IMS Collections
Vol. 0 (2009) 191–216

arXiv: math.PR/0000018

Nonparametric Estimation of Hemodynamic Response Function: A Frequency Domain Approach

Ping Bai[1], **Young Truong**[2] **and Xuemei Huang**[3]

University of North Carolina at Chapel Hill

Abstract: Hemodynamic response function (HRF) has played an important role in many recent functional magnetic resonance imaging (fMRI) based brain studies where the main focus is to investigate the relationship between stimuli and the neural activity. Standard statistical analysis of fMRI data usually calls for a "canonical" model of HRF, but it is uncertain how well this fits the actual HRF. The objective of this paper is to exploit the experimental designs by modeling the stimulus sequences using stochastic point processes. The identification of the stimulus-response relationship will be conducted in the frequency domain, which will be facilitated by fast Fourier transforms (FFT). The usefulness of this approach will be illustrated using both simulated and real human brain data. Under regularity conditions, it is shown that the estimated HRF possesses an asymptotic normal distribution.

Contents

[1]Department of Statistics and Operational Research, University of North Carolina, Chapel Hill, NC 27599-3260, email: pingbai@gmail.com
[2]Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599-3260, email: truong@bios.unc.edu
[3]Department of Neurology, University of North Carolina, Chapel Hill, NC 27599-3260, and Department of Neurology, Milton S. Hershey Medical Center H037, 500 University Drive, Penn State University, email: xhuang3@hmc.psu.edu

AMS 2000 subject classifications: Primary 60K35, 60K35; secondary 60K35
Keywords and phrases: fMRI, hemodynamic response function, transfer function model, point process, spectral analysis

1. Introduction

Consider a statistical problem in which data are acquired by applying stimuli at times $\tau_1 < \tau_2 < \cdots$ and simultaneously a varying response $Y(t)$ is recorded. Suppose it is desired to conduct the associated statistical inference based on the model:

$$Y(t) = \sum_j h(t - \tau_j) + \epsilon(t), \tag{1.1}$$

where $h(\cdot)$ is an unknown function and $\epsilon(t)$ is a stationary, zero mean, noise series with power spectrum given by $s_{\epsilon\epsilon}(\cdot)$. It is assumed that the function $h(t) = 0$ for $t < 0$ and will have finite duration.

This type of problems has played an important role in the fields of psychology, neurobiology, neurology, radiology, biomedical engineering, and many others, where data acquisition is carried out in functional magnetic resonance imaging (fMRI) experiments. As a noninvasive technique, fMRI allows us to study dynamic physiological processes at a time scale of seconds. The basis of fMRI is the Blood Oxygenation Level Dependent (BOLD) effect [40]. Due to differential magnetic susceptibility of oxygenated (oxygen-rich) hemoglobin and deoxygenated hemoglobin, the BOLD effect reflects the changes in hemodynamics which in turn yields greater MRI intensity when brain activity increases (see [30]). It is this hemodynamic response to the underlying neuronal activity that makes the fMRI signal in brain areas of activation a blurred and delayed version of the stimuli. Figure 1 shows the recorded BOLD signals (solid line) triggered by a single event (dashed line on the left panel) and a sequence of consecutive of stimuli (dashed line on the right panel) respectively. Both of them show the blur-and-delay effect caused by the hemodynamic response.

In practice, the BOLD effect is modeled through the convolution of the stimulus sequence $X(\cdot)$ and a hemodynamic response function (HRF) $h(\cdot)$ given by

$$\text{BOLD}(t) = h \otimes X(t) = \int h(t-u)X(u)du, \tag{1.2}$$

whose discrete time version is represented by model (1.1). Typically, an fMRI dataset consists of a 3D grid of voxels, each containing a time series of measurements that reflect brain activity. For each of roughly 200,000 voxels that lie inside the brain images, we wish to carry out the estimation of the HRF which will subsequently be applied to infer those voxels that were activated under certain experimental conditions.

The rest of this paper is organized as follows. Our estimate is described in Section 2, which is based on frequency domain methods applied to the point processes and ordinary time series. A brief survey of HRF modeling is provided in Section 3. Section 4 illustrates the performance of our proposed method through a simulated data analysis. For the purpose of comparison, a real fMRI data set is analyzed using the proposed method and a popular fMRI tool in Section 5. Discussions and concluding remarks are given in Section 6. Proofs are given in the last section of the paper.

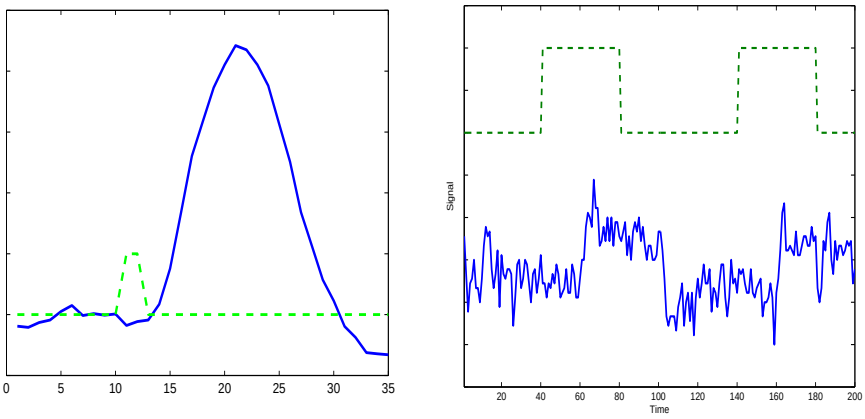

FIG 1. *Left Panel: The recorded BOLD signal (solid line) triggered by a single event (dashed line). Right Panel: The recorded BOLD signal (solid line) triggered by a typical block-design sequence (dashed line).*

2. A Frequency Domain Method for Estimating HRF

Model (1.1) has the structure of linear time invariant system carrying the stimuli $X(t)$ onto an response time series $Y(t)$. These models are generally studied by the frequency domain methods based on cross-spectral analysis (see [10]).

Define the *system transfer function* by

$$H(f) = \sum_u h(u) \exp(-iuf), \qquad f \in \mathbb{R}.$$

Define the finite Fourier transform of $Y(t)$ by

$$\varphi_0^T(f) \equiv \varphi_Y^T(f) = \sum_{t=0}^{T-1} \exp(-ift) Y(t)$$

with a similar definition for $\varphi_\epsilon^T(f)$, $f \in \mathbb{R}$. Also, define

$$\varphi_1^T(f) \equiv \varphi_X^T(f) = \sum_{t=0}^{T-1} \exp(-ift) X(t) = \sum_j \exp(-if\tau_j), \qquad f \in \mathbb{R},$$

the last sum is over the available stimuli before time $T-1$. It follows from (1.1) that

$$\varphi_0^T(f) = H(f)\varphi_1^T(f) + \varphi_\epsilon^T(f), \qquad f \in \mathbb{R}. \tag{2.1}$$

Now let m_f denote the integer $m \in \{0, 1, \ldots, T-1\}$ such that $2\pi m/T$ is closest to the (angular) frequency $f \in (0, \pi/2)$. Let K denote a positive integer. Then, for smooth $H(\cdot)$,

$$\begin{aligned} \varphi_0^T\left(\frac{2\pi}{T}(m_f + k)\right) &\approx H(f)\varphi_1^T\left(\frac{2\pi}{T}(m_f + k)\right) \\ &\quad + \varphi_\epsilon^T\left(\frac{2\pi}{T}(m_f + k)\right), \quad k = 0, \pm 1, \ldots, \pm K, \end{aligned} \tag{2.2}$$

hold for the $2K+1$ nearest frequencies around f of the form $2\pi(m_f+k)/T$. Thus a reasonable estimate of $H(f)$ can be obtained by regressing $\varphi_0^T(2\pi(m_f+k)/T)$ on $\varphi_1^T(2\pi(m_f+k)/T)$ for $k=0,\pm1,\ldots,\pm K$, which is given by

$$\hat{H}(f)=\hat{s}_{01}(f)/\hat{s}_{11}(f), \qquad f\in\mathbb{R}, \tag{2.3}$$

where

$$\hat{s}_{jj'}(f)=(2K+1)^{-1}\sum_{k=-K}^{K}\tilde{s}_{jj'}\left(\frac{2\pi}{T}(m_f+k)\right), \tag{2.4}$$

$$\tilde{s}_{jj'}(f)=(2\pi T)^{-1}\varphi_j^T(f)\overline{\varphi_{j'}^T(f)}, \qquad f\in\mathbb{R}, \quad j,j'\in\{0,1\}. \tag{2.5}$$

Here $\overline{a}$ is the conjugation of $a\in\mathbb{C}$. This is similar to the linear regression setting and therefore the residual sum of squares (RSS) is given by

$$\hat{s}_{\epsilon\epsilon}(f)=\frac{2K+1}{2K+1-1}\left(\hat{s}_{00}(f)-\hat{s}_{01}(f)\hat{s}_{11}^{-1}(f)\hat{s}_{10}(f)\right). \tag{2.6}$$

Note that

$$\hat{s}_{\epsilon\epsilon}(f)\propto\hat{s}_{00}(f)\Big(1-\frac{|\hat{s}_{01}(f)|^2}{\hat{s}_{11}(f)\hat{s}_{00}(f)}\Big)=\hat{s}_{00}(f)(1-R_{01}^2(f)),$$

where

$$|\hat{R}_{01}(f)|^2=\frac{|\hat{s}_{01}(f)|^2}{\hat{s}_{00}(f)\hat{s}_{11}(f)}, \qquad f\in\mathbb{R},$$

is the squared coherence, which lies between 0 and 1, the closer it is to 1 the stronger is the linear relationship between the two series.

Let $s_{\epsilon\epsilon}(\cdot)$ denote the power spectrum of the noise series. It can be shown that (see Section 7) the estimate $\hat{H}(f)$ is asymptotically complex normal with mean $H(f)$ and variance $s_{\epsilon\epsilon}(f)/K\hat{s}_{11}(f)$. And $\hat{H}(f_1)$, $\hat{H}(f_2)$, …, $\hat{H}(f_M)$ are asymptotically independent normal for distinct f_1, f_2, …, f_M [7].

In practice, we use a smoother estimate known as window estimate by observing that (2.4) can be written more generally as

$$\hat{s}_{jj'}(f)=\sum_{k\neq0}b^{-1}W\left(b^{-1}\left(f-\frac{2\pi k}{T}\right)\right)\tilde{s}_{jj'}\left(\frac{2\pi k}{T}\right), \qquad f\in\mathbb{R}, \tag{2.7}$$

where $W(\cdot)$ is a non-negative function called the weight or window function, and $b\equiv b_T\searrow0$ is the smoothing parameter. It has been shown that (2.7) has better sampling properties than (2.4) as an estimate of the cross-spectrum of the bivariate time series. See [10] and Section 7. From now on, our estimate of $H(f)$ will be based on (2.3) and (2.7).

We remark that when $j=j'$, then (2.5) becomes

$$\tilde{s}_{jj}(f)=(2\pi T)^{-1}\varphi_j^T(f)\overline{\varphi_j^T(f)}=(2\pi T)^{-1}|\varphi_j^T(f)|^2, \qquad f\in\mathbb{R}, \tag{2.8}$$

which is the *periodogram* of the series $Y(t)$ when $j=0$, or of the series $X(t)$ when $j=1$. The periodogram is an important statistic in spectral time series analysis.

Under certain conditions, $\hat{R}_{01}(f)$ is asymptotically normal with mean $R_{01}(f)$ and variance proportional to constant $(1-R_{01}^2(f))/(Tb)$. Moreover, if $R_{01}(f)=0$, then

$$F(f)=\frac{c|\hat{R}_{01}(f)|^2}{1-|\hat{R}_{01}(f)|^2}\sim F_{2,2c}, \tag{2.9}$$

where $c = (bT/\gamma) - 1$ and $\gamma = \int \lambda^2$ with λ being the lag-window generator [39].

This result can be used to test for a response to the stimulus by computing a test statistic for significant activation $F(f_a)$ at the fundamental frequency of activation f_a. Under the null hypothesis of no activation, the F-statistic at the fundamental frequency of activation, $F(f_a)$, has a F distribution with 2 and $2c$ degrees of freedom. Large values of $F(f_a)$ indicate a large effect at the fundamental frequency.

The estimate of the impulse response function $h(\cdot)$ is then given by

$$\hat{h}(u) = \frac{1}{Q}\sum_{q=0}^{Q} \hat{H}\left(\frac{2\pi q}{Q}\right) \exp\left(i\frac{2\pi uq}{Q}\right),$$

where $Q \equiv Q_T$ denote a sequence of positive integers tending to ∞ with T. Under certain conditions, $(\hat{h}(u_1), \ldots, \hat{h}(u_J))$ is asymptotically normal with mean $(h(u_1), \ldots, h(u_J))$ and covariance matrix

$$\frac{2\pi}{bT}\int W(\lambda)^2 d\lambda \cdot \frac{1}{Q^2}\int \exp\left(i(u_j - u_k)\lambda\right)\frac{s_{\epsilon\epsilon}(\lambda)}{s_{11}(\lambda)}d\lambda, \qquad j,k = 1,2,\ldots J.$$

See Section 7 for more details.

3. A Brief Survey of HRF Modeling

The basis of model (1.1) is the linearity of BOLD fMRI responses when multiple stimuli are presented in succession. This was first studied by Boynton, et al. [6]. The linearity arises from the fact that a stimulus induces the neural activity in a specific region of the brain. This then brings blood flow changes (hemodynamics) in that region, while BOLD fMRI responses are measured from these blood flow changes. In addition to giving this clear picture of how BOLD fMRI works, the linear transform model is important in two respects. Firstly, the assumption of linearity of the fMRI response and neural activity makes it possible to determine changes in neural activity by the amplitude changes in hemodynamic response. Secondly, this linear transform model also shows that when multiple stimuli are presented in succession, the hemodynamic response would be the summation of the individual responses generated by the single stimulus respectively.

Modeling the relationship between the fMRI response and stimuli is a key step towards detecting fMRI activity. Standard statistical analysis is carried out based on the following model:

$$Y(t) = \beta\sum_j h(t - \tau_j) + \epsilon(t), \tag{3.1}$$

where the HRF $h(\cdot)$ is pre-specified and β is a voxel specific parameter, to be utilized for testing fMRI activity [15, 21, 23, 24, 31, 33, 43]. The assumptions made about the shape of the HRF vary among different methods. Some of them are very stringent, while others are relatively more flexible. Typically, a "canonical" HRF is employed to process fMRI data. Some studies have reported variation in the shape of HRF across subjects [2, 4], and within the same subject across regions [11, 32, 38].

Detecting fMRI activity has also been evolved from using block-designs (where the stimulus times τ_i occur consecutively to form a block) to event-related fMRI

(ER-fMRI) designs [41]. In the latter case, stimuli (or events) are applied for short bursts in a stochastic manner. The recorded BOLD fMRI signals measure transient changes in brain activity associated with discrete events. This feature makes ER-fMRI a useful tool to estimate the change in the MR signal triggered by neuronal activity.

As an early study of ER-fMRI, Dale and Buckner [16] correlated the selective averaging (or time averaging) data and the fixed HRF induced data in a voxelwise manner. Friston et al. [24] employed a Poisson function with a fixed parameter for the HRF. In the general linear model (GLM) framework, Friston et al. [22] estimated the HRF by two given temporal basis functions. To enhance its flexibility, this idea was extended by Josephs et al. [31] to include more basis functions. These are very important contributions since the basis sets allow one to estimate an HRF of arbitrary shape for different events at different voxels of the brain, and at the same time the inferences can be easily made. Many studies on modeling HRF have since focused on the refinement and improvement of the basis sets idea. For example, Woolrich et al. [44] introduced a technique by applying some constraints to avoid nonsensical HRF, which is a big problem when using simple basis functions. More recently, Lindquist and Wager [34] proposed another method, using three superimposed inverse logistic functions, to model the HRF. This paper also described some most popular HRF modeling techniques, such as smooth finite impulse response (FIR) filter [29], canonical HRF with time and dispersion derivatives [14] and the canonical SPM HRF [25]. A flexible method based on splines has been considered by Zhang et al. [47].

From a Bayesian perspective, Genovese [26] and Gössl et al. [28] proposed to model the HRF by a number of parameters and prior distributions are given to each parameter. See also Woolrich et al. [44] and Lindquist and Wager [34]. Inferences of the parameters were then made at each voxel using Markov Chain Monte Carlo (MCMC) technique. The disadvantage of these methods is the slow performance of general MCMC techniques for the inferences.

The above methods are referred to as the time-domain methods. We now consider the frequency-domain approach. Assuming a periodic stimulus design, fMRI time series analysis can be greatly simplified in the frequency domain, which is more natural as the problem of modeling the relationship between the response and the stimuli is reduced to a few parameters related to the stimulus frequency information. One of the first frequency domain approaches is given by Lang and Zeger [33], who used model (3.1) along with a two-parameter gamma function to model the HRF. The two parameters vary at different voxels and hence the estimated HRF varies from voxel to voxel. It was reported that this approach has an identifiability problem of the parameters. The issue was addressed in Marchini and Ripley [37] using a fixed HRF approach.

We remark that a common theme among the time-domain methods for testing activity is the two-step procedure: the extra step is required for modeling parametrically the temporal noise series. This will affect the power of the test. While the above frequency approaches avoided the noise modeling part, they lack the ability to address: (1) the varying HRF issue, and (2) different types of stimulus designs. Moreover, the linear transformed model of fMRI response has not been tested and some studies [16, 30] reported the presence of non-linearity.

In Section 2, we described a regression approach based on model (1.1) for addressing these problems. The application of point processes to model the stimuli is novel in the HRF modeling literature. The procedure is also greatly simplified. Namely, it enables us to estimate the HRF directly and simultaneously test the lin-

earity assumption in a single step. The idea of using point processes can be traced back to the work of Brillinger [8, 9] who applied this to identification problems emerged in neurophysiological neural spike train analysis.

4. Simulated Numerical Results for HRF estimation

We illustrate the usefulness of our procedure in a simulated study. Here we use one of the HRF's from the literature [16, 27] to generate the response. The stimulus sequence contains a brief trial (1 second) intermixed with events "on" or "off" (Figure 2a). In this experiment, each trial lasts for 18 seconds and there are sixteen runs. So the "average" (because of the random "on" or "off") frequency of the event is $18/288 = 0.0625$. The estimated power spectrum (Figure 2b) and the frequency of the event are precisely estimated. The second peak corresponds to the frequency of every other event. In the first experiment (Figure 3a-b), the response is given by $Y(t) = a \int h(t-u)X(u)du + \epsilon(t)$ with $a = 0.5$ and the noise ϵ is generated from an AR(1) with coefficient 0.7: $\epsilon(t) = 0.7\epsilon(t-1) + z(t)$, $z(t) \sim N(0, .3^2)$. In the second experiment (Figure 3c-d), the noise is generated from an ARMA(2,2): $\epsilon(t) - 0.8897\epsilon(t-1) + 0.4858\epsilon(t-2) = z(t) - 0.2279z(t-1) + 0.2488z(t-2)$, $z(t) \sim N(0, .3^2)$. The AR(1) time series model was chosen to represent the default settings in Statistical Paramtric Mapping (SPM) [21] and FMRIB Software Library (FSL) [43], while the ARMA case was mainly for testing teh strengths of our method under other types of correlated structures. The coefficients were selected to illustrate the performance of the procedure under moderate serially correlated noise. Large coherency at the stimulus frequency of 0.0625 indicates that the activation is strong, and there is some linearity in the response and the stimulus series. This is also confirmed by the highly significant F-statistics (Figures 4a-d). The significant level is a dashed line that sits near the bottom of the graph. The variability of the proposed estimated is illustrated in Figures 5a-c using various noise level with SD=0.3, 0.5. We remark that the number of runs (=16) used in these simulations is based on recently published articles.

We further apply the procedure to examine the main application of fMRI to detect regions of activation. These are illustrated in Figures 6, 7. In these experiments, the responses are generated from

$$Y(t) = a \int h(t-u)X(u)\, du + \epsilon(t),$$

with varying a to show contrast of the regions so that the sub-region has a higher value of a. The noise component is ARMA(2,2), the same as in the previous experiment with SD=0.3. The regions of activation are clearly captured (Figure 6) when the contrast ratio is high. The effect of the contrast ratio on the detection of region of activation is depicted in Figures 7. It is evident that the level of detection depends on the contrast ratio.

5. A Real Data Analysis

5.1. Experiment Paradigm and Data Description

In this study, an fMRI data set was obtained from one human subject performing a predefined event sequence as visually instructed. The stimulus sequence includes two different events: right-hand and left-hand finger tapping. Each finger tapping

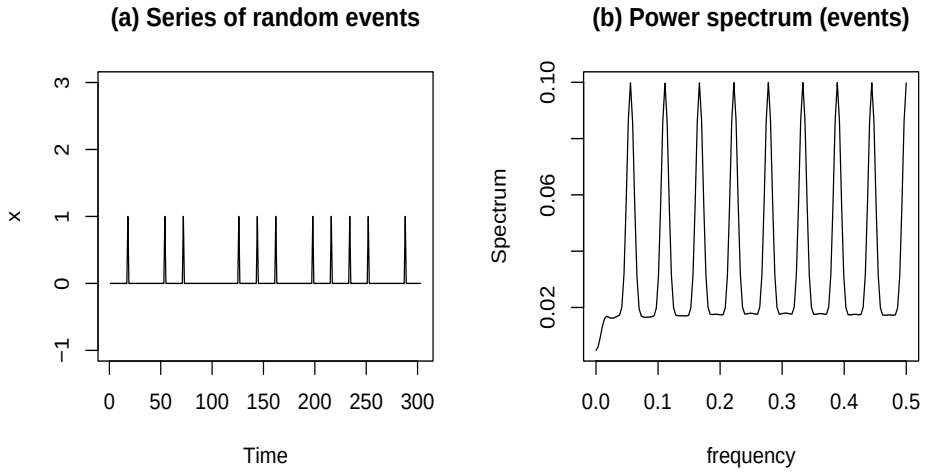

FIG 2. Plot of the stimulus series (16 trials) with on-off pattern stimuli (every 18 seconds). The whole duration lasts about $16 \times 18 = 288$ seconds. (b) The estimated power spectrum provides frequency of the occurred events. The frequency associated with the first peak is about $18/288 = 0.0625$. The second peak gives frequency of every other event, etc.

movement lasted around 1 second. The order of the sequence was predefined in a random way. To avoid the overlapping of consecutive events, the time interval between two successive events was randomly selected from Uniform$[18, 22]$. A typical sequence of stimuli is $\{R, L, R, R, L, L\}$.

During the experiment, 47 MR scans were acquired on a modified 3T Siemens MAGNETOM Vision system. Each acquisition consisted of 49 contiguous slices. Each slice contained 64×64 voxels. Hence there were $64 \times 64 \times 49$ voxels from each scan. The size of each voxel is 3mm $\times$ 3mm $\times$ 3mm. Each acquisition took 2.9388 seconds, with the scan to scan repetition time (TR) set to be 3 seconds.

5.2. Analysis and Results

The data set was preprocessed using SPM5 [21]. The preprocessing included realignment, slice timing correction, coregistration and spatial smoothing. We then analyzed the processed data set using both our proposed method and SPM5. When using SPM5, we used a canonical HRF with time and dispersion derivatives to model the hemodynamic response [25] and its functional form is shown in Figure 5. A t-statistic map was generated to show the activations triggered by the stimuli and part of them is shown on the first row of Figure 8.

When using the proposed method to detect which regions of the brain were activated by the finger tapping movements, we generated a spatial color map of the p-value for each voxel. The p-values were calculated based on the test defined by (2.9). Thus the activation regions are identified by the F statistics that are significant. The p-map generated this way is shown on the second row in Figure 8.

The four image slices represent the spatial maps of the right-hand activation. The red areas illustrate activated brain regions. Brighter color indicates higher intensity. Our p-maps demonstrate the classic brain activation patterns during hand movement as described above. However, the t-maps of the same four slices generated using SPM5 do not show any activation, as seen from the first row of Figure 8.

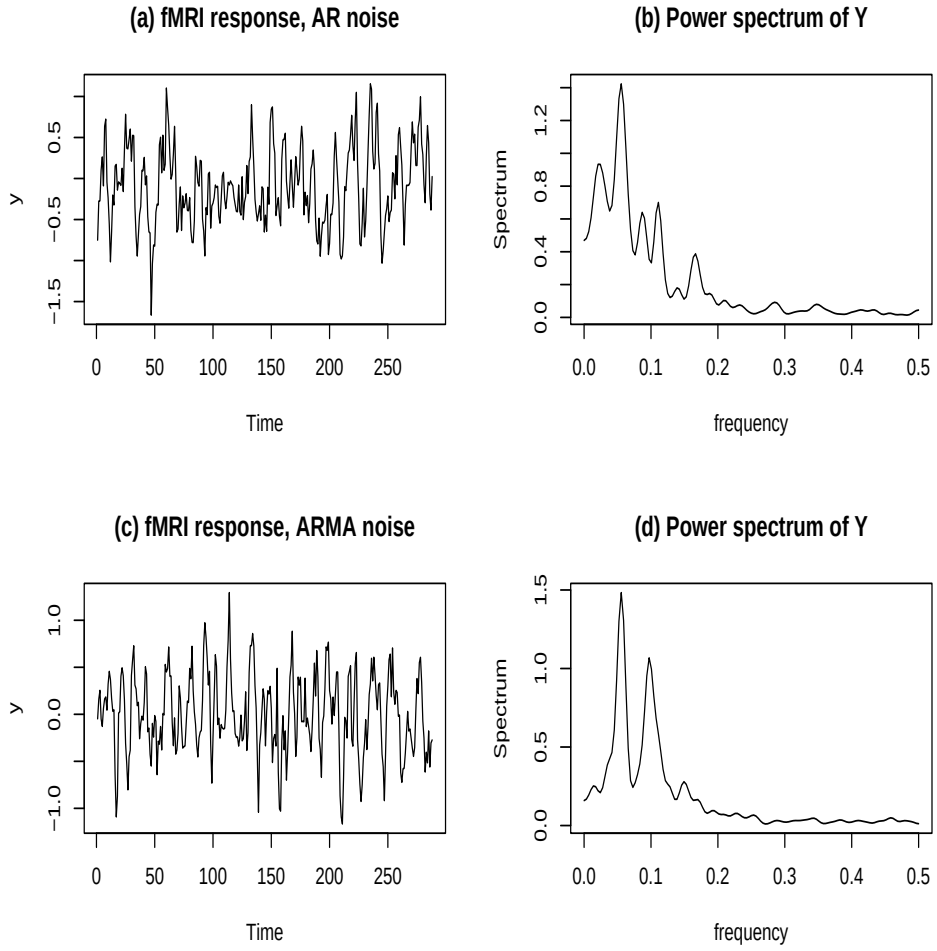

FIG 3. The fMRI responses. In these experiments, the responses are generated from $Y(t) = (0.5)\int h(t-u)X(u)\,du + \epsilon(t)$ with $\epsilon(t) = 0.7\epsilon(t-1) + z(t)$, $z \sim N(0, .3^2)$ in (a) and (b); $\epsilon(t) - 0.8897\epsilon(t-1) + 0.4858\epsilon(t-2) = z(t) - 0.2279z(t-1) + 0.2488z(t-2)$, $z \sim N(0, .3^2)$ in (c) and (d). The stimuli are the same as in Fig. 2.

Next we plot the estimated HRFs at voxels which are shown to be activated according to Figure 8. Figure 9 displays the estimated HRFs for voxels (with $F > 20$) selected from primary motor cortex (PMC). Figure 10 displays the estimated HRFs in cerebellum, and Figure 11 shows the estimated HRFs in the supplementary motor area (SMA). These figures were obtained by first computing the F-statistics (2.9) followed with the selection of those voxels with the F values greater than 20. This threshold was chosen to adjust for the multiple comparison effect and was carried out by computing the F statistics over those voxels that are known to be not activated by the stimuli. For example, we used the WFU PickAtlas software [36] to generate region of interest (ROI) mask in the cerebro spinal fluid area of the brain. Then computed the F statistics over this area, followed with a density estimate (eg, kernel method, or simply histogram) to select the thresholding value. There are about 20,000 voxels in the cerebro spinal fluid area which can be used to calibrate the null distribution for detecting fMRI activity.

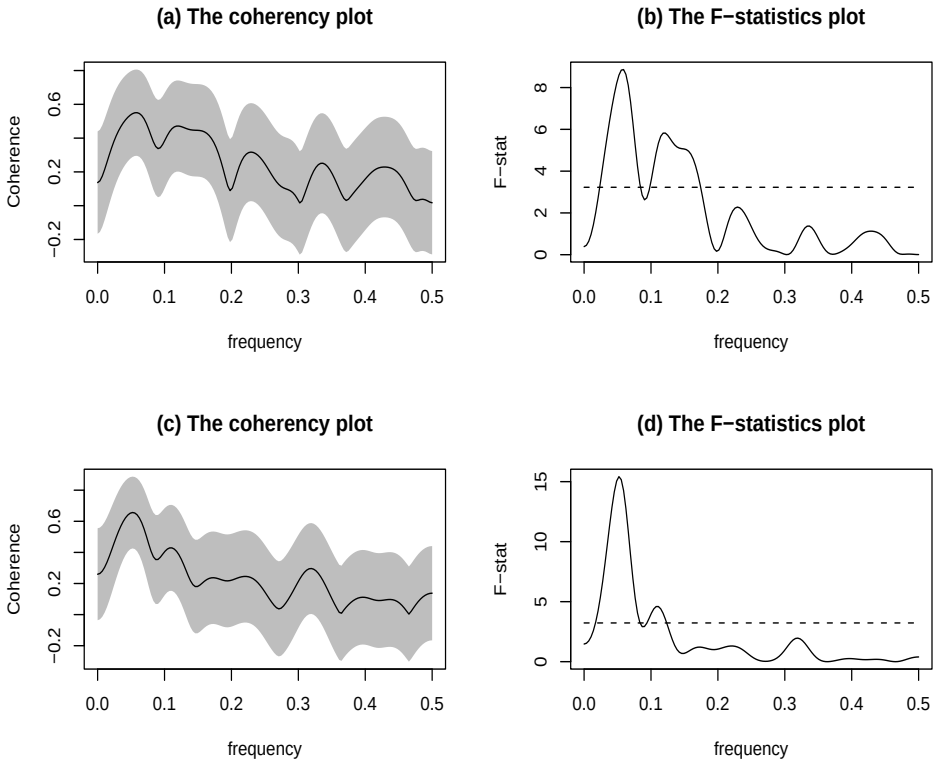

FIG 4. (a) and (c): The estimated coherency function with pointwise confidence intervals. Large coherence values at the event frequencies indicate perfect linear time invariant system used in this simulation. (b) and (d): The F-test for coherency with the dashed-line showing the significant level. (a) and (b) have the same conditions as (a) and (b) of Fig. 3; similarly for (c) and (d).

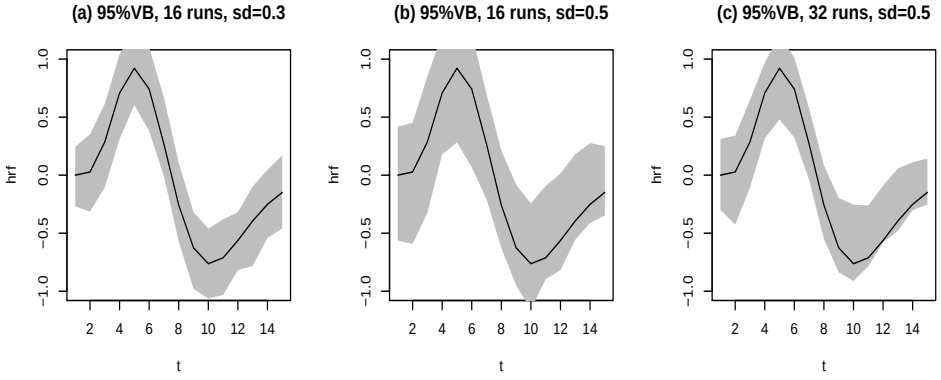

FIG 5. (a) 95% variation bands (VB) are obtained from estimates of the HRF using 100 random samples simulated from the model by trimming off the 2.5% both ends. (b) Same as in (a) except the noise SD has increased to 0.5. This clearly indicates the variance of the proposed estimate depends on the variance of the noise. (c) Same as in (b) by doubling the number of runs. This illustrates the variance of the estimate is inversely proportional to the duration of the experiment. In these experiments, the responses are generated from $Y(t) = \int h(t-u)X(u)\,du + \epsilon(t)$ with $\epsilon \sim \text{ARMA}(2,2)$. The parameters are the same as in Fig. 3.

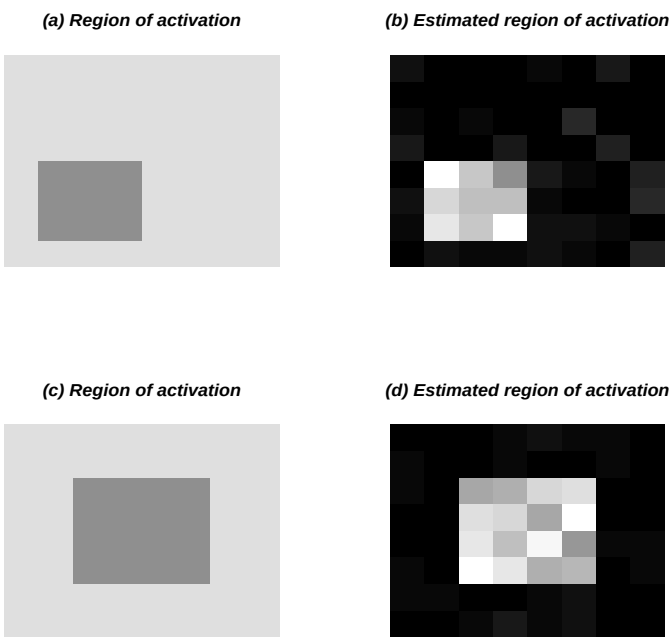

FIG 6. Estimated region of activation. In these experiments, the responses are generated from $Y(t) = a \int h(t-u)X(u)\,du + \epsilon(t)$ with $a = 0.1$ outside and $a = 1.0$ inside the sub-region, respectively. There are two slices with different region sizes and locations. Each slice is 8×8. (a) and (c) are the true regions, (b) and (d) are the estimated regions. The noise ϵ is generated from an ARMA(2,2). The stimuli are the same as in Fig. 2.

All of them have the form that agrees with empirical experience. It is well established that the contralateral cerebral hemisphere motor areas such as primary motor cortex (PMC), and ipsilateral cerebellar areas play dominant role in motor functions in normal human subjects [1, 3, 17, 19]. Our new methods validate unequivocally this known motor activation pattern with single finger movement in a single subject, whereas traditional SPM5 failed to do. Adequate imaging analysis techniques to demonstrate the involvement of those structures during motor function is superior important. PMC is the primary brain region directly control human movement, basal ganglia and cerebellar modulate its functions through a number of cortical motor associated areas of the brain (such as SMA). Dysfunctions of these structures have known to cause a variety of movement disorders such as Parkinson's disease and cerebellar ataxia [1]. Our methods might provide "higher resolution" statistical analysis methods for clinical and neuroscientist to define the roles of these structures in disease stages using fMRI.

5.3. Implications

(1) We demonstrated that our method handles statistical issues related to event-related experiments well. (2) It is nonparametric in the sense that the functional form of HRF is not specified a priori. Hence it is an useful diagnostic tool for other approaches that may be biased because of misspecification of HRF. (3) Variation of HRF in the brain has been under active research [18], and the nonparametric

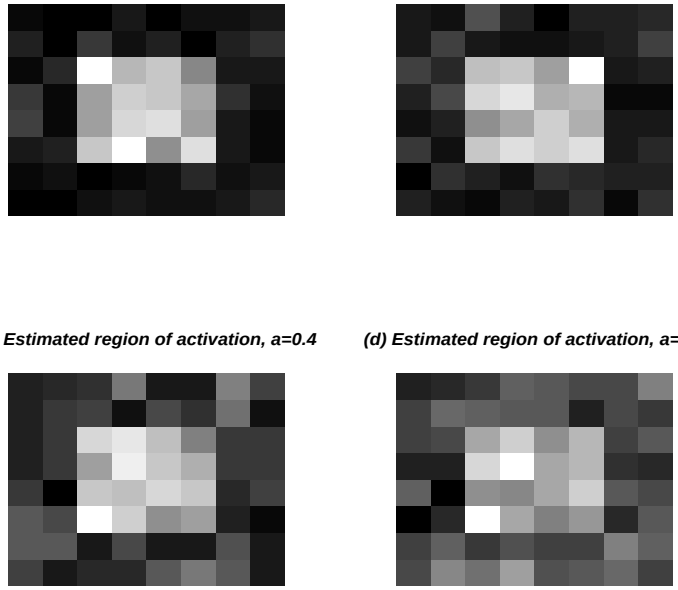

FIG 7. Estimated region of activation with varying contrast ratios. In these experiments, the responses are generated from $Y(t) = a\int h(t-u)X(u)\,du + \epsilon(t)$ with $a = 0.2, 0.3, 0.4, 0.5$ outside the sub-region and $a = 1.0$ inside the sub-region. The noise ϵ is generated from an ARMA(2,2). These results illustrate that accuracy of the estimates depends on the signal-to-noise (or contrast) ratio: The contrast ratio is proportional to $1/a$. (a) Here $a = 0.2$ implies that the signal is weaker than that in Fig. 6, but the contrast is still high and so the estimated region can still be clearly identified. (b) The contrast here is weaker with $a = 0.3$. (c) Weaker contrast with $a = 0.4$, and (d) fuzzy region due to the weakest contrast used in this experiment.

approach offers a systematic way to study the variation without requiring HRF to have the same shape over all voxels. (4) The linear relationship specified through the BOLD signal can be examined statistically by carrying out a formal test of hypothesis. This is important in verifying the linearity assumption employed in SPM [20, 23, 24, 45] in the process of constructing the T-map. (5) It is relatively easy to interpret the results using our approach as no prior specification of HRF is required (as is done in SPM [21]/FSL [43]/AFNI [15]).

5.4. Discussions

There are many ways to detect fMRI activity. The critical problem is to estimate the statistical significance, which depends on the estimation of both the magnitude of the response to the stimulus and the serial dependence of the time series and especially on the assumptions made in that estimation. Nonparametric spectral density estimation is shown to be self-calibrating and accurate when compared to several other time-domain approaches [12, 13], SPM: [20, 23, 24, 45, 46]. In particular, spectral technique to detect periodic and event-related activations has a distribution theory with significance levels down to 1 in 100,000, levels which are needed when a whole brain image is under consideration. The technique is especially resistant to high frequency artifacts that are found in some datasets and

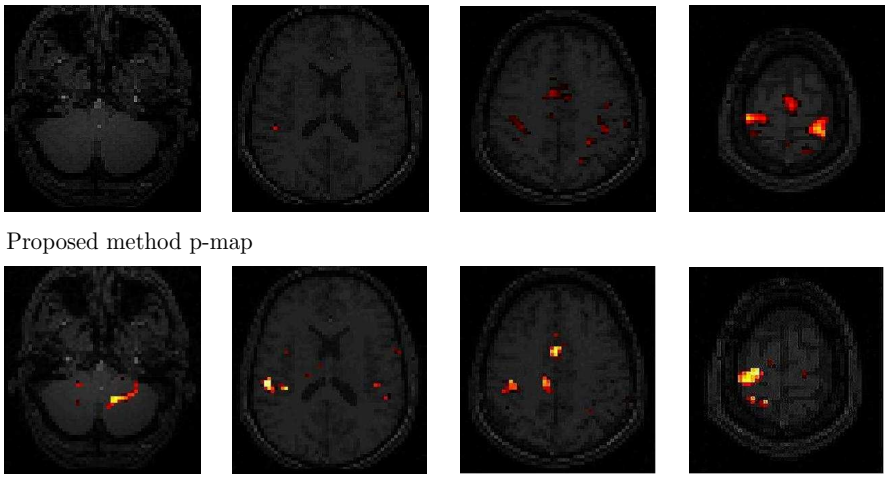

FIG 8. *The four related slices that contain the areas activated by right-hand finger tapping. The first row consists of the t-maps generated by SPM5 and they do not show any activation. The second row contains the p-maps generated by the proposed method. The first slice indicates the activated areas in cerebellum. The second slice contains basal ganglia. The third slice contains supplementary motor area (SMA) and the fourth slice shows primary motor cortex (PMC).*

it was demonstrated that time-domain approaches may be sufficiently susceptible to these effects to give misleading results. Also, these techniques are capable for detecting activations in clumps of a few (even one) voxel in periodic designs, yet produce essentially no false positive detections at any voxels in null datasets [37].

6. Concluding Remarks

It is now widely accepted that fMRI modeling requires flexible HRF modeling, with the HRF varying spatially and between subjects. Flexibility in linear modeling has been introduced with the use of basis functions [22]. However, basis functions suffer from a number of limitations. They impose a hard constraint on the allowed HRF shape and often the extent of the constraint is difficult to control and/or interpret. To overcome these problems, we formulated a procedure based on model (1.1) and FFT. The usefulness has been demonstrated empirically.

We remark that time-domain methods such as SPM [21], FSL [43], FIR [34, 35] and local adaptive spline estimate [47] in modeling the HRF are generally very sensitive to the linearity assumption and the error structures they employ. Any approach proposed within the time-domain may have difficulty providing resistant estimates. There is also no guarantee that the parametric noise model chosen will be sufficiently flexible to capture the true form of the correlation structure even if artifacts are removed and a model selection procedure is employed [18, 37]. Therefore significant loss in statistical efficiency can occur if these assumptions are invalidated. In contrast, if these assumptions are valid then the use of a frequency approach will result in a comparatively small loss in efficiency [10]. When considering voxel time series from fMRI datasets there can be no guarantees that the correct time domain

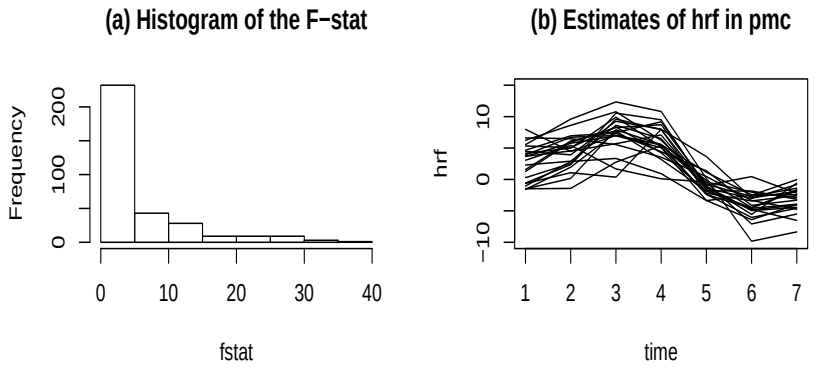

FIG 9. *The histogram for the F-stat and the HRF estimates in the primary motor cortex (PMC) area with $F > 20$. Each finger-tapping task lasted around 1 second. The order of the sequence was predefined in a random way. The time interval between two successive events was randomly selected from Uniform (18,22). Each acquisition took 2.9388 seconds, with the scan to scan repetition time (TR) set to 3 seconds.*

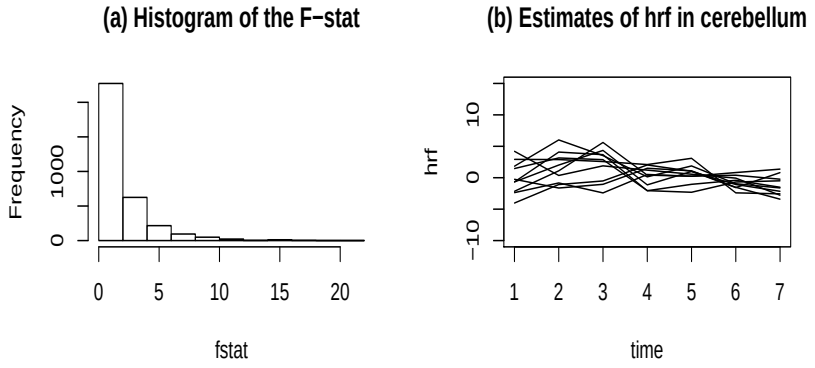

FIG 10. *The histogram for the F-stat and the HRF estimates in the cerebellum area with $F > 15$.*

approach has been chosen and a frequency approach seems the most prudent in this context. It is often demonstrated that the assumptions of commonly proposed time-domain models are not resistant to high frequency artifacts.

It is generally believed that the direct analysis of nonperiodic designs will not be as simple as that of the periodic designs, since the response due to the stimulus will be spread over a range of frequencies. Marchini and Ripley [37] suggested that this may be addressed by combining their method with the iteratively reweighted least squares [33] in the spectral domain and the basis functions [22]. However, this method will not be easily extended to model the HRF discussed in this paper. By formulating the problem using point processes, the frequency method advocated by [37] can be easily generalized to handle event-related designs. We also observe that our method is applicable to block designs since the stimuli can be put next to each other to form a block. Thus this unified approach significantly improves the estimation of the HRF described in [33, 37].

The flexible frequency approach proposed here acts as an insurance policy against the results being badly affected by artifacts, and is guaranteed to be near-optimal under all realistic operational conditions. It also offers a quick and accurate way to

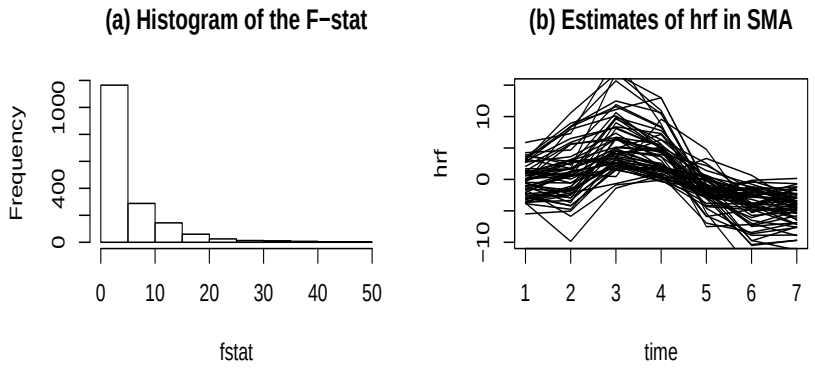

FIG 11. *The histogram for the F-stat and the HRF estimates in the cerebellum area with* $F > 15$.

check the calibration of the procedure. Further investigations will be carried out for an extensive comparative study on these maps. A concern about our procedure is the choice of weight function $W(\cdot)$ and bandwidth b given in (2.7). The former is less crucial and it can be addressed by choosing one of commonly used weight functions described in Newton [39]. Bandwidth selection appears to be more serious and it would seem to require adaptive methods such as cross-validation. Based on our experience and the fact that the HRF (blood flow) is relatively smooth, the choice of bandwidth therefore plays a less significant role. Nevertheless, we do observe that the spectral properties of the stimuli can be closely examined by the designs of the experimental protocols, which to some extent can help determine the smoothness of the estimate of HRF. This project is currently underway along with the use of splines for estimating the spectra.

7. Sampling Properties of the Estimates

The sampling properties of the HRF estimate will rely on the spectral properties of the stimulus $X(t)$, which is a point process. They also depend on the spectral properties of the response $Y(t)$ and the noise series $\epsilon(t)$, which are real-valued ordinary stationary time series. Thus this section starts with a brief summary of the spectral properties of stationary time series which will be denoted by $X(t)$. This is followed by a discussion of the cumulants, which are essential for establishing asymptotic distributions of the estimates. Subsequent sections describe sampling properties of various statistics involved in establishing the properties of HRF estimate.

7.1. Point Process

Consider a point process $X(t)$ with points occurring at times $0 \le \tau_1 \le \tau_2 \le \cdots$ with $X(t)$ denoting the number of points in the interval $(0, t]$. When it exists, the rate of the process at time t is given by

$$p_X(t) = \lim_{v \downarrow 0} \frac{1}{v} \mathsf{E}\big(X(t+v) - X(t)\big).$$

The expected number of points in the small interval $(t, t+v]$ is given by $p_X(t)v + o(v)$. Suppose orderliness, that is, the points of its realizations are isolated, multiple

points do not occur. Then $dX(t) = 0$ or 1 and one must have

$$\mathsf{P}\{dX(t) = 1\} = p_X(t)\,dt.$$

The rate function $p_X(t)$ is seen to have an interpretation as a probability.

In the second-order case one defines the second-order product density as

$$p_{XX}(t_1, t_2) = \lim_{v_1, v_2 \downarrow 0} \frac{1}{v_1 v_2} \mathsf{E}\big(X(t_1 + v_1) - X(t_1)\big)\big(X(t_2 + v_2) - X(t_2)\big), \quad t_1 \neq t_2.$$

In view of the orderliness of the process, $\mathsf{P}\{dX(t) = 1 \text{ and } dX(t) = 1\} = \mathsf{P}\{dX(t) = 1\}$, the case $t_1 = t_2$ can be included via

$$\begin{aligned}\mathsf{P}\{dX(t_1) = 1 \text{ and } dX(t_2) = 1\} &= \mathsf{E}\big(dX(t_1)dX(t_2)\big) \\ &= \big(p_{XX}(t_1, t_2) + \delta(t_1 - t_2)p_X(t_2)\big)\,dt_1\,dt_2,\end{aligned}$$

where $\delta(\cdot)$ is the Dirac delta function: $\delta(\cdot) \geq 0$ and $\int \delta(t)\varphi(t)\,dt = \varphi(0)$ for infinitely differentiable real-valued function φ with compact support. It is useful to recall here that δ is the (generalized) derivative of the Heaviside function: $H(\cdot) = 1_{(0,\infty)}(\cdot)$. See [42].

The covariance density of the process is defined by

$$q_{XX}(t_1, t_2) = p_{XX}(t_1, t_2) - p_X(t_1)p_X(t_2),$$

with the interpretation

$$\operatorname{cov}\{dX(t_1), dX(t_2)\} = \big(q_{XX}(t_1, t_2) + \delta(t_2 - t_1)p_X(t_2)\big)\,dt_1\,dt_2.$$

The conditional intensity of the process is defined by $p_{XX}(t)/p_X(t)$ with the interpretation

$$\mathsf{P}\{dX(t_2) = 1 \mid dX(t_1) = 1\} = \big(p_{XX}(t_1, t_2)/p_X(t_1)\big)\,dt_2.$$

A point process is said to be stationary when its probability properties are unaffected by simple shifts of time. In this case one has

$$\mathsf{P}\{dX(t) = 1\} = p_X\,dt,$$

$$\mathsf{P}\{dX(t_1) = 1 \text{ and } dX(t_2) = 1\} = \big(p_{XX}(t_2 - t_1) + \delta(t_2 - t_1)p_X\big)\,dt_1\,dt_2$$

$$\operatorname{cov}\{dX(t_1), dX(t_2)\} = \big(q_{XX}(t_2 - t_1) + \delta(t_2 - t_1)p_X\big)\,dt_1\,dt_2.$$

By analogy with what is done in the ordinary time series case one may define the power spectrum at frequency f by

$$s_{XX}(f) = \frac{1}{2\pi} \int e^{-ifu}\big(\operatorname{cov}\{dX(t + u), dX(t)\}/dt\big)\,du.$$

For multivariate process $X(t) = \{X_1(t), \ldots, X_m(t)\}$, it may be convenient to consider

$$\mathsf{P}\{dX_j(t) = 1\} = C_j\,dt, \qquad j = 1, \ldots, m,$$

and

$$\operatorname{cov}\{dX_j(t + u), dX_k(t)\} = C_{jk}(du)\,dt, \qquad j, k = 1, \ldots, m.$$

The power spectrum at frequency f is defined by

$$s_{jk}(f) = \frac{1}{2\pi} \int e^{-ifu} C_{jk}(du), \qquad j, k = 1, \ldots, m.$$

7.2. Stationary Time Series

Let $X(t) = (X_1(t), \ldots, X_r(t))$, $t = 0, 1, 2, \ldots$, denote a vector-valued stationary time series. Set

$$C_{jk}(u) = \operatorname{cov}\{X_j(t+u), X_k(t)\}, \qquad j, k = 1, \ldots, r.$$

The power spectrum at frequency f is defined by

$$s_{jk}(f) = \frac{1}{2\pi} \sum_u e^{-ifu} C_{jk}(u), \qquad f \in \mathbb{R}, \quad j, k = 1, \ldots, r.$$

7.3. Cumulants and Spectra

Definition 1. *Let $X_1, X_2, \ldots, X_r$ denote random variables with finite rth moment. The rth order joint cumulant of $X_1, X_2, \ldots, X_r$ is defined by*

$$\operatorname{cum}(X_1, X_2, \ldots, X_r) = \sum (-1)^{p-1}(p-1)! \left(\prod_{j \in \nu_1} X_j \right) \ldots \left(\prod_{j \in \nu_p} X_j \right),$$

where the summation extends over all partitions $\nu_1, \ldots, \nu_p$, $p = 1, \ldots, r$ of $\{1, 2, \ldots, r\}$.

Remarks

1. When $X_1 = X_2 = \cdots = X_r$, the definition gives the cumulant of order r of a univariate random variable.
2. $\operatorname{cum}(X_1, X_2, \ldots, X_r)$ is also given by the coefficient of $(i)^r t_1 \ldots t_r$ in the Taylor series expansion of $\log(E \exp i \sum_1^r X_j t_j)$.

Given r time series $X_1(t), X_2(t), \ldots, X_r(t)$ with each having finite rth moment, we define

$$C_{1,\ldots,r}(t_1, t_2, \ldots, t_r) = \operatorname{cum}(X_1(t_1), X_2(t_2), \ldots, X_r(t_r)).$$

For stationary time series,

$$C_{1,\ldots,r}(t_1, t_2, \ldots, t_r) = C_{1,\ldots,r}(t_1 - t_r, t_2 - t_r, \ldots, t_{r-1} - t_r, 0),$$

which is a function of $r-1$ variables. In this case, the rth order cumulant spectrum, $s_{1,\ldots,r}(f_1, f_2, \ldots, f_{r-1})$, is defined by

$$s_{1,\ldots,r}(f_1, f_2, \ldots, f_{r-1}) =$$

$$(2\pi)^{-k+1} \sum_{u_1, u_2, \ldots, u_{r-1}} C_{1,\ldots,r}(u_1, u_2, \ldots, u_{r-1}) \exp\left(-i \sum_{j=1}^{r-1} u_j f_j \right),$$

$$f_1, f_2, \ldots, f_{r-1} \in \mathbb{R}, \quad r \geq 2.$$

For a more detailed discussion of cumulants and their spectra, see [10].

7.4. Fast Fourier Transforms

Let $a_j(\cdot) : \mathbb{R} \to \mathbb{R}$, $j = 1, 2$, denote **tapering functions**. The discrete Fourier transform for the univariate series X_j is defined by

$$\varphi_j^T(f) \equiv \varphi_{X_j}^T(f) = \sum_t a_j(t/T) X_j(t) \exp(-ift), \qquad f \in \mathbb{R}, \quad j = 1, 2.$$

For vector-valued series $\mathbf{X}$, it is given by

$$\boldsymbol{\varphi}^T(f) \equiv \boldsymbol{\varphi}_{\mathbf{X}}^T(f) = \begin{pmatrix} \varphi_1^T(f) \\ \varphi_2^T(f) \end{pmatrix}, \qquad d \in \mathbb{R}.$$

Set $a_j^T(t) = a_j(t/T)$, $j = 1, 2$. For $j_m \in \{1, 2\}$, $m = 1, \ldots, M$,

$$A_{j_1,\ldots,j_M}^T(f) = \sum_t \left(\prod_{m=1}^M a_{j_m}^T(t) \right) \exp(-ift), \quad f \in \mathbb{R}.$$

Condition 1. *The tapering function $a(\cdot) : \mathbb{R} \to \mathbb{R}$ has a compact support with bounded first derivative. Furthermore,*

$$\int a(u)\, du = 1 \qquad \textit{and} \qquad \int |a(u)|\, du < \infty.$$

Condition 2. *The covariance function satisfies*

$$\sum_u C_{jk}(u) < \infty,$$

and

$$\sum_{u_1,\ldots,u_{M-1}} C_{j_1 \ldots j_M}(u_1, \ldots, u_{M-1}) < \infty, \qquad j_1, \ldots, j_M = 1, 2.$$

The second part of the above condition is necessary for establishing the asymptotic properties of the estimates to be considered in this section.

Lemma 1. *Suppose Conditions 1 and 2 hold. Then*

$$\sup_{f_1,\ldots,f_M} \Big| \mathrm{cum}(\varphi_{j_1}^T(f_1), \ldots, \varphi_{j_M}^T(f_M)) - (2\pi)^{M-1} A_{j_1,\ldots,j_M}^T(f_1 + \cdots + f_M)\, s_{j_1,\ldots,j_M}(f_1, \ldots, f_M) \Big| = o(T).$$

Condition 3. *The covariance function satisfies*

$$\sum_u |u| c_{jk}(u) < \infty,$$

and

$$\sum_{u_1,\ldots,u_{M-1}} |u_1 \cdots u_{M-1}| C_{j_1 \ldots j_M}(u_1, \ldots, u_{M-1}) < \infty, \qquad j_1, \ldots, j_M = 1, 2.$$

Lemma 2. *Under Conditions 1 and 3,*

$$\sup_{f_1,\ldots,f_M} \Big| \mathrm{cum}(\varphi_{j_1}^T(f_1), \ldots, \varphi_{j_M}^T(f_M)) - (2\pi)^{M-1} A_{j_1,\ldots,j_M}^T(f_1 + \cdots + f_M)\, s_{j_1,\ldots,j_M}(f_1, \ldots, f_M) \Big| = O(1).$$

Proof. We now prove Lemmas 1 and 2. If follows from

$$|a_j(t+u)a_k(t+v) - a_j(t)a_k(t)| \le |a_j(t+u)a_k(t+v) - a_j(t+u)a_k(t)| + |a_j(t+u)a_k(t) - a_j(t)a_k(t)|$$

and Condition 1 that there is a constant K_1 such that

$$|\sum_t a^T_{j_1}(t+u_1)\cdots a^T_{j_{M-1}}(t+u_{M-1})a^T_{j_M}(t)\exp(-ift) - A^T_{j_1\ldots j_M}(f)| \le K_1(|u_1|+\cdots+|u_{M-1}|).$$

By the cumulant property,

$$\begin{aligned}
&\operatorname{cum}(\varphi^T_{j_1}(f_1),\ldots,\varphi^T_{j_M}(f_M))\\
&= \sum_{t_1}\cdots\sum_{t_M} a^T_{j_1}(t_1)\cdots a^T_{j_M}(t_M)\exp\left(-i\sum_{m=1}^{M} f_m t_m\right)\\
&\qquad\times C_{j_1,\ldots,j_M}(t_1-t_M,\ldots,t_{M-1}-t_M)\\
&= \sum_{u_1=-2(T-1)}^{2(T-1)}\cdots\sum_{u_{M-1}=-2(T-1)}^{2(T-1)}\exp\left(-i\sum_{m=1}^{M-1} f_m t_m\right)C_{j_1,\ldots,j_M}(u_1,\ldots,u_{M-1})\\
&\qquad\times\sum_t a^T_{j_1}(t+u_1)\cdots a^T_{j_{M-1}}(t+u_{M-1})a^T_{j_M}(t)\exp\left(-i\sum_{m=1}^{M} f_m t\right)\\
&= \epsilon_T + \sum_{u_1=-2(T-1)}^{2(T-1)}\cdots\sum_{u_{M-1}=-2(T-1)}^{2(T-1)}\exp\left(-i\sum_{m=1}^{M-1} f_m t_m\right)\\
&\qquad C_{j_1,\ldots,j_M}(u_1,\ldots,u_{M-1})A^T_{j_1\ldots j_M}(f_1+\cdots+f_M),
\end{aligned}$$

where

$$|\epsilon_T| \le K_2\sum_{u_1=-2(T-1)}^{2(T-1)}\cdots\sum_{u_{M-1}=-2(T-1)}^{2(T-1)}(|u_1|+\cdots+|u_{M-1}|)C_{j_1,\ldots,j_M}(u_1,\ldots,u_{M-1}).$$

It now follows from Condition 3,

$$T^{-1}|\epsilon_T| \le K_2\sum_{u_1=-2(T-1)}^{2(T-1)}\cdots\sum_{u_{M-1}=-2(T-1)}^{2(T-1)}T^{-1}(|u_1|+\cdots+|u_{M-1}|)C_{j_1,\ldots,j_M}(u_1,\ldots,u_{M-1}),$$

$T^{-1}(|u_1|+\cdots+|u_{M-1}|)\to 0$ and the dominated convergence theorem that

$$|\epsilon_T| = o(T). \tag{7.1}$$

Lemmas 1 and 2 follow from this and

$$\begin{aligned}
&s_{j_1,\ldots,j_M}(f_1,\ldots,f_{M-1})\\
&\quad= (2\pi)^{M-1}\sum\cdots\sum\exp\left(-i\sum_{1}^{M-1} f_m u_m\right)C_{j_1,\ldots,j_M}(u_1,\ldots,u_{M-1}) + o(1).
\end{aligned}$$

7.5. Complex Normal

Let X denote an k-dimensional random vector whose components are complex-valued random variables. If, for some $\mu \in \mathbb{C}^k$ and $k \times k$ Hermitian non-negative definite matrix Σ (that is, $\Sigma = \bar{\Sigma}^\top$),

$$\begin{pmatrix} \operatorname{Re} X \\ \operatorname{Im} X \end{pmatrix} \sim N_{2k}\left(\begin{pmatrix} \operatorname{Re}\mu \\ \operatorname{Im}\mu \end{pmatrix}, \tfrac{1}{2} \begin{pmatrix} \operatorname{Re}\Sigma & -\operatorname{Im}\Sigma \\ \operatorname{Im}\Sigma & \operatorname{Re}\Sigma \end{pmatrix} \right),$$

we say X has a complex normal distribution with mean μ and covariance matrix Σ, and is abbreviated by $X \sim N_k^c(\mu, \Sigma)$.

The FFT is asymptotically normal with mean specified according to the frequency f as described below.

Theorem 7.1. *Under Conditions 1 and 2, $\varphi_j^T(f)$ is asymptotically*

1. $N_1^c(0, 2\pi T s_{jj}(f) A_{jj}(0))$ *if* $f \neq 0 \mod \pi$,
2. $N_1(T c_j A_{jj}(0), 2\pi T s_{jj}(0) A_{jj}(0))$ *if* $f = 0, \pm 2\pi, \ldots$,
3. $N_1(0, 2\pi T s_{jj}(\pi) A_{jj}(0))$ *if* $f = \pm\pi, \ldots$.

Note that $A_{jj}(0) = \int a_j^2$. The above result implies that the real and the imaginary part of $\varphi_j^T(f)$ are approximately independent. Each is approximately normal with mean and variance $\pi T s_{jj}(f) \int a_j^2$.

Proof. To prove Theorem 7.1, we note that by Condition 1, $A^T_{j_1,\ldots,j_M}(f) = O(T)$. Recall that the Gaussian distribution has cumulants of order greater than 2 vanishes. The desired result now follows from Lemmas 1, 2 and that fact that

$$\begin{aligned} & T^{-M/2} \operatorname{cum}(\varphi^T_{j_1}(f_1), \ldots, \varphi^T_{j_M}(f_M)) \\ & \quad = T^{-M/2}(2\pi)^{M-1} A^T_{j_1,\ldots,j_M}(f_1 + \cdots + f_M)\, s_{j_1,\ldots,j_M}(f_1, \ldots, f_M) + o(T^{1-M/2}) \\ & \quad \to 0 \qquad \text{for } M > 2 \text{ as } T \to \infty. \end{aligned}$$

7.6. Asymptotics for Periodograms

The distributions of the FFT suggests the following statistic:

$$\tilde{s}_{jj}(f) = |\varphi_j^T(f)|^2 \Big/ \left(2\pi \sum_t [a_j(t/T)]^2 \right), \qquad f \in \mathbb{R}.$$

This is called **periodogram** and is an estimate of the spectral density function s_{jj}. For more historical remarks, see [10]. Note that if there is no tapering function, the periodogram is given by

$$\tilde{s}_{jj}(f) = (2\pi T)^{-1} |\varphi_j^T(f)|^2, \qquad f \in \mathbb{R}.$$

Let $f_m = 2\pi m/T$, $m = 0, \pm 1, \pm 2, \ldots, \pm T/2$ denote the Fourier frequencies. The result below describes the asymptotic distribution of the periodograms.

Theorem 7.2. *Under Conditions 1–3, $\tilde{s}_{jj}(f_m)$, $m = 1, \ldots, M = T/2$, are asymptotically independent $s_{jj}(f_m)\chi_2^2/2$. Also $\tilde{s}_{jj}(f)$ is asymptotically $s_{jj}(f)\chi_1^2$ for $f = \pm\pi, \pm 3\pi, \ldots$, independent of the $\tilde{s}_{jj}(f_m)$, $m = 1, \ldots, T/2$.*

Proof. The proof follows from Theorem 7.1 and the definition of the chi-square distribution.

The above result shows that the asymptotic variance of the periodogram is approximately $s_{jj}(f)^2$, which is usually positive. Thus the periodogram is not a consistent estimate of the spectral density function. The following section will present a class of consistent estimates obtained by smoothing the periodograms.

7.7. *Window Estimates — The Smoothed Periodograms*

A class of consistent estimates can be obtained by using a running mean or local average of the periodograms. Specifically, set

$$\hat{s}(f_m) = (2K+1)^{-1} \sum_{k=-K}^{K} \tilde{s}_{jj}\left(\frac{2\pi}{T}(m+k)\right).$$

It follows from the asymptotic distributional properties of the periodograms (Theorem 7.2) that $\hat{s}(f_m)$, $m = 1, \ldots, T/2$, are asymptotically independent with $\hat{s}(f) \sim s(f)\chi^2_{4K+2}/(4K+2)$ if $f \neq 0$, and $\hat{s}(0) \sim s(0)\chi^2_{2K}/(2K)$. An important implication of the above result is that consistency can be achieved by letting $K \to \infty$ and $K/T \to 0$ as $T \to \infty$.

More generally, let $W(\cdot)$ denote a weight function. Set

$$\hat{s}_{jj'}(f) = \sum_{k \neq 0} b_T^{-1} W\left(b_T^{-1}\left(f - \frac{2\pi k}{T}\right)\right) \tilde{s}_{jj'}\left(\frac{2\pi k}{T}\right), \tag{7.2}$$

where

$$\tilde{s}_{jj'}(f) = (2\pi T)^{-1} \varphi_j^T(f) \overline{\varphi_{j'}^T(f)}, \qquad f \in \mathbb{R},$$

and b_T is referred to as the **bandwidth** or **window width** that will be specified more clearly later. Certain properties of the weight function $W(\cdot)$ will be required in order to assure that the above estimate is consistent.

Condition 4. *The weight function $W(\cdot) : \mathbb{R} \to \mathbb{R}$ is a symmetric probability density function with a compact support $[-\pi, \pi]$.*

Under this condition, the bias of the window estimate is given by

$$E(\hat{s}_{jj}(f)) = \int W(\lambda) s_{jj}(f - b_T\lambda)\, d\lambda + O(T^{-1}b_T^{-1}).$$

In fact, more properties can be obtained and are stated in the following result.

Theorem 7.3. *Under Conditions 1–3 and suppose that the spectral density function s_{jj} does not vanish. Let $b_T \to 0$ and $b_T T \to \infty$ as $T \to \infty$. Then, $\hat{s}(f_m)$, $m = 1, \ldots, M$, are asymptotically normal with mean zero and covariance structure given by*

$$\lim_{T\to\infty} b_T T \operatorname{cov}(\hat{s}(f_1), \hat{s}(f_2)) = \begin{cases} 0 & \text{if } f_1 \neq f_2, \\ 2\pi s(f)^2 \int W^2 & \text{otherwise.} \end{cases} \tag{7.3}$$

Proof. Direct computation shows that

$$\operatorname{cov}(\tilde{s}_{jj}(f_1), \tilde{s}_{jj}(f_2)) = s_{jj}(f_1)\left\{\left(\frac{\sin T(f_1+f_2)/2}{T\sin(f_1+f_2)/2}\right)^2 + \left(\frac{\sin T(f_1-f_2)/2}{T\sin(f_1-f_2)/2}\right)^2\right\} + O(1/T).$$

Moreover,

$$\begin{aligned}\operatorname{cov}(\hat{s}_{jj}(f_1), \hat{s}_{jj}(f_2)) = {} & 2\pi T^{-1}\int W^T(f_1-\lambda)W^T(f_2-\lambda)s_{jj}(\lambda)^2\,d\lambda \\ & + 2\pi T^{-1}\int W^T(f_1-\lambda)W^T(f_2+\lambda)s_{jj}(\lambda)^2\,d\lambda \\ & + O(b_T^{-2}T^{-2}) + O(T^{-1}),\end{aligned}$$

where

$$W^T(\lambda) = b_T^{-1}\sum_{k=-\infty}^{\infty} W(b_T^{-1}(\lambda+2\pi k)).$$

The indicated covariance structure (7.3) is an easy consequence of these results.

To obtain the asymptotic normality, we need to show that all cumulants of order higher than 2 tend to zero as $T\to\infty$. This is carried out by directly computing the cumulants of the window estimates in a manner similar to the proof of Lemma 1.

7.8. Estimating the Transfer Function

Theorem 7.4. *Suppose that*

1. $\epsilon(t)$, $t = 0, 1, \ldots$ *satisfy Condition 2 and have mean zero,*
2. $X(t)$ *is uniformly bounded and* $s_{11} \neq 0$,
3. $\sum_u |u|h(u) < \infty$,
4. W *in Condition 4 is a uniform kernel.*

Let $b_T \to 0$, $b_T T \to \infty$, $b_T^5 T \to 0$ *as* $T\to\infty$. *Then* $\hat{H}(f_1), \ldots, \hat{H}(f_M))$ *is complex normal with mean* $(E\hat{H}(f_1), \ldots, E\hat{H}(f_M))$ *and covariance matrix whose entries are given by*

$$\operatorname{cov}(\hat{H}(f_1), \hat{H}(f_2)) = \eta(f_1-f_2)\frac{2\pi s_{\epsilon\epsilon}(f_1)}{b_T T s_{11}(f_1)}\int W^2,$$

where $\eta(0) = 1$ *and* $\eta(f) = 0$ *for* $f \neq 0$.

The weight function W is assumed to be uniform on $[-\pi, \pi]$ in order to simplify the presentation of the above asymptotic properties of the estimate of the transfer function. A more general approach can be found in [10].

Proof. We begin the proof of Theorem 7.4 with two lemmas.

Lemma 3. *Let* $(\mathbf{V}_n)$ *denote a sequence of random vectors converging in distribution to* $\mathbf{V}$. *Then there exists a probability space such that* $\mathbf{V}_n$ *converges to* $\mathbf{V}$ *almost surely.*

Proof. The proof can be found in [5].

Lemma 4. *Let* $(\mathbf{V}_n)$ *denote a sequence of random vectors in* $\mathbb{R}^p$ *converging in distribution to* $N_p^c(\mathbf{0}, \mathbf{I}_p)$ *and* $(\mathbf{U}_n)$ *a sequence of* $p\times p$ *unitary matrices. Then* $\mathbf{U}_n\mathbf{V}_n$ *converges to* $N_p^c(\mathbf{0}, \mathbf{I}_p)$ *as* $n\to\infty$.

Proof. This follows from Lemma 3.

Before proceeding to the proof, we remark that the following argument is simplified by assuming the series X to be non-random. The result nevertheless holds for general random X. Let φ_j^T, $j = 0, 1$, be the Fourier transform of Y and X, respectively. Let $2\pi k/T$ denote the Fourier frequency that is nearest to λ. Then

$$\begin{aligned}&\varphi_0^T(2\pi(k+l)/T)\\ &\quad = H(2\pi(k+l)/T)\varphi_1^T(2\pi(k+l)/T) + \varphi_\epsilon^T(2\pi(k+l)/T) + O(1)\\ &\quad = H(\lambda)\varphi_1^T(2\pi(k+l)/T) + \varphi_\epsilon^T(2\pi(k+l)/T) + O(1), \quad l = 0, \pm1, \ldots, \pm m,\end{aligned}$$

where $O(1)$ is uniformly in l. Now let $\mathbf{D}_0$ denote the $1 \times (2m+1)$ matrix given by

$$\mathbf{D}_0 = \frac{1}{\sqrt{2\pi T}}\left[\ \varphi_0^T(2\pi(k-m)/T) \quad \cdots \quad \varphi_0^T(2\pi k/T) \quad \cdots \quad \varphi_0^T(2\pi(k+m)/T)\ \right].$$

Define $\mathbf{D}_1$ and $\mathbf{D}_\epsilon$ similarly. Then

$$\mathbf{D}_0 = H(f)\mathbf{D}_1 + \mathbf{D}_\epsilon + O(T^{-1/2}).$$

Let $\mathbf{U} \equiv \mathbf{U}^T = [\mathbf{U}_1, \mathbf{U}_0]$ be a $(2m+1)\times(2m+1)$ unitary matrix whose first column is $\mathbf{U}_1 = \mathbf{D}_1^H(\mathbf{D}_1\mathbf{D}_1^H)^{-1/2}$, where $\mathbf{D}^H = \overline{\mathbf{D}}^\top$ is the conjugate transpose of $\mathbf{D}$. Then

$$\mathbf{D}_0\mathbf{U} = H(f)\mathbf{D}_1\mathbf{U} + \mathbf{D}_\epsilon\mathbf{U} + O(T^{-1/2}).$$

The first and the remaining columns of these matrices yield

$$[\hat{H}(f) - H(f)]\hat{s}_{11}(f)^{1/2}(2m+1)^{1/2} = \mathbf{D}_\epsilon\mathbf{U}_1 + O(T^{-1/2}), \tag{7.4}$$

$$\mathbf{D}_0\mathbf{U}_0 = \mathbf{D}_\epsilon\mathbf{U}_0 + O(T^{-1/2}). \tag{7.5}$$

By the property of the unitary matrix,

$$\begin{aligned}(2m+1)\hat{s}_{00} = \mathbf{D}_0\mathbf{D}_0^H &= \mathbf{D}_0\mathbf{U}_1\mathbf{U}_1^H\mathbf{D}_0^H + \mathbf{D}_0\mathbf{U}_0\mathbf{U}_0^H\mathbf{D}_0^H\\ &= \mathbf{D}_0\mathbf{D}_1^H(\mathbf{D}_1\mathbf{D}_1^H)^{-1}\mathbf{D}_1\mathbf{D}_0^H + \mathbf{D}_0\mathbf{U}_0\mathbf{U}_0^H\mathbf{D}_0^H.\end{aligned}$$

Thus

$$\hat{s}_{\epsilon\epsilon} = \mathbf{D}_0\mathbf{U}_0\mathbf{U}_0^H\mathbf{D}_0^H = \mathbf{D}_\epsilon\mathbf{U}_0\mathbf{U}_0^H\mathbf{D}_\epsilon^H + O_p(T^{-1/2}). \tag{7.6}$$

Now, according to Theorem 7.1, $\mathbf{D}_\epsilon \to_d N_{2m+1}^c N(\mathbf{0}, s_{\epsilon,\epsilon}(f)\mathbf{I})$ and therefore $s_\epsilon(f)^{-1/2}\mathbf{D}_\epsilon \to_d N_{2m+1}^c N(\mathbf{0}, \mathbf{I})$. By Lemma 4, $s_\epsilon(f)^{-1/2}\mathbf{D}_\epsilon\mathbf{U} \to_d N_{2m+1}^c N(\mathbf{0}, \mathbf{I})$, or $\mathbf{D}_\epsilon\mathbf{U} \to_d N_{2m+1}^c N(\mathbf{0}, s_\epsilon(f)\mathbf{I})$. This, together with (7.4) and (7.6) yield the desired result. This completes the proof of the theorem.

7.9. Estimating the Hemodynamic Response Function

From

$$H(f) = \sum_u h(u)\exp(-iuf),$$

we see that the hemodynamic response function is given by

$$h(u) = \frac{1}{2\pi}\int_0^{2\pi} H(f)\exp(iuf)\,df, \qquad u = 0, \pm1, \pm2, \ldots.$$

Let $\hat{H}(f)$ denote an estimate of $H(f)$ given by the last section, and let $Q \equiv Q_T$ denote a sequence of positive integers tending to ∞ with T. As an estimate of $h(u)$ by approximating the integral using finite sums, we define

$$\hat{h}(u) = \frac{1}{Q}\sum_{q=0}^{Q} \hat{H}\left(\frac{2\pi q}{Q}\right) \exp\left(i\frac{2\pi uq}{Q}\right), \qquad u = 0, \pm 1, \pm 2, \ldots .$$

Theorem 7.5. *Suppose that*

1. $\epsilon(t)$, $t = 0, 1, \ldots$ *satisfy Condition 2 and have mean zero,*
2. $X(t)$ *is uniformly bounded and* $s_{11} \neq 0$,
3. $\sum_u |u|h(u) < \infty$,
4. W *in Condition 4 is a uniform kernel.*

Let $Qb \to 0$ *as* $T \to \infty$. *Then*

$$E\hat{h}(u) = h(u) + \sum_{q\neq 0} h(u + qQ) + O(b) + O(T^{-1/2}).$$

In particular, $\hat{h}(u)$ *is asymptotically unbiased. Furthermore,* $\hat{h}(u_1), \ldots, \hat{h}(u_M)$ *are asymptotically normal with mean* $h(u_1), \ldots, h(u_M)$ *and covariance structure*

$$\operatorname{cov}(\hat{h}(u), \hat{h}(v)) = \frac{2\pi}{QbT}\Lambda^T(u, v)\int W^2 + O\big(T^{-1}\big), \quad u, v = 0, \pm 1, \pm 2, \ldots,$$

where

$$\Lambda^T(u, v) = \frac{1}{Q}\sum_{q=0}^{Q} \exp\left(i\frac{2\pi(u - v)q}{Q}\right) s_{\epsilon\epsilon}(2\pi q/Q)/s_{11}(2\pi q/Q).$$

Proof. The proof of Theorem 7.5 is tedious and very computational. We outline the argument here. The proof starts by assuming the X series to be non-random. The asymptotic normality then follows from the computation of the joint cumulants, it is shown that cumulants of order greater than 2 of $\hat{h}(u_1), \ldots, \hat{h}(u_M)$ tend to zero as $T \to \infty$. The desired result for the random X follows by invoking a standard technique in nonparametric regression for handling ratio of two random variates.

Acknowledgements

We like to thank Mechelle Lewis, Andrew Smith, Suman Sen and Roxanne Poole for their comments and helps to make the data available to us. We are grateful to three referees for their critical and constructive comments that improve the readability of the paper significantly. Finally, we thank Allen Song for the physics behind MRI.

References

[1] Afifi, A. K. and Bergman, R. A. (1998). *Functional Neuroanatomy: Text and Altas*. McGraw-Hill, USA.
[2] Aguirre, G. K., Zarahn, E. and D'Esposito, M. (1998). The Variability of Human, BOLD Hemodynamic Responses. *NeuroImage*, **8**, 360–369.
[3] Alexander, G. E., DeLong, M. R. and Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.*, **9**, 357–381.
[4] Bénar, C. G., Gross, D. W., Wang, Y., Petre, V., Pike, B., Dubeau, F. and Gotman, J. (2002). The BOLD Response to Interictal Epileptiform Discharges. *NeuroImage*, **17**, 1182–1192.
[5] Billingsley, P. (1995). *Probability and Measures, 3rd ed.*. Wiley, New York, USA.

[6] Boynton, G. M., Engel, S. A., Glover, G. H. and Heeger, D. J. (1996). Linear Systems Analysis of Functional Magnetic Resonance Imaging in Human V1. *The Journal of Neuroscience*, **16**, 4207–4221.

[7] Brillinger, D. R. (1974). Cross-Spectral Analysis of Processes with Stationary Invrements Including the Stationary G/G/∞ Queue. *AOP*, **2**, 815–827.

[8] Brillinger, D. R. (1975). The Identification of Point Process Systems. *AOP*, **3**, 909–929.

[9] Brillinger, D. R. (1978). A note on the estimation of evoked response. *Biological Cybernetics*, **31:3**, 141–144.

[10] Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*. Holden-Day, San Francisco, CA.

[11] Buckner, R. L. and Koutstaal, W. and Schacter, D. L. and Dale, A. M. and Rotte, M. and Rosen, B. R. (1998). Functional-Anatomic Study of Episodic Retrieval: II. Selective Averaging of Event-Related fMRI Trials to Test the Retrieval Success Hypothesis. *NeuroImage*, **7**, 163–175.

[12] Bullmore, E., Brammer, M., Williams, S. C. R., Rabe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R. and Sham, P. (1996). Statistical methods of estimation and inference for functional MR image analysis. *IEEE Trans. Med. Img.*, **35**, 261–277.

[13] Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E. and Brammer, M. J. (1999). Statistical methods of estimation and inference for functional MR image analysis. *IEEE Trans. Med. Img.*, **18**, 32–42.

[14] Caulhoun, V. D., Stevens, M. C., Pearlson, G. D. and Kiehl, K. A. (2004). fMRI analysis with the general linear model: removal of latency-induced amplitude bias by incorporation of hemodynamic derivative terms. *NeuroImage*, **22**, 252–257.

[15] Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, **29**, 162–173. http://afni.nimh.nih.gov/afni.

[16] Dale, A. M. and Buckner, R. L. (1997). Selective Averaging of Rapidly Presented Individual Trials Using fMRI. *Human Brain Mapping*, **5**, 329–340.

[17] DeLong, M. R., Alexander, G. E., Georgopoulos, A. P., Crutcher, M. D., Mitchell, S. J. and Richardson, R. T. (1984). Role of basal ganglia in limb movements. *Hum. Neurobiol.*, **2**, 235–244.

[18] Duann, J. R., Jung, T. P., Kuo, W. J., Yeh, T. C., Makeig, S., Hsieh, J. C. and Sejnowski, T. J. (1998). Single-Trial Variability in Event-Related BOLD Signals. *NeuroImage*, 823–835.

[19] Fried, I., Katz, A., McCarthy, G., Sass, K. J., Williamson, P. and Spencer, S. S. and Spencer, D. D. (1991). Functional organization of human supplementary motor cortex studied by electrical stimulation. *J. Neurosci.*, **11**, 3656-3666.

[20] Friston, K., Holmes, A., Worsley, K., Poline, J. B., Grasby, P., Williams, S., Frackowiak, R. and Turner, R. (1995). Analysis of fMRI time series revisited. *NeuroImage*, **2**, 45–53.

[21] Friston, K. J., Ashburner, J., Kiebel, S. J., Nichols, T. E. and Penny, W. D. (2007). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, Academic Press. http://www.fil.ion.ucl.ac.uk/spm/.

[22] Friston, K. J., Frith, C. D., Turner, R. and Frackowiak, R. S. J. (1995). Characterizing Evoked Hemodynamics with fMRI. *NeuroImage*, **2**, 157–165.

[23] Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D. and Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping*, **2**, 189–210.

[24] Friston, K. J., Jezzard, P. and Turner, R. (1994). Analysis of Functional MRI Time-Series. *Human Brain Mapping*, **1**, 153–171.

[25] Friston, K. J., Josephs, O., Rees, G. and Turner, R. (1998). Nonlinear event-related responses in fMRI. *Magnetic Resonance in Medicine*, **39**, 41–52.

[26] Genovese, C. R. (2000). A Bayesian Time-Course Model for Functional Magnetic Resonance Imaging Data (with discussion). *JASA*, **95:451**, 691–703.

[27] Glover, G. H. (1999). Deconvolution of impulse response in event-related fmri. *NeuroImage*, **9**, 416–429.

[28] Gössl, C., Fahrmeir, L. and Auer, D. P. (2001). Bayesian Modeling of the hemodynamic response function in BOLD fMRI. *NeuroImage*, **14**, 140–148.

[29] Goutte, C., Nielsen, F. A. and Hansen, L. K. (2000). Modeling the haemodynamic response in fMRI using smooth FIR filters. *IEEE Transactions on Medical Imaging*, **19**, 1188–1201.

[30] Huettel, S. A., Song, A. W. and McCarthy, G. (2004). *Functional Magnetic Resonance Imaging*. Sinauer Associates, Inc..

[31] Josephs, O., Turner, R. and Friston, K. (1995). Event-Related fMRI. *Human Brain Mapping*, **5**, 243–248.

[32] Kang, J. K., Bénar, C. G., Al-Asmi, A., Khani, Y. A., Pike, G. B., Dubeau, F. and Gotman, J. (2003). Using patient-specific hemodynamic response functions in combined EEG-fMRI studies in epilepsy. *NeuroImage*, **20**, 1162-1170.

[33] Lange, N. and Zeger, S. L. (1997). Non-linear Fourier Time Series Analysis for Human Brain Mapping by Functional Magnetic Resonance Imaging (with discussion). *JRSS*, **46:1**, 1–29.

[34] Lindquist, M. and Wager, T. D. (2007). Validity and power in hemodynamic response modeling: A comparison study and a new approach. *Human Brain Mapping*, Available online.

[35] Lindquist, M. and Wager, T. D. (2006). Validity and Power in Hemodynamic Response Modeling:

A comparison study and a new approach. *Human Brain Mapping*, DOI: 10.1002/hbm.20310.

[36] Maldjian, J. A., Laurienti, P. J. and Burdette, J. H. (2004). Precentral Gyrus Discrepancy in Electronic Versions of the Talairach Atlas. *NeuroImage*, **21:1**, 450–455.

[37] Marchini, J. L. and Ripley, B. D. (2000). A new statistical approach to detecting significant activation in functional MRI. *NeuroImage*, **12**, 366–380.

[38] Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E. and Buckner, R. L. (2000). Characterizing the Hemodynamic Response: Effects of Presentation Rate, Sampling Procedure, and the Possibility of Ordering Brain Activity Based on Relative Timing. *NeuroImage*, **11**, 735-759.

[39] Newton, H. J. (1988). *Timeslab: A Time Series Analysis Laboratory*. Wadsworth & Brooks/Cole.

[40] S. Ogawa, D. W. Tank, R. Menon, J. M. Ellermann, S. G. Kim, H. Merkle and K. Uğurbil (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *PNAS*, **89**, 5951–5955.

[41] Rosen, B. R., Buckner, R. L. and Dale, A. M. (1998). Event-related functional MRI: past, present, and future. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 773–780.

[42] Schwartz, L. (1966). *Mathematics for the Physical Sciences*, Hermann.

[43] Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I. and Flitney, D. E., Niazy, R., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M. and Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, **23(S1)**, 208–219. http://www.fmrib.ox.ac.uk/fsl/.

[44] Woolrich, M. W., Behrens, T. E. and Smith, S. M. (2004). Constrained linear basis sets for HRF modelling using Variational Bayes. *NeuroImage*, **21**, 1748–1761.

[45] Worsley, K. and Friston, K. (1995). Analysis of fMRI time series revisited-again. *NeuroImage*, **2**, 173–181.

[46] Zarahn, E., Aguirre, G. K. and D'Esposito, M. (1997). Empirical analyses of BOLD fMRI statistics. *NeuroImage*, **5**, 179–197.

[47] Zhang, C. M., Jiang, Y. and Yu, T. (2007). A comparative study of one-level and two-level semiparametric estimation of hemodynamic response function for fMRI data. *Statistics in Medicine*, **26**, 3845–3861.