasks (position and meaning, each representing the activity recorded in the switch condition minus the activity in the no-switch condition). Moreover, the difference in the relative spectral estimates between position and meaning features was computed. The results are shown in Figure 3 (power at 2 Hertz), Figure 4 (power at 4 Hertz) and Figure 5 (power at 6 Hertz).

First, we note that the relative powers at 2 Hertz and 4 Hertz oscillatory brain activity show similar spatial distribution, although the power at 4 Hertz is of smaller magnitude. For 2 Hertz, one notes that the power for the meaning
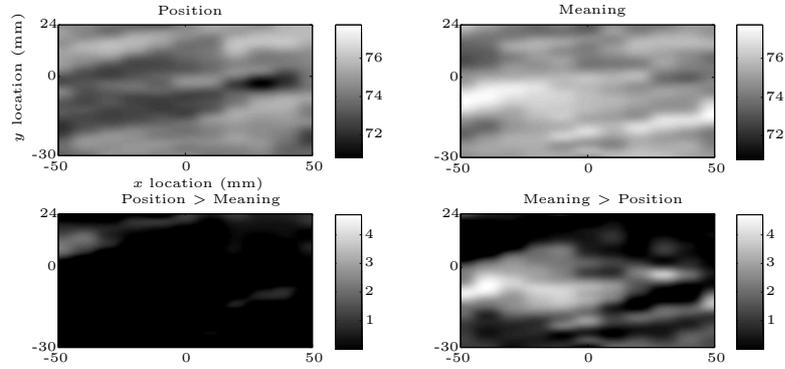
Figure 3. Spatial distribution of spectral power at 2 Hertz. Smoothing bandwidth in frequency $B_T = 5$. *Top Left*: **Position** Task. Location-dependent normalized temporal spectral estimate of the difference time series between switch and no-switch cues. *Top Right*: **Meaning** Task. Location-dependent normalized temporal spectral estimate of the difference time series between switch and no-switch cues. *Bottom Left*: Bright spots indicate power at position (spatial) is greater than at the meaning (verbal) task. *Bottom Right*: Bright spots indicate power is greater at the meaning (verbal) than at the position (spatial) task.
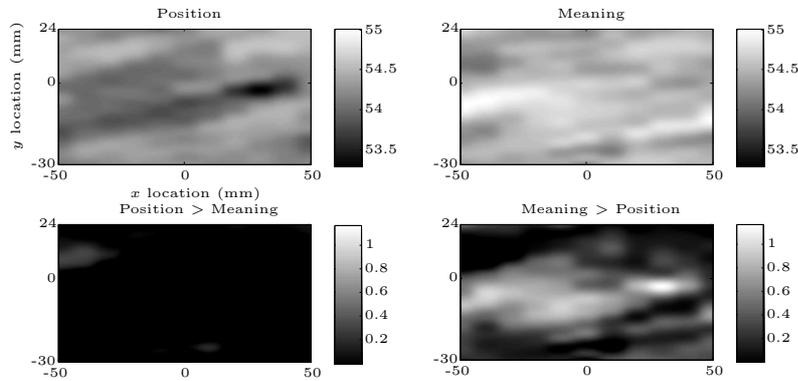

Figure 4. Spatial distribution of spectral power at 4 Hertz. Smoothing bandwidth in frequency $B_T = 5$. *Top Left*: **Position** Task. Location-dependent normalized temporal spectral estimate of the difference time series between switch and no-switch cues. *Top Right*: **Meaning** Task. Location-dependent normalized temporal spectral estimate of the difference time series between switch and no-switch cues. *Bottom Left*: Bright spots indicate power at position (spatial) is greater than at the meaning (verbal) task. *Bottom Right*: Bright spots indicate power is greater at the meaning (verbal) than at the position (spatial) task.
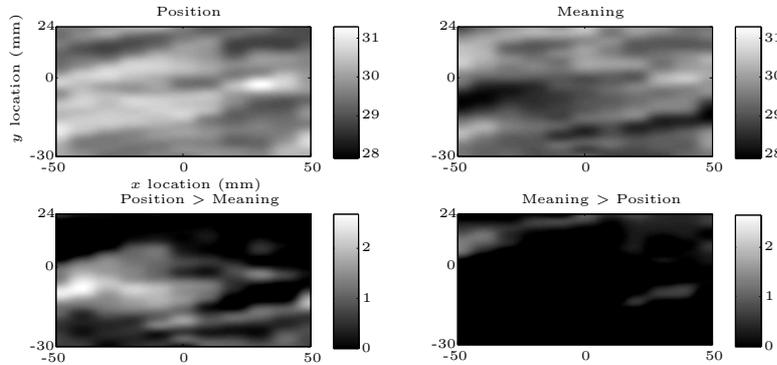
Figure 5. Spatial distribution of spectral power at 6 Hertz. Smoothing bandwidth in frequency $B_T = 5$. *Top Left*: **Position** Task. Location-dependent normalized temporal spectral estimate of the difference time series between switch and no-switch cues. *Top Right*: **Meaning** Task. Location-dependent normalized temporal spectral estimate of the difference time series between switch and no-switch cues. *Bottom Left*: Bright spots indicate power at position (spatial) is greater than at the meaning (verbal) task. *Bottom Right*: Bright spots indicate power is greater at the meaning (verbal) than at the position (spatial) task.

task is substantially greater than that for the position task over a wide spatial distribution, although the hottest spots were highly spatially localized to the superior-frontal gyrus and the left posterior-central gyrus. Interestingly, the spatial distribution at 6 Hertz showed a reverse picture: areas showing a large amount of 2 Hertz and 4 Hertz activity showed a particularly low amount of the 6 Hertz activity. Although this phenomenon may be partly related to the fact that spectra are normalized during the procedure, this fact alone does not account for the similar distribution of 2 and 4 Hertz activity. We postulate that the larger activity in the 2-4 Hz bands for verbal, and 6 Hz band for spatial, reflects the fact that theta activity is more obvious for the verbal condition (which is in fact the most difficult one). Note that this particular data analysis took advantage of the location-specific distribution of the power estimates. Of course, the main result shows a dominant lateralization in the left hemisphere, especially for the meaning task. The results show that slow (theta) activity is localized to the left hemisphere for the switch-to-meaning condition (when compared to "switch-to-position" condition). This may reflect the fact that verbal processes are left-hemisphere dominant. Alternatively, it may reflect the fact the left prefrontal cortex may be more involved when switching to a more difficult task (in this case, the meaning task) is required.

5. Conclusion

This paper rigorously develops a general Cramér representation for a wide class of spatio-temporal processes, motivated by the need to investigate the variation of the oscillatory activity across cortical regions. Previous work in this area has not taken specific advantage of the spatial nature of brain imaging data. A non-parametric estimation method is proposed to estimate the location-dependent temporal spectrum, a quantity that describes the spatial variation of temporal spectral behavior. For a class of temporal linear processes with spatially varying coefficients, asymptotic normality and mean-square consistency are established for the proposed estimator. Using the proposed model and procedure, we estimated the location-dependent temporal spectrum for the EROS data set.

The proposed approach is potentially useful for neuroscience research. It has been shown that different cognitive processes (or, most likely, their components and sub/micro processes) involve brain networks that oscillate at particular frequencies. However, there have been no approaches that address the issue of variation in the spectrum of spatially adjacent locations. In part, this happened because the types of data available in neuroscience did not possess good spatial resolution to start with (or, if they did, then they did not have sufficient temporal resolution for power-frequency analysis). The proposed approach responds to the development of new imaging tools employed by neuroscientists. For instance, in our example, we were able to take advantage of the combination of good temporal and spatial resolution of EROS. We showed that the two different tasks induce similar spatial distributions of spectral power at specific frequencies. We are careful to note the limitations of the applicability of this model to neuroscience, because (spatial) proximity does not imply the existence of functional connections.

There are theoretical limitations of our current work. The time series at each location is constrained to be a linear process and this excludes interesting non-linear time series models which could be better suited at modelling complex phenomena such as brain processes. However, linear models such as the autoregressive moving average (ARMA) are shown to be useful for modelling the temporal autocorrelation structure in fMRI and also for feature selection in EEG classification studies. Moreover, the spatio-temporal model and asymptotic framework in this paper serves as a starting point for a number of future research directions. First, we are developing a data-adaptive spatial bandwidth selection procedure for estimating the location-dependent temporal spectrum. Next, to enable applicability of the asymptotic distribution of the proposed estimator to tests of difference between the conditions, we are investigating the use of a spatial bootstrap to approximate the sampling distribution of our estimator. Moreover, motivated by

the potential usefulness of higher order spectra (bispectrum, in particular) for different types of neuroscience applications (e.g., studies of functional connectivity, classifying sleep depth of anesthesized patients, etc.), one may develop a location-dependent temporal bispectrum under the framework that was developed in this paper. Finally, we are generalizing the proposed model to the setting where, at each location, the time series is non-stationary. We may consider some non-stochastic temporal processes such as the autoregressive moving average models with time-varying coefficients, temporal processes with stochastic representations that use localized waveforms (see Nason, von Sachs and Kroisandt (2000) and Ombao, H., Raz, J. von Sachs, R. and Guo, W. (2002)) and the Dahlhaus (1997) model of locally stationary processes.

Acknowledgements

References

Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. Wiley, New York.

Brillinger, D. R. (1969). Asymptotic properties of spectral estimates of second order. *Biometrika* **56**, 375-390.

Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods.* Second edition. Springer, New York.

Cohen, L. (1972). Changes in neuron structure during action potential propagation and synaptic transmission. *Phys. Rev.* **53**, 373-417.

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

Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *Ann. Statist.* **28**, 1762-1794.

Dummermuth, H. and Molinari, G. (1987). Spectral Analysis of the EEG: Some Fundamentals Revisited and Some Open Problems. *Neuropsychobiology* **17**, 85-99.

Fabiani, M., Low, K., Wee, E., Sable, J. and Gratton, G. (2006). Reduced suppression or labile memory? Mechanisms of inefficient filtering of irrelevant information in older adults. *J. Cognitive Neuroscience* **18**, 637-650.

Fuentes, M. (2002). Spectral methods for nonstationary spatial processes. *Biometrika* **89**, 197-210.

Gratton, G., Villa, A., Fabiani, M., Colombis, G., Palin, E., Bolcioni, G., and Fiori, M. (1992). Functional correlates of a three-component spatial model of the alpha rhythm. *Brain Research* **582**, 159-162.

Gratton, G. Fabiani, M. Corballis, P., Hood, D., Goodman-Wood, M., Hirsh, J., Kim, J., Friedman, D. and Gratton, E. (1997). Fast and localized event-related optical signals (EROS) in the human occipital cortex: Comparison with the visual evoked potential and fMRI. *NeuroImage* **6**, 168-180.

Gratton, G. and Fabiani, M, (1998). Dynamic brain imaging: Event-related optical signal (EROS) of the time course annd localization of cognitive-related activity. *Psychonomic Bulletin and Review* **5**, 535-563.

Gratton, G., Goodman-Wood, M. and Fabiani, M. (2001). Comparison of neuronal and hemo-dynamic measures of the brain response to visual stimulation: an optical imaging study. *Human Brain Mapping* **13**, 13-25.

Hoogenboom, N., Schoffelen, J-M., Oostenveld, R., Parkes, L. and Fries, P. (2006). Localizing human visual gamma-band activity in frequency, time and space. *NeuroImage* **29**, 764-773.

Koenig, T., Marti-Lopez, F. and Valdes-Sosa, P. (2001). Topographic time-frequency decompo-sition of the EEG. *NeuroImage* **14**, 383-390.

Low, K., Leaver, E., Kramer, A., Fabiani, M. and Gratton, G. (2006). Fast optical imaging of frontal cortex during active and passive oddball tasks. *Psychophysiology* **43**, 127-136.

Muthuswamy, J. and Thakor, N. (1998). Spectral analysis methods for neurological signals. *J. Neuroscience Methods* **83**, 1-14.

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.

Ombao, H., Raz, J. von Sachs, R. and Guo, W. (2002). The SLEX model of a non-stationary random process. *Ann. Inst. Statist. Math.* **54**, 171-200.

Priestley, M. B. (1981). *Spectral Analysis and Time Series* 1. Academic, New York.

Rykhlevskaia, E., Fabiani, M. and Gratton, G. (2006). Lagged covariance structure models for studying functional connectivity. *NeuroImage* **30**, 1203-1218.

Rosenblatt, M. (1985). *Stationary Sequences and Random Fields.* Birkhäuser, Boston.

Shao, X. and Wu, W. B. (2007). Asymptotic spectral theory for nonlinear time series. *Ann. Statist.* **35**, 1773-1801.

Stepnoski, R., Prta, A., Raccuia-Behling, F. Blonder, G., Slusher, R. and Kleinfeld, D. (1991). Noninvasive detection of changes in membrane potential in cultured neurons by light scat-tering. *Proc. Natl. Acad. Sci.* **88**, 9382-9386.

Department of Community Health and Center for Statistical Sciences, Brown University, 121 South Main Street, Providence, RI 02912, U.S.A.

E-mail: ombao@stat.brown.edu

Department of Statistics, University of Illinois at Urbana-Champaign, 104-A Illini Hall, 725 South Wright Street, Champaign, IL 61820, U.S.A.

E-mail: xshao@uiuc.edu

Department of Psychology and Department of Psychiatry and Behavioral Sciences, Stanford University, 780 Welch Road, Palo Alto, CA 94304, U.S.A.

E-mail: elenary@stanford.edu

Beckman Institute, University of Illinois at Urbana-Champaign, 405 N. Matthews Ave., Urbana, IL 61801, U.S.A.

E-mail: mfabiani@uiuc.edu

Beckman Institute, University of Illinois at Urbana-Champaign, 405 N. Matthews Ave., Urbana, IL 61801, U.S.A.

E-mail: grattong@uiuc.edu