Nonparametric Spectrum Estimation for Spatial Data

P.M. Robinson*

London School of Economics

September 5, 2005

Abstract

Smoothed nonparametric kernel spectral density estimates are considered for stationary data observed on a d-dimensional lattice. The implications for edge effect bias of the choice of kernel and bandwidth are considered. Under some circumstances the bias can be dominated by the edge effect. We show that this problem can be mitigated by tapering. Some extensions and related issues are disussed.

MSC: 62M30, 62M15

Keywords: nonparametric spectrum estimation; edge effect; tapering.

*Address: Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE; Telephone number: #44-(0)20-7955 7516; Fax number: #44-(0)20-7955 6592; email: p.m.robinson@lse.ac.uk. This research was supported by ESRC Grant R000239936.

1. INTRODUCTION

Let $\{x_t\}$ be a weakly dependent, zero-mean covariance stationary process, on a d-dimensional lattice $\mathbb{Z}^d$, for $d \geq 2$, such that t represents the multiple index $(t_1, ..., t_d)$. Defining the lag-u autocovariance $\gamma_u = cov(x_t, x_{t+u})$, we assume that x_t has a spectral density $f(\lambda)$, for $\lambda = (\lambda_1, ..., \lambda_d) \in \Pi^d$, $\Pi = (-\pi, \pi]$; this is given by

$$f(\lambda) = (2\pi)^{-d} \sum_{u \in \mathbb{Z}^d} \gamma_u e^{-iu.\lambda}, \tag{1.1}$$

where $u.\lambda = u_1\lambda_1 + ... + u_d\lambda_d$, and the expansion is well-defined under the condition

$$\sum_{u \in \mathbb{Z}^d} |\gamma_u| < \infty. \tag{1.2}$$

We are concerned with smoothed nonparametric estimation of $f(\lambda)$ given observations on t on the rectangular grid $\mathbb{N} = \{t : t_j \in [1, n_j], j = 1, ..., d\}$. A classical class of estimates is of weighted sample autocovariance type, depending on user choice of kernel function and bandwidth number. Define the lag-u sample autocovariance by

$$c_u = \frac{1}{n} \sum_{t(u)}{}' x_t x_{t+u}, \quad u \in \mathbb{N}^*, \tag{1.3}$$

where $\sum'_{t(u)}$ is a sum over $t_j, t_j + u_j \in [1, n_j]$, $j = 1, ..., d$, and $\mathbb{N}^* = \{u : 1 - n_j \leq u_j \leq n_j - 1, j = 1, ..., d\}$. A weighted autocovariance estimate of $f(\lambda)$ is given by

$$\tilde{f}(\lambda) = (2\pi)^{-d} \sum_{u \in \mathbb{N}^*} w_n(u) c_u e^{-iu.\lambda}, \tag{1.4}$$

where $w_n(u)$ is a suitable n-dependent weight function and $n = \Pi_{j=1}^d n_j$; "n-dependent" is a convenient short-hand for "dependent on $n_j, j = 1, ..., d$", that is justified because in asymptotic theory we regard each n_j as increasing with the overall sample size n, so we can write $n_j = n_j(n)$. In particular we consider $w_n(u)$ of form

$$w_n(u) = \prod_{j=1}^{d} k(u_j/m_j), \tag{1.5}$$

where $k(v)$ is an even, bounded, real-valued function such that

$$\lim_{v \to 0} \left\{ \frac{1 - k(v)}{|v|^q} \right\} = k_q, \tag{1.6}$$

for some $q > 0$, $0 < k_q < \infty$, and the $m_j = m_j(n)$ are non-negative integers such that $m_j \to \infty$ as $n \to \infty$, $j = 1, ..., d$.

Condition (1.6) controls the bias, ensuring in particular that $w_n(u) \to 1$ for all fixed u as $n \to \infty$. However, c_u is a biased estimate of γ_u unless $u = (0, ..., 0)$, and for $d \geq 2$ its bias is liable to be significant. We may write

$$Ec_u = \gamma_u \left\{1 + \eta_n(u)\right\}, \tag{1.7}$$

where

$$\eta_n(u) = \prod_{j=1}^{d} \left(1 - \frac{|u_j|}{n_j}\right) - 1. \tag{1.8}$$

For fixed u,

$$\eta_n(u) = \left(- \sum_{j=1}^{d} \frac{|u_j|}{n_j} \right) (1 + o(1)) \tag{1.9}$$

as $n \to \infty$. For u such that

$$\left| \sum_{j=1}^{d} \frac{u_j}{n_j} \right| \geq c \sum_{j=1}^{d} n_j^{-1}, \tag{1.10}$$

where $c > 0$, e.g. for $u_j > 0$, all j, we can apply the inequality between arithmetic and geometric means,

$$\sum_{j=1}^{d} n_j^{-1} \geq dn^{-1/d}, \tag{1.11}$$

to deduce that the bias in c_u for γ_u is of exact order $n^{-1/d}$.

This is the so-called "edge effect". Guyon (1982) found that the usual parametric Whittle estimates for lattice data have bias due to edge effect of exact order $n^{-1/d}$. The implication is that when $d = 2$ one has to "incorrectly" center the Whittle estimates before norming by $n^{\frac{1}{2}}$ and establishing asymptotic normality. For $d \geq 3$ matters are even worse, the Whittle estimates no longer being $n^{\frac{1}{2}}$-consistent.

Here we focus principally on the implications of the edge effect for the bias of smoothed nonparametric estimates of $f(\lambda)$. Intuitively, one expects the problem to be less serious because unlike in parametric estimation one would never aspire to a bias of order n^{-1}. This conjecture is confirmed in case of $\tilde{f}(\lambda)$, but this estimate does give non-negligible weight to u satisfying (1.10). We describe circumstances in which the edge effect does and does not dominate its bias.

A simple way of avoiding edge effect bias is to replace c_u by

$$\bar{c}_u = \left\{ \prod_{j=1}^{d} \left(1 - \frac{|u_j|}{n_j} \right) \right\}^{-1} c_u, \tag{1.12}$$

as advocated by Guyon (1982) in parametric Whittle estimation. There is now no bias, $E\bar{c}_u = \gamma_u$. However, the $\bar{c}_u$ lack a non-negative definiteness property of the c_u that contributes to guaranteeing non-negative estimates of the non-negative function $f(\lambda)$. Defining the periodogram

$$I(\lambda) = (2\pi)^{-d} n^{-1} \left| \sum_{t \in \mathbb{N}} x_t e^{it.\lambda} \right|^2, \tag{1.13}$$

we can write

$$\tilde{f}(\lambda) = \int_{\Pi^d} W_n(\lambda - \nu) I(\nu) d\nu, \tag{1.14}$$

where

$$W_n(\lambda) = \prod_{j=1}^{d} K_{m_j}(\lambda_j), \tag{1.15}$$

for

$$K_{m_j}(\lambda_j) = (2\pi)^{-1} \sum_{u_j \in \mathbb{Z}} k\left(\frac{u_j}{m_j}\right) e^{-i\lambda_j u_j}, \quad j = 1, ..., d. \tag{1.16}$$

Since $I(\lambda) \geq 0$ for all λ, choosing non-negative $K_{m_j}(\lambda_j)$ for all $\lambda_j \in \Pi$, $j = 1, ..., d$, thus ensures that $\tilde{f}(\lambda) \geq 0$ for all λ. The time series spectral analysis literature provides such choices, for example the modified Bartlett weights

$$k(v) = (1 - |v|) 1(|v| \leq 1) \tag{1.17}$$

lead to

$$K_{m_j}(\lambda_j) = \frac{\sin^2 \frac{1}{2}\lambda_j m_j}{2\pi m_j \sin^2 \frac{1}{2}\lambda_j} \geq 0, \tag{1.18}$$

as desired. An alternative choice is the Parzen weights.

$$\begin{aligned} k(v) &= 1 - 6v^2 + 6\,|v|^3\,, \quad |v| \leq \frac{1}{2}, \\ &= 2(1-|v|)^3, \quad \frac{1}{2} < |v| \leq 1, \\ &= 0, \quad |v| > 1, \end{aligned} \tag{1.19}$$

which again produces non-negative $K_{m_j}(\lambda_j)$ (see Anderson, 1971, p.518). In general, if $q > 2$ in (1.6) the $K_{m_j}(\lambda_j)$ need not be non-negative; this is the case if higher-order kernels are used, or the "flat top" kernels of Politis and Romano (1996) where effectively $q = \infty$. On the other hand even if the $K_{m_j}(\lambda_j)$ are non-negative, if c_u is replaced in (1.4) by the unbiased $\bar{c}_u$, a non-negative estimate of $f(\lambda)$ is not guaranteed.

The following section discusses $\tilde{f}(\lambda)$, principally focussing on bias but for completeness also recording a standard asymptotic approximation to the variance of $\tilde{f}(\lambda)$; as usual, on combining these results consistency can be deduced, and furthermore an approximation for the mean squared error of $\tilde{f}(\lambda)$ and an optimal choice of bandwidth. In Section 3 we introduce and analyze a tapered estimate, $\hat{f}(\lambda)$, of $f(\lambda)$, also employing a kernel and bandwidth similar to those in $\tilde{f}(\lambda)$. Dahlhaus and Künsch (1987) noted that Guyon's (1982) use of $\bar{c}_u$ in place of c_u loses the minimum-distance character of Whittle estimation. They pointed out that employing instead a periodogram based on tapered x_t avoids this draw-back, and can reduce edge effect bias sufficiently that, for $d = 2, 3, 4$, the usual $n^{\frac{1}{2}}$-consistency property of Whittle estimation is maintained. Correspondingly, our $\hat{f}(\lambda)$ is guaranteed non-negative, and we find that it reduces the bias due to edge effect. Soulier (1996) considered the effect of tapering on long memory random fields.

Section 4 consists of a Monte Carlo study of finite sample behaviour, and Section 5 discusses related issues and extensions.

2. UNTAPERED SPECTRUM ESTIMATES

We introduce the following assumptions.

Assumption 1: $k(v)$ is a real, even function such that $|k(v)| \leq 1$, (1.6) holds and

$$\int_{-\infty}^{\infty} |k(v)|\, dv < \infty. \tag{2.1}$$

Assumption 2: As $n \to \infty$

$$m_j \to \infty, \quad n_j \to \infty, \quad j = 1, ..., d. \tag{2.2}$$

Assumption 3: x_t is a covariance stationary process and

$$\sum_{u \in \mathbb{Z}^d} \left(\sum_{j=1}^{d} |u_j|^{\max(q,1)} \right) |\gamma_u| < \infty, \tag{2.3}$$

where q satisfies (1.6).

Theorem 1 *Let Assumptions 1-3 hold. Then as $n \to \infty$,*

$$E\tilde{f}(\lambda) = f(\lambda) + \alpha_{1n} + \alpha_{2n} + o\left(\sum_{j=1}^{d} \left(m_j^{-q} + n_j^{-1} \right) \right), \tag{2.4}$$

where

$$\alpha_{1n} = (2\pi)^{-d} k_q \sum_{j=1}^{d} m_j^{-q} \sum_{u \in \mathbb{Z}^d} |u_j|^q \gamma_u e^{-iu.\lambda}, \tag{2.5}$$

$$\alpha_{2n} = -(2\pi)^{-d} \sum_{j=1}^{d} n_j^{-1} \sum_{u \in \mathbb{Z}^d} |u_j| \gamma_u e^{-iu.\lambda}. \tag{2.6}$$

Proof: Using (1.7),

$$\begin{aligned} E\tilde{f}(\lambda) &= (2\pi)^{-d} \sum_{u \in \mathbb{N}^*} w_n(u) E\tilde{c}_u e^{-iu.\lambda} \\ &= (2\pi)^{-d} \sum_{u \in \mathbb{N}^*} w_n(u) \gamma_u (1 + \eta_n(u)) e^{-iu.\lambda}. \end{aligned} \tag{2.7}$$

The difference between this and $f(\lambda)$ is

$$(2\pi)^{-d}\sum_{u\in\mathbb{N}^*}\{w_n(u)-1\}\gamma_u e^{-iu.\lambda} \tag{2.8}$$

$$+(2\pi)^{-d}\sum_{u\in\mathbb{N}^*}w_n(u)\gamma_u\eta_n(u)e^{-iu.\lambda} \tag{2.9}$$

$$-(2\pi)^{-d}\sum_{u\in\mathbb{Z}^d-\mathbb{N}^*}\gamma_u e^{-iu.\lambda}. \tag{2.10}$$

Now

$$w_n(u)-1=\sum_{j=1}^{d}\left\{k\left(\frac{u_j}{m_j}\right)-1\right\}+v_n, \tag{2.11}$$

where v_n is linear in products of two or more of the $k(u_j/m_j)-1$. For any subset L of $\{1,...,d\}$,

$$\sum_{u\in\mathbb{N}}\prod_{j\in L}\left\{k\left(\frac{u_j}{m_j}\right)-1\right\}\gamma_u e^{-iu.\lambda}=\left(\prod_{j\in L}m_j\right)^{-q}\sum_{u\in\mathbb{N}}\left\{\prod_{j\in L}\frac{k(u_j/m_j)-1}{|u_j/m_j|^q}|u_j|^q\right\}\gamma_u e^{-iu.\lambda}, \tag{2.12}$$

proceeding as Hannan (1970, p.284). It follows from Assumption 2 that (2.8) is $\alpha_{1n}(1+o(1))$. Next

$$\eta_n(u)=-\sum_{j=1}^{d}\frac{|u_j|}{n_j}+s_n, \tag{2.13}$$

where s_n is linear in products of two or more $|u_j|/n_j$. We have

$$\sum_{u\in\mathbb{N}}w_n(u)\gamma_u\prod_{j\in L}\frac{|u_j|}{n_j}e^{-iu.\lambda}\sim\left(\prod_{j\in L}n_j\right)^{-1}\sum_{u\in\mathbb{Z}^d}\left(\prod_{j\in L}|u_j|\right)\gamma_u e^{-iu.\lambda}. \tag{2.14}$$

Thus (2.9) is $\alpha_{2n}(1+o(1))$. Finally, (2.10) is

$$O\left((2\pi)^{-d}\sum_{j=1}^{d}n_j^{-1}\sum{}'|u_j||\gamma_u|\right)=o\left(\sum_{j=1}^{d}n_j^{-1}\right), \tag{2.15}$$

where $\sum'$ is the sum over u such that $|u_j|\geq n_j$ and $u_k\in\mathbb{Z}$, $k\neq j$. □

Under Assumption 1, $\alpha_{1n},\alpha_{2n}\to 0$ as $n\to\infty$, so it follows from Theorem 1 that $\tilde{f}(\lambda)$ is asymptotically unbiased. Our interest is in the relative magnitude of α_{1n},α_{2n};

α_{1n} corresponds to the usual bias term stressed in the time series literature, while α_{2n} might be called the "edge effect term". Clearly α_{2n} is dominated by α_{1n} if and only if

$$\frac{\sum_{j=1}^{d} n_j^{-1}}{\sum_{j=1}^{d} m_j^{-q}} \to 0, \quad \text{as } n \to \infty. \tag{2.16}$$

In the time series case $d = 1$, (2.16) reduces to a condition standardly imposed in studies of bias (see e.g. Grenander and Rosenblatt, 1957, Chapter 4; Parzen, 1957; Anderson, 1971, Chapter 9). However, in practice the statistician is faced with fixed n_j, selects particular m_j, and never knows whether α_{1n} or α_{2n} is numerically the major source of bias. For $d > 1$, one should perhaps be less content with simply assuming (2.16) and thereby automatically recognizing α_{1n} as dominant.

Bias is often studied with a view to establishing consistency, and a choice of bandwidth that minimizes mean squared error (MSE). The latter involves the variance of $\tilde{f}(\lambda)$. We introduce two further assumptions.

Assumption 4: x_t is fourth-order stationary, (1.2) holds, and also

$$\sum_{s,t,u \in \mathbb{Z}^d} |\kappa_{stu}| < \infty,$$

where κ_{stu} is the fourth cumulant of x_0, x_s, x_t, x_u.

Assumption 5: As $n \to \infty$

$$\frac{m_j}{n_j} \to 0, \quad j = 1, ..., d.$$

The following theorem routinely extends classical results for $d = 1$ (see e.g. Anderson (1971, p.520), Grenander and Rosenblatt (1957, p.134), Hannan (1970, p.280), Parzen (1957); see also Brillinger's (1970) and Zhurbenko's (1986) discussion of spectral estimates for random fields). Thus the proof is omitted.

Theorem 2 *Let Assumptions 1, 2, 4 and 5 hold. Then as* $n \to \infty$

$$V\hat{f}(\lambda) = \beta_n + o\left(\prod_{j=1}^{d} \frac{m_j}{n_j}\right),$$

where

$$\beta_n = \left\{\prod_{j=1}^{d} \frac{m_j}{n_j}\right\} \int_{\mathbb{R}^d} k^2(u)du \; f(\lambda)^2 \left\{1 + 1(\lambda = 0, \bmod \pi)\right\},$$

where $1(.)$ *is the indicator function.*

Thus, under Assumptions 1-5, $\tilde{f}(\lambda)$ is mean-square consistent for $f(\lambda)$, and moreover

$$MSE\tilde{f}(\lambda) \sim (\alpha_{1n} + \alpha_{2n})^2 + \beta_n, \quad \text{as } n \to \infty.$$

Under (2.16), MSE$\tilde{f}(\lambda) \sim \alpha_{1n}^2 + \beta_n$ as $n \to \infty$, and as indicated by Zhurbenko (1986, p.164) this is minimized by $m_j \sim a_j n^{1/(d+2q)}$, $j = 1, ..., q$, where the a_j are certain positive constants. With this choice of the m_j, $\alpha_{1n}^2 + \beta_n$ has rate $n^{-2q/(d+2q)}$. However, if the n_j are such that (2.16) does not hold for these m_j, then the contribution of $\alpha_{2n}^2 + 2\alpha_{1n}\alpha_{2n}$ to MSE$\tilde{f}(\lambda)$ (of order $\left(\Sigma_{j=1}^{d} n_j^{-1}\right)^2$) will match or dominate $\alpha_{1n}^2 + \beta_n$. Notice that Assumption 5 implies (2.16) when $q \leq 1$ (as for the Bartlett weights (1.17)), but not when $q > 1$ (as for the Parzen weights (1.19)).

3. TAPERED SPECTRUM ESTIMATES

We introduce a taper function $h(v)$, satisfying

Assumption 6: $h(v)$ is Lipschitz-continuous on $[0, 1]$ and satisfies

$$h(0) = 0, \tag{3.1}$$

$$h(1-v) = h(v), \quad 0 \leq v < \frac{1}{2}, \tag{3.2}$$

$$\int_0^1 h^2(v)dv > 0. \tag{3.3}$$

An example of a taper satisfying Assumption 6 is the cosine bell

$$h(v) = \frac{1}{2}(1 - \cos(2\pi v)). \tag{3.4}$$

Define for integer s

$$h_{j,s} = h\left(\frac{s - \frac{1}{2}}{n_j}\right) \tag{3.5}$$

and thence the tapered sample autocovariances

$$\hat{c}_u = \frac{1}{H_n} \sum_{t(u)}{}' \left(\prod_{j=1}^{d} h_{j,t_j} h_{j,t_j+u_j} \right) x_t x_{t+u}, \tag{3.6}$$

where

$$H_n = \prod_{j=1}^{d} \sum_{t_j=1}^{n_j} h_{j,t_j}^2. \tag{3.7}$$

Consider the estimate

$$\hat{f}(\lambda) = (2\pi)^{-d} \sum_{u \in \mathbb{N}^*} w_n(u) \hat{c}_u e^{-iu.\lambda}. \tag{3.8}$$

We introduce:

Assumption 7: $f(\lambda)$ is twice boundedly differentiable on Π^d.

Assumption 8: For all sufficiently large n,

$$K_{m_j}(\lambda_j) \geq 0, \quad j = 1, ..., d. \tag{3.9}$$

Theorem 3 *Let Assumptions 1-3 and 6-8 hold. Then as $n \to \infty$*

$$E\hat{f}(\lambda) = f(\lambda) + \alpha_{1n}(1 + o(1)) + O\left(\sum_{j=1}^{d} n_j^{-2} \right). \tag{3.10}$$

Proof: As in (1.14) we may write

$$\hat{f}(\lambda) = \int_{\Pi^d} W_n(\lambda - \nu) I_h(\nu) d\nu, \tag{3.11}$$

where

$$I_h(\lambda) = (2\pi)^{-d} H_n^{-1} \left| \sum_{t \in \mathbb{N}} \left(\prod_{j=1}^{n_j} h_{j,t_j} \right) x_t e^{it.\lambda} \right|^2 . \tag{3.12}$$

We have

$$EI_h(\nu) = \int_{\Pi^d} f(\zeta) \prod_{j=1}^{d} g_j(\nu_j - \zeta_j) d\zeta_j \tag{3.13}$$

for

$$g_j(\nu_j) = \left(2\pi \sum_{t=1}^{n_j} h_{j,t_j}^2 \right)^{-1} \left| \sum_{j=1}^{n_j} h_{j,t_j} e^{it_j \nu_j} \right|^2 . \tag{3.14}$$

Then we may write

$$E\hat{f}(\lambda) - f(\lambda) = a + b, \tag{3.15}$$

where

$$a = \int_{\Pi^d} W_n(\lambda - \nu) \int_{\Pi^d} \prod_{j=1}^{d} g_j(\zeta_j) \left\{ f(\nu - \zeta) - f(\nu) \right\} d\zeta d\nu, \tag{3.16}$$

$$b = \int_{\Pi^d} W_n(\lambda - \nu) \left\{ f(\nu) - f(\lambda) \right\} d\nu. \tag{3.17}$$

Now $b =$ (2.8)+(2.10), and is thus $\alpha_{1n}(1+o(1))$. By Taylor's theorem and Assumption 7,

$$\left| f(\nu - \zeta) - f(\nu) + \sum_{j=1}^{d} \zeta_j \frac{\partial f(\nu)}{\partial \nu_j} \right| \leq C \left\| \zeta \right\|^2 , \tag{3.18}$$

where C denotes a generic arbitrarily large positive constant. Since the $g_j(\zeta_j)$ are even functions, the triangle inequality, Assumption 8 and (3.18) give

$$|a| \leq C \int_{\Pi^d} W_n(\lambda - \nu) \sum_{j=1}^{d} \int_{\Pi^d} \left\| \zeta_j \right\|^2 g_j(\zeta_j) d\zeta_j. \tag{3.19}$$

As in Dahlhaus and Künsch (1987), summation by parts and taking $h(\nu) = 0$, $\nu \notin [0, 1]$, give

$$\sum_{t_j=1}^{n_j} h_{j,t_j} e^{it_j \zeta_j} = \left\{ \exp(-i\zeta_j) - 1 \right\}^{-1} \sum_{t_j=0}^{n_j} D(h_{j,t_j}) \exp(it_j \zeta_j), \tag{3.20}$$

where $D(h_{j,t_j}) = h_{j,t_j+1} - h_{j,t_j}$. Since Assumption 6 implies $\sum_{t_j=1}^{n_j} h_{j,t_j}^2 \geq n_j/C$, the j-th term in the sum in (3.19) is bounded by

$$\begin{aligned} Cn_j^{-1} \int_\Pi \left| \sum_{t_j=0}^{n_j} D(h_{j,t_j}) \exp(it_j\zeta_j) \right|^2 d\zeta_j &\leq Cn_j^{-1} \sum_{t_j=0}^{n_j} D(h_{j,t_j})^2 \\ &\leq Cn_j^{-2}, \end{aligned} \tag{3.21}$$

from Assumption 6. By Assumption 8

$$\int_{\Pi^d} W_n(\nu) d\nu = k(0)^d = 1 \tag{3.22}$$

to complete the proof. □

Assumption 7 is stronger than Assumption 3 when $q = 1$, but weaker than Assumption 3 when $q = 2$. Assumption 8 could be relaxed but it implies non-negative estimates of $f(\lambda)$, and facilitates a simple proof. It would be possible to show under slightly stronger conditions that the $\left(\sum_{j=1}^{d} n_j^{-2}\right)$ remainder term in (3.10) is exact. We are content with a bound here as it is sufficient to demonstrate improvement over Theorem 1, and to show that under Assumption 4 the remainder is dominated by α_{1n} when $q \leq 2$, as is true for $k(v)$ given by the Parzen weights (1.19). The remainder term could be reduced by allowing the $K_{m_j}(\lambda_j)$ to have a higher-order kernel property, or to correspond to the kernels of Politis and Romano (1996), but then $\hat{f}(\lambda) \geq 0$ would no longer be guaranteed.

For completeness we record an approximation to the variance of $\hat{f}(\lambda)$ (cf. Hannan, 1970, p.270).

Theorem 4 *Let Assumptions 1-3 and 6-8 hold. Then as* $n \to \infty$

$$V\hat{f}(\lambda) = \frac{\int_0^1 h^4(v)dv}{\left\{\int_0^1 h(v)^2 dv\right\}^2} \beta_n + o\left(\prod_{j=1}^{d} \frac{m_j}{n_j}\right). \tag{3.23}$$

Since the coefficient of β_n in (3.23) exceeds 1 unless $h(v)$ is constant, Theorem 4 demonstrates the well-known cost of tapering.

4. MONTE CARLO STUDY OF FINITE SAMPLE PERFORMANCE

Finite sample bias and standard deviation were examined by a Monte Carlo simulation. Simple moving average (MA) models were simulated for various values of d on regular lattices, with $n_1 = n_2 ... = n_d$ (as in Robinson and Vidal Sanz, 2005). For $d = 2, 3$ we considered the symmetric multilateral MA model

$$x_t = \varepsilon_t + \rho \sum_{\substack{j_1=-1 \\ (j_1,j_2,j_3)\neq(0,0,0)}}^{1} \cdots \sum_{j_d=-1}^{1} \varepsilon_{t-j},$$

having spectral density

$$f(\lambda) = (2\pi)^{-d} \left\{1 + \rho v_d(\lambda_1, ..., \lambda_d)\right\}^2,$$

where $v_d(\lambda_1, ..., \lambda_d) = \prod_{j=1}^{d}(1 + 2\cos\lambda_j) - 1$. For $d = 2$ we generated data for both $\rho = 0.05$ and 0.1, with $n_1 = n_2 = 11, 15, 19$ (so $n = 121, 225, 361$). For $d = 3$ we generated data for both $\rho = 0.015$ and 0.03 with $n_1 = n_2 = 5, 7$ (so $n = 125, 343$). For $d = 4$ we considered the temporal spatial model

$$x_t = \varepsilon_t + \rho \sum_{\substack{j_1=-1 \\ (j_1,j_2,j_3)\neq(0,0,0)}}^{1} \sum_{j_2=-1}^{1} \sum_{j_3=-1}^{1} \sum_{j_4=1}^{1} \varepsilon_{t-j},$$

having spectral density

$$f(\lambda) = (2\pi)^{-4} h(\lambda_1, \lambda_2, \lambda_3, \lambda_4),$$

where

$$h(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = 1 + \rho^2 v_3\left(\lambda_1, \lambda_2, \lambda_3\right) + 2\rho v_3(\lambda_1, \lambda_2, \lambda_3)\cos\lambda_4.$$

We generated data for both $\rho = 0.015$ and 0.03 with $n_1 = n_2 = n_3 = n_4 = 5, 7$ (so $n = 625, 2401$).

We computed $\tilde{f}(\lambda)$ and $\hat{f}(\lambda)$ at both $\lambda = (0, ..., 0)$ and $(\pi/2, ..., \pi/2)$, using the Parzen weights (1.19) in both cases, and the cosine bell taper (3.4) for $\hat{f}(\lambda)$. For each combination, two values of $m_1 = ... = m_d$ were employed. The Monte Carlo biases and standard deviations, on the basis of 100 replications, are presented in Tables 1 and 2 respectively.

(Tables 1 and 2 about here)

The Parzen weights are ones for which $q = 2$, and so tapering is expected to reduce large sample bias. This is only partially borne out in the samples used in our simulations. For $d = 2$, tapering always reduces bias in case $\lambda = (0, ..., 0)$, but sometimes produces the opposite effect when $\lambda = (\pi/2, ..., \pi/2)$. For $d = 3$ tapering has virtually no effect when $\lambda = (0, ..., 0)$, and sometimes reduces, sometimes increases, bias when $\lambda = (\pi/2, ..., \pi/2)$. For $d = 4$, $\tilde{f}$ and $\hat{f}$ are virtually the same. The larger m_j in each pair tends to perform best, though there is little evidence of bias reduction with increase of n. As expected, bias tends to increase with ρ, and is always negative at the modal value $\lambda = (0, ..., 0)$. So far as standard deviations are concerned the predicted inflation due to tapering is noticeable; there is also generally an increase with m_j. Standard deviation tends also to increase with ρ, and to be larger at $\lambda = (0, ..., 0)$ than at $\lambda = (\pi/2, ..., \pi/2)$.

5. FINAL COMMENTS

1. There may be cancellations in the bias contributions of Theorems 1 and 3. For example, since $k_q > 0$, when $\gamma_u \geq 0$ for all u and $\lambda = 0$ we have $\alpha_{1n} > 0$ and $\alpha_{2n} < 0$.

2. Nonparametric spectral estimation is of considerable importance in inference for semiparametric models. Deriving asymptotic normality of a (possibly implicitly-defined) estimate of a vector-valued parameter typically requires establishing asymptotic normality of a statistic of form $n^{-\frac{1}{2}} \sum_{t\in\mathbb{N}} x_t$, where x_t can now be a column vector. Under a variety of weak dependence conditions we have

$$n^{-2} \sum_{t\in\mathbb{N}} x_t \rightarrow_d \mathcal{N}(0, 2\pi f(0)). \tag{5.1}$$

The construction of valid rules of inference requires using a consistent estimate of $f(0)$ with (5.1). Studentizing mean-like statistics by a nonparametric spectrum estimate was developed by Jowett (1955), Hannan (1957), Brillinger (1979), and has latterly been heavily employed in the econometric literature, see e.g. Newey and West (1987), Andrews (1991). Possible estimates are $\tilde{f}(0), \hat{f}(0)$ with $x_t x_{t+u}$ replaced in (1.3), (3.5) by $x_t x'_{t+u}$, the prime denoting transposition. If non-negative $K_{m_j}(\lambda_j)$ are used, $\tilde{f}(0)$ and $\hat{f}(0)$ will be non-negative definite, as is desirable for the construction of test statistics or interval estimates from these variance estimates. Their bias components are analogous to those of Theorems 1 and 3, and in connection with our discussion of these note that (1.17), where $q = 2$, was stressed by Newey and West (1987), and (1.19), where $q = 2$, is one of the possibilities mentioned by Andrews (1991).

3. Sometimes there is interest in spectral estimation for an unobservable sequence, in particular for the errors in a time series regression model, for example in the context of efficient semiparametric estimation of such a model (see e.g. Hannan, 1970, Chapter 7). Tapered and untapered spectral estimates based on residuals will incur an additional additive contribution to the bias, which in case of least squares correction for an unknown mean of x_t is of order $\Pi_{j=1}^{d}(m_j/n_j)$ (cf. Anderson, 1971, p.542). Denote this term α_{3n}. It always dominates α_{2n} when $d = 1$, but not necessarily when $d > 1$. Consider the case $n_j = n^{\phi_j}$,

$j = 1, ..., d$, where $0 < \phi_j \leq \ldots \leq \phi_d$, $\sum_{j=1}^{d} \phi_j = 1$, and $m_j = n_j^{\psi_j}$. Then $\alpha_{1n} \sim n^{-q \min_j(\phi_j \psi_j)}$, $\alpha_{2n} \sim n^{-\phi_1}$, $\alpha_{3n} \sim n^{\sum_{j=1}^{d} \phi_j \psi_j - 1}$. Then α_{2n} dominates α_{1n} if $\phi_1 < q \min_j(\phi_j \psi_j)$, and dominates α_{3n} if $\phi_1 < 1 - \sum_{j=1}^{d} \phi_j \psi_j$. If all n_j increase at the same rate, i.e. $\phi_j \equiv 1/d$, this requires respectively $\min_j \psi_j > 1/q$ and $\sum_{j=1}^{d} \psi_j < d - 1$; a necessary condition for both inequalities to hold is $d > q/(q-1)$, e.g. $d > 2$ for $q = 2$. For the tapered estimate $\hat{f}(\lambda)$, on the other hand, a necessary condition for the $O\left(\sum_{j=1}^{d} n_j^{-2}\right)$ edge effect term to dominate both the "leading" bias term in Theorem 2 and an $O\left(\Pi_{j=1}^{d}(m_j/n_j)\right)$ mean-correction term is $d > 2q/(q-2)$, under the same circumstances.

ACKNOWLEDGEMENTS

I thank Fabrizio Iacone for carrying out the simulations reported in Section 4, and two referees for helpful comments.

REFERENCES

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

Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817-858.

Brillinger, D.R., 1970. The frequency analysis of relations between stationary spatial series. Proc. Twelfth Bien. Sem. Canadian Math. Cong., Montreal, 39-81.

Brillinger, D.R., 1979. Confidence intervals for the cross-covariance function. Sel. Statist. Canada 5, 3-16.

Dahlhaus, R., Künsch, H., 1987. Edge effects and efficient parameter estimation for stationary random fields. Biometrika 74, 877-882.

Grenander, U., Rosenblatt, M., 1957. Statistical Analysis of Stationary Time Series. Wiley, New York.

Guyon, X., 1982. Parameter estimation for a stationary process on a d-dimensional lattice, Biometrika 69, 95-106.

Hannan, E.J., 1957. The variance of the mean of a stationary process. R. Statist. Soc. Ser. B 19, 282-285.

Hannan, E.J., 1970. Multiple Time Series. Wiley, New York.

Jowett, G.H., 1955. The comparison of means of sets of observations from sections of independent stochastic series. J. R. Statist. Soc. Ser. B 17, 208-227.

Newey, W.K., West, K.D., 1987. A simple positive semidefinite heteroskedadticity and autocorrelation consistent covariance matrix. Econometrica 56, 203-208.

Parzen, E., 1957. On consistent estimates of the spectrum of a stationary time series. Ann. Math. Statist. 28, 329-348.

Politis, D.N., Romano, J.P., 1996. On flat-top kernel spectral density estimation for honogeneous random fields. J. Statist. Plann. Inf. 51, 41-53.

Robinson, P.M., Vidal Sanz, J., 2005. Modified Whittle estimation of multilateral models on a lattice. Preprint.

Soulier, P., 1996. Nonparametric estimation of a strongly dependent stationary Gaussian field. REBRAPE 10, 69-86.

Yuan, J., Subba Rao, T., 1993. Spectral estimation for random fields with application to Markov modelling and texture classification. In R. Chellapa and A. Jain, eds. *Markov Random Fields: Theory and Applications.* Academic Press, Boston, 179-210.

Zhurbenko, I.G., 1986. The Spectral Analysis of Time Series. North-Holland, Amsterdam.

Table 1

Monte Carlo bias of $\tilde{f}(0) =: \hat{f}(0,...,0)$, $\hat{f}(0) =: \hat{f}(0,...,0)$, $\tilde{f}(\pi/2) =: \tilde{f}(\pi/2,...,\pi/2)$, $\hat{f}(\pi/2) =: \hat{f}(\pi/2,...,\pi/2)$, using Parzen weights and cosine bell taper for various values of d, ρ, n_j, m_j.

$d = 2$:

n_j	m_j	$\rho = 0.05$				$\rho = 0.1$			
		$\tilde{f}(0)$	$\hat{f}(0)$	$\tilde{f}(\pi/2)$	$\hat{f}(\pi/2)$	$\tilde{f}(0)$	$\hat{f}(0)$	$\tilde{f}(\pi/2)$	$\hat{f}(\pi/2)$
11	4	-.00127	-.0112	.0004	.0001	-.0298	-.0267	.0014	.0011
11	7	-.0081	-.0045	.0002	.0007	-.0186	-.0122	.0007	.0011
15	5	-.0100	-.0066	.0001	.0019	-.0231	-.0175	.0010	.0028
15	8	-.0070	-.0021	-.0001	.0021	-.0155	-.0071	.0002	.0026
19	6	-.0034	-.0054	-.0003	-.0001	-.0191	-.0139	.0003	.0006
19	9	-.0071	-.0030	-.0007	.0000	-.0151	-.0078	-.0004	.0005

$d = 3$:

n_j	m_j	$\rho = 0.015$				$\rho = 0.03$			
		$\tilde{f}(0)$	$\hat{f}(0)$	$\tilde{f}(\pi/2)$	$\hat{f}(\pi/2)$	$\tilde{f}(0)$	$\hat{f}(0)$	$\tilde{f}(\pi/2)$	$\hat{f}(\pi/2)$
5	2	-.0034	-.0034	.0001	.0002	-.0080	-.0080	.0002	.0003
5	3	-.0028	-.0029	-.0036	.0003	-.0068	-.0069	-.0085	.0003
7	3	-.0029	-.0029	-.0001	.0000	-.0008	-.0068	.0000	.0000
7	4	-.0024	-.0024	-.0039	-.0001	-.0057	-.0017	-.0088	.0000

$d = 4$:

n_j	m_j	$\rho = 0.015$				$\rho = 0.03$			
		$\tilde{f}(0)$	$\hat{f}(0)$	$\tilde{f}(\pi/2)$	$\hat{f}(\pi/2)$	$\tilde{f}(0)$	$\hat{f}(0)$	$\tilde{f}(\pi/2)$	$\hat{f}(\pi/2)$
5	2	-.0006	-.0006	.0000	.0000	-.0013	-.0013	.0000	.0000
5	3	-.0005	-.0005	.0010	.0000	-.0012	-.0012	.0000	.0000
7	3	-.0005	-.0005	.0000	.0000	-.0012	-.0012	.0000	.0000
7	4	-.0004	-.0004	.0000	.0000	-.0010	-.0010	.0000	.0000

Table 2

Monte Carlo standard deviation of $\tilde{f}(0) =: \hat{f}(0,...,0)$, $\hat{f}(0) =: \hat{f}(0,...,0)$, $\tilde{f}(\pi/2) =: \tilde{f}(\pi/2,...,\pi/2)$, $\hat{f}(\pi/2) =: \hat{f}(\pi/2,...,\pi/2)$, using Parzen weights and cosine bell taper for various values of d, ρ, n_j, m_j.

$d = 2$:

n_j	m_j	$\rho = 0.05$				$\rho = 0.1$			
		$\tilde{f}(0)$	$\hat{f}(0)$	$\tilde{f}(\pi/2)$	$\hat{f}(\pi/2)$	$\tilde{f}(0)$	$\hat{f}(0)$	$\tilde{f}(\pi/2)$	$\hat{f}(\pi/2)$
11	4	.0011	.0209	.0045	.0082	.0168	.0321	.0049	.0090
11	7	.0141	.0360	.0069	.0125	.0302	.0569	.0071	.0130
15	5	.0109	.0241	.0045	.0109	.0171	.0378	.0048	.0113
15	8	.0176	.0380	.0067	.0167	.0283	.0609	.0068	.0167
19	6	.0095	.0207	.0040	.0082	.0149	.0329	.0041	.0084
19	9	.0140	.0296	.0062	.0120	.0224	.0477	.0062	.0119

$d = 3$:

n_j	m_j	$\rho = 0.015$				$\rho = 0.03$			
		$\tilde{f}(0)$	$\hat{f}(0)$	$\tilde{f}(\pi/2)$	$\hat{f}(\pi/2)$	$\tilde{f}(0)$	$\hat{f}(0)$	$\tilde{f}(\pi/2)$	$\hat{f}(\pi/2)$
5	2	.0007	.0021	.0012	.0016	.0008	.0023	.0006	.0017
5	3	.0012	.0033	.0006	.0017	.0016	.0041	.0008	.0017
7	3	.0009	.0022	.0013	.0013	.0012	.0029	.0005	.0013
7	4	.0013	.0031	.0005	.0018	.0019	.0043	.0007	.0018

$d = 4$:

n_j	m_j	$\rho = 0.015$				$\rho = 0.03$			
		$\tilde{f}(0)$	$\hat{f}(0)$	$\tilde{f}(\pi/2)$	$\hat{f}(\pi/2)$	$\tilde{f}(0)$	$\hat{f}(0)$	$\tilde{f}(\pi/2)$	$\hat{f}(\pi/2)$
5	2	.0000	.0002	.0000	.0002	.0000	.0002	.0000	.0002
5	3	.0001	.0003	.0001	.0002	.0001	.0004	.0001	.0002
7	3	.0001	.0002	.0000	.0001	.0001	.0003	.0000	.0001
7	4	.0001	.0004	.0001	.0002	.0001	.0005	.0001	.0002