The Annals of Applied Probability
1999, Vol. 9, No. 1, 242–273

PREDICTING RANDOM FIELDS WITH INCREASING DENSE OBSERVATIONS[1]

BY MICHAEL L. STEIN

University of Chicago

This work investigates some spectral characteristics of the errors of optimal linear predictors for weakly stationary random fields. More specifically, for errors of optimal linear predictors, results here explicitly bound the fraction of the variance attributable to some set of frequencies. Such a bound is first obtained for random fields on $\mathbb{R}^d$ observed on the infinite lattice δJ for all J on the d-dimensional integer lattice. If the spectral density exists, then the faster the spectral density tends to 0 at high frequencies, the more quickly this bound tends to 0 as $\delta \downarrow 0$. Under certain conditions on the spectral density, a similar result is given for processes on $\mathbb{R}$ where both observations and predictands are confined to a finite interval and observations may not be evenly spaced. These results provide a powerful tool for studying a problem the author has previously addressed using different methods: the properties of linear predictors calculated under an incorrect spectral density. Specifically, this work gives a number of new rates of convergence to optimality for predictors based on an incorrect spectral density when the ratio of the incorrect to the correct spectral density tends to 1 at high frequencies.

1. Introduction. In a series of previous papers (Stein 1990a, b, 1993, 1997), I have argued that the low frequency behavior of a random field does not have much impact on optimal linear predictions of the field in some region when it is densely sampled in this region. These works addressed this problem by considering what happens when using an incorrect spectral density whose high frequency behavior is similar to that of the actual spectrum. This work takes a more direct approach to studying the effect of low frequency behavior on prediction problems for mean 0 weakly stationary random fields by bounding the fraction of variance of prediction errors attributable to various frequency ranges.

To clarify what I mean by the fraction of the variance attributable to a set of frequencies, consider a mean 0 weakly stationary random field Z with spectral measure F. The spectral representation of Z is given by $Z(x) = \int_{\mathbb{R}^d} \exp(i\omega^T x) Y(d\omega)$, where Y is a mean 0 complex random measure such that, for Borel sets Δ_1 and Δ_2, $E\{Y(\Delta_1)\overline{Y(\Delta_2)}\} = F(\Delta_1 \cap \Delta_2)$. For real con-

Received July 1997; revised April 1998.

[1]Supported in part by NSF Grants DMS-92-04504 and DMS-95-04470. Computations for this work were done using computer facilities supported in part by NSF Grant DMS-89-05292, awarded to the Department of Statistics at University of Chicago, and by University of Chicago Block Fund.

AMS 1991 *subject classifications*. Primary 62M20; secondary 62M40, 41A25.

Key words and phrases. Approximation in Hilbert spaces, design of time series experiments, fixed-domain asymptotics, infill asymptotics, kriging, sampling theorem.

stants $a_1, \dots, a_n$, the random variable $X = \sum_{j=1}^n a_j Z(x_j)$ can then be written as $X = \int_{\mathbb{R}^d} V(\omega) Y(d\omega)$, where $V(\omega) = \sum_{j=1}^n a_j \exp(i\omega^T x_j)$. Furthermore, $\mathrm{Var}(X) = \int_{\mathbb{R}^d} |V(\omega)|^2 F(d\omega)$. Let Z_B be the random field obtained by filtering out frequencies outside of B in the spectral representation: $Z_B(x) = \int_B \exp(i\omega^T x) Y(d\omega)$. Then $\mathrm{Var}\{\sum_{j=1}^n a_j Z_B(x_j)\} = \int_B |V(\omega)|^2 F(d\omega)$ and we might reasonably call $\int_B |V(\omega)|^2 F(d\omega) / \int_{\mathbb{R}^d} |V(\omega)|^2 F(d\omega)$ the fraction of the variance of X attributable to the frequencies in B.

The main results in this work (Theorems 1 and 3) put bounds on this ratio when the random variable X is the error of an optimal linear predictor. Loosely speaking, what the results say are that if observations are densely packed in the region in which we wish to predict the random field, then this ratio will be small if the set B does not contain frequencies of too large a magnitude. To give an explicit example of this phenomenon, suppose Z is a mean 0 process on $\mathbb{R}$ and $\mathrm{Cov}\{Z(x), Z(y)\} = e^{-\alpha|x-y|}$ for some $\alpha > 0$ so that Z has spectral density $f(\omega) = \alpha / \{\pi(\alpha^2 + \omega^2)\}$. Consider predicting $Z(\frac{1}{2}\delta)$ based on observing $Z(0)$ and $Z(\delta)$. The optimal linear predictor is then easily shown to be $\frac{1}{2}\mathrm{sech}(\frac{1}{2}\delta\alpha)\{Z(0) + Z(\delta)\}$ and the variance of the prediction error is $\tanh(\frac{1}{2}\delta\alpha)$ [Daley (1991), page 134]. Moreover, for covariance functions of this form, these results are unchanged if further observations are added outside of $[0, \delta]$. For this prediction problem, the function V corresponding to the prediction error is $V(\omega) = \frac{1}{2}\mathrm{sech}(\frac{1}{2}\delta\alpha)(1 + e^{i\omega\delta}) - e^{i\omega\delta/2}$ and straightforward calculations yield

$$
\begin{aligned}
|V(\omega)|^2 &= \left\{\mathrm{sech}\left(\tfrac{1}{2}\delta\alpha\right) - \cos\left(\tfrac{1}{2}\omega\delta\right) - 1\right\}^2 \\
&= 4\,\mathrm{sech}^2\left(\tfrac{1}{2}\alpha\delta\right)\left\{\sin^2\left(\tfrac{1}{4}\delta\omega\right) + \sinh^2\left(\tfrac{1}{4}\delta\alpha\right)\right\}^2.
\end{aligned}
$$

The second expression for $|V(\omega)|^2$ is useful for numerical calculations because it does not involve taking differences of nearly equal numbers. Define the ratio

$$
r(T, \delta) = \frac{\int_{-T}^{T} |V(\omega)|^2 f(\omega)\, d\omega}{\int_{-\infty}^{\infty} |V(\omega)|^2 f(\omega)\, d\omega}.
$$

Then if $\delta \downarrow 0$ and $\delta T \downarrow 0$, it can be shown that $r(T, \delta) \sim (\delta^3 T / 16\pi)(\alpha^2 + \frac{1}{3}T^2)$. We see that $r(T, \delta)$ is small whenever δT is small. For $\alpha = 1$, Figure 1 plots $r(T, \delta)$ computed numerically for $\delta = 0.1$, 0.05, and 0.025 as a function of T. As expected, $r(T, \delta)$ is small whenever δT is small and $r(T, \delta)$ decreases rapidly with δ for fixed T.

Theorem 1 in Section 2 provides an elementary but powerful approach to bounding the fraction of variance of prediction errors attributable to a set of frequencies when Z is observed at δJ for all $J \in \mathbb{Z}^d$, the integer lattice in d dimensions. Prediction problems based on such an infinite lattice of observations are relatively easy due to the fact that the space of possible linear predictors is isomorphic to a space of periodic functions [Hannan (1970)]. For any mean 0 weakly stationary process on $\mathbb{R}^d$, Theorem 1 gives a simple and generally sharp uniform (over all possible linear predictions) bound on the fraction of the variance of a prediction error attributable to some set of

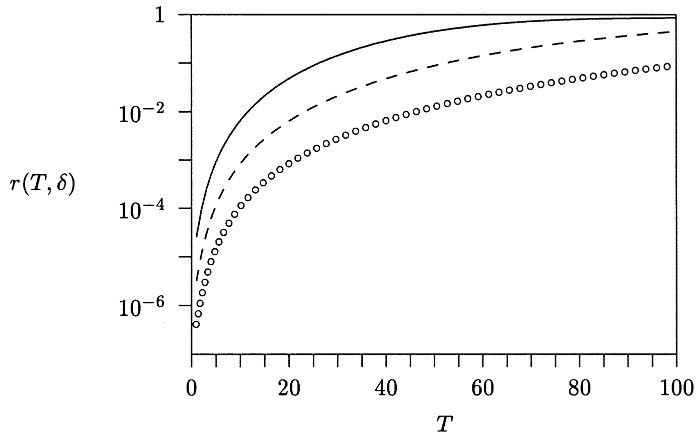

FIG. 1. *Plots of* $r(T, \delta)$ *when* $K(x) = e^{-|x|}$ *for various values of* δ: *solid line*, $\delta = 0.1$; *dashed line*, $\delta = 0.05$; *open circles*, $\delta = 0.025$.

frequencies. For a fixed and bounded set of frequencies, the faster the spectral density $f(\omega)$ tends to 0 as $|\omega| \to \infty$, the faster this bound tends to 0 as $\delta \downarrow 0$.

While Theorem 1 is of independent interest, it can be used to obtain results on the effect of misspecifying the spectral density on prediction. For example, Theorem 2 gives rates of convergence as $\delta \downarrow 0$ for the uniform asymptotic optimality of predictions when using f_1 instead of the correct f_0 under the conditions $f_0(\omega)(1 + |\omega|)^\alpha$ is bounded away from 0 and ∞ and there is a positive γ such that $(1 + |\omega|)^\gamma\{f_1(\omega) - f_0(\omega)\}/f_0(\omega)$ is bounded. Furthermore, the rates obtained are shown to be the best possible unless perhaps $\alpha = \gamma$ or $\alpha = 2\gamma$.

Section 3 uses a similar approach to obtain some results for the much more difficult setting when the observations and predictands are confined to a finite interval in $\mathbb{R}$. Theorem 3, which is proven in Section 4, gives the crucial bound on the fraction of the variance of prediction errors attributable to a frequency band of the form $(-r, r)$. As in the infinite lattice setting, this bound yields new results on the effects of using a misspecified spectral density in prediction. These results hold much more generally than the only previous bounds for a nonperiodic process observed on a finite interval, given in Stein (1990b). In the course of proving Theorem 3, I obtain a result that is relevant to the design of time series experiments as studied by Sacks and Ylvisaker (1966, 1968, 1970, 1971). The Appendix describes this connection.

Optimal linear prediction of a random field with mean 0 is generally called simple kriging in geology and hydrology [Cressie (1993), Christakos (1992)] and is extensively used in the atmospheric sciences, where it is called objective analysis [Daley (1991)]. Ordinary kriging, which is just best linear unbiased prediction when the mean of the random field is an unknown constant [Christensen (1991), Robinson (1991)], is more commonly used in geological applications. While the focus here is on simple kriging, many of the results obtained also apply to ordinary kriging. In particular, combining results of Section 3 with those in the Appendix yields bounds on the effect of adopting an incorrect spectral density on ordinary kriging predictions.

2. Infinite lattice. Consider a mean 0 weakly stationary random field Z on $\mathbb{R}^d$ with spectrum F and covariance function $K(x) = \int_{\mathbb{R}^d} \exp(i\omega^T x)F(d\omega)$. Let H_0 be the linear envelope of all finite sums of the form $\sum c_j Z(x_j)$ with c_j's real and x_j's in $\mathbb{R}^d$. For h and k in H_0, define their inner product as $E(hk)$ and let $H(F)$ be the Hilbert space given by the closure of H_0 with respect to this inner product. Similarly, let $\mathscr{V}_0$ be the linear envelope of functions expressible as finite sums of the form $\sum c_j \exp(i\omega^T x_j)$ with c_j's real and x_j's in $\mathbb{R}^d$. For φ and μ in $\mathscr{V}_0$, take their inner product as $\langle \varphi, \mu \rangle_F = \int_{\mathbb{R}^d} \varphi(\omega)\overline{\mu(\omega)}F(d\omega)$ and define $\|\varphi\|_F = \langle \varphi, \varphi \rangle_F^{1/2}$. Let $\mathscr{V}(F)$ be the closure of $\mathscr{V}_0$ with respect to this inner product. Identify $\sum c_j Z(x_j)$ with $\sum c_j \exp(i\omega^T x_j)$ and extend this correspondence to limits of such sums under their respective norms. Denoting by $\hat{h}$ the element $\mathscr{V}(F)$ corresponding to $h \in H(F)$, this correspondence is a unitary isomorphism of $H(F)$ and $\mathscr{V}(F)$: $E(hk) = \langle \hat{h}, \hat{k} \rangle_F$.

Next, let $H_\delta(F)$ be the subspace of $H(F)$ generated by $Z(\delta J)$ for $J \in \mathbb{Z}^d$ and $\mathscr{V}_\delta(F)$ the corresponding subspace of $\mathscr{V}(F)$. Define $e(h, \delta)$ to be the error of the best linear predictor of h when $Z(\delta J)$, $J \in \mathbb{Z}^d$, is the set of observations. Equivalently, $e(h, \delta)$ is the difference between h and its projection on $H_\delta(F)$. If $V \in \mathscr{V}(F)$ corresponds to the error of a best linear predictor, then V is orthogonal to $\mathscr{V}_\delta(F)$; that is, $\langle V, U \rangle_F = 0$ for all $U \in \mathscr{V}_\delta(F)$.

Let $A(r) = (-\pi r, \pi r]^d$. For a symmetric Borel set $B \subset A(\delta^{-1})$, define

$$M_\delta(F, B) = \sup \frac{\sum'_J F(S + 2\pi\delta^{-1}J)}{\sum_{J \in \mathbb{Z}^d} F(S + 2\pi\delta^{-1}J)},$$

where the supremum is over symmetric Borel subsets S of B, $\sum'_J$ indicates summation over $J \in \mathbb{Z}^d \setminus \{0\}$, $S + x$ means the set S translated by x and $0/0$ is defined to be 0. The following result puts a bound on how much of the variance of a prediction error can come from the frequencies in the set B.

THEOREM 1. *Suppose $V \in \mathscr{V}(F)$ is orthogonal to $\mathscr{V}_\delta(F)$. Then, for any symmetric Borel set $B \subset A(\delta^{-1})$,*

$$\int_B |V(\omega)|^2 F(d\omega) \le M_\delta(F, B)\|V\|_F^2. \tag{2.1}$$

Furthermore, if $\sum'_J F(\cdot + 2\pi\delta^{-1}J)$ is absolutely continuous with respect to $F(\cdot)$ on B,

$$\sup \frac{\int_B |V(\omega)|^2 F(d\omega)}{\|V\|_F^2} = M_\delta(F, B), \tag{2.2}$$

where the supremum is over all $V \in \mathscr{V}(F)$ orthogonal to $\mathscr{V}_\delta(F)$ satisfying $\|V\|_F^2 > 0$. If there are no such V, define this supremum as 0.

PROOF. Note that (2.1) is trivial if $\|V\|_F^2 = 0$ or $M_\delta(F, B) = 1$, so assume otherwise. Consider the function $V_{\delta, B}(\omega)$ with period $2\pi\delta^{-1}$ in each coordinate, $V_{\delta,B}(\omega) = V(\omega)$ for $\omega \in B$ and $V_{\delta,B}(\omega) = 0$ for $\omega \in A(\delta^{-1})\setminus B$. Then $M_\delta(F, B) < 1$ and B symmetric imply $V_{\delta,B} \in \mathscr{V}_\delta(F)$. Thus

$$\begin{aligned}
0 &= \int_{\mathbb{R}^d} V(\omega)\overline{V_{\delta,B}(\omega)}F(d\omega) \\
&= \int_B |V(\omega)|^2 F(d\omega) + \sum_J{}' \int_B V(\omega + 2\pi\delta^{-1}J)\overline{V_{\delta,B}(\omega)}F(d\omega + 2\pi\delta^{-1}J),
\end{aligned}$$

so that

$$\begin{aligned}
&\int_B |V(\omega)|^2 F(d\omega) \\
&\quad \le \sum_J{}' \int_B \left|V(\omega + 2\pi\delta^{-1}J)V_{\delta,B}(\omega)\right| F(d\omega + 2\pi\delta^{-1}J) \\
&\quad \le \Bigg\{ \sum_J{}' \int_B \left|V(\omega + 2\pi\delta^{-1}J)\right|^2 F(d\omega + 2\pi\delta^{-1}J) \\
&\qquad\qquad \times \sum_J{}' \int_B |V_{\delta,B}(\omega)|^2 F(d\omega + 2\pi\delta^{-1}J) \Bigg\}^{1/2} \\
&\quad \le \left[\left\{ \|V\|_F^2 - \int_B |V(\omega)|^2 F(d\omega) \right\} \frac{M_\delta(F,B)}{1 - M_\delta(F,B)} \int_B |V(\omega)|^2 F(d\omega) \right]^{1/2},
\end{aligned}$$

where the second inequality is by Cauchy–Schwarz. Straightforward algebra yields (2.1).

To prove (2.2), note that it is trivial if $M_\delta(F, B) = 0$, so assume otherwise. Denote by φ the Radon–Nikodym derivative of $\sum_J' F(\cdot + 2\pi\delta^{-1}J)$ with respect to $\sum_{J \in \mathbb{Z}^d} F(\cdot + 2\pi\delta^{-1}J)$ on B. Given $\varepsilon \in (0, M_\delta(F, B))$, let B_ε be the subset of B on which $\varphi(\omega) > M_\delta(F, B) - \varepsilon$. By the definition of $M_\delta(F, B)$, $\sum_{J \in \mathbb{Z}^d} F(B_\varepsilon + 2\pi\delta^{-1}J) > 0$, so that, by assumption, $F(B_\varepsilon) > 0$. Next, define a function V by

$$V(\omega + 2\pi\delta^{-1}J) = \begin{cases} 1, & \text{for } J = 0,\ \omega \in B_\varepsilon, \\ 1 - 1/\varphi(\omega), & \text{for } J \ne 0,\ \omega \in B_\varepsilon, \end{cases}$$

and 0 otherwise. Then it is straightforward to show that $V \in \mathscr{V}(F)$, V is orthogonal to $\mathscr{V}_\delta(F)$ and $\|V\|_F^2 > 0$. Furthermore $\int_B |V(\omega)|^2 F(d\omega) = F(B_\varepsilon)$ and $\varphi(\omega)/\{1 - \varphi(\omega)\}$ is the Radon–Nikodym derivative of $\sum_J' F(\cdot + 2\pi\delta^{-1}J)$ with respect to F on B_ε, so that

$$\begin{aligned}
\|V\|_F^2 &= F(B_\varepsilon) + \int_{B_\varepsilon} \left\{ \frac{1}{\varphi(\omega)} - 1 \right\}^2 \sum_J{}' F(d\omega + 2\pi\delta^{-1}J) \\
&= F(B_\varepsilon) + \int_{B_\varepsilon} \left\{ \frac{1}{\varphi(\omega)} - 1 \right\} F(d\omega) \\
&\le \frac{F(B_\varepsilon)}{M_\delta(F,B) - \varepsilon}.
\end{aligned}$$

It follows that $\int_B |V(\omega)|^2 F(d\omega)/\|V\|_F^2 \geq M_\delta(F, B) - \varepsilon$, which implies (2.2) since ε can be taken arbitrarily small. □

The condition preceding (2.2) that $\sum'_J F(\cdot + 2\pi\delta^{-1}J)$ be absolutely continuous with respect to F on B is always satisfied if $M_\delta(F, B) < 1$. It may not hold if $M_\delta(F, B) = 1$, in which case, (2.2) can be false. In particular, if $M_\delta(F, B) = 1$ and the support of F does not intersect B, then $\int_B |V(\omega)|^2 F(d\omega)$ is trivially always 0 so that the left-hand side of (2.2) is in fact 0, not 1.

Let us now consider some specific cases of Theorem 1. Suppose F has density f with respect to Lebesgue measure and there exist positive constants C_0 and C_1 and $\alpha > d$ such that

$$C_0(1 + |\omega|)^{-\alpha} \leq f(\omega) \leq C_1(1 + |\omega|)^{-\alpha}. \tag{2.3}$$

Then if $B(r)$ is the ball of radius r centered at the origin with $r < \pi/\delta$ so that $B(r) \subset A(\delta^{-1})$,

$$\begin{aligned} M_\delta(F, B(r)) &\leq \frac{C_1}{C_0}(1 + r)^\alpha \sup_{\omega \in B(r)} \sum_J{}' |\omega + 2\pi\delta^{-1}J|^{-\alpha} \\ &\leq \frac{C_1}{C_0}\left\{\frac{(1 + r)\delta}{\pi}\right\}^\alpha \xi_d(\alpha), \end{aligned} \tag{2.4}$$

where $\xi_d(\alpha) = \sum'_J |J|^{-\alpha}$. As a second example, suppose there exist positive constants α, C_0 and C_1 such that

$$C_0 \exp(-\alpha|\omega|) \leq f(\omega) \leq C_1 \exp(-\alpha|\omega|).$$

Then, for $r < \pi/\delta$,

$$\begin{aligned} &M_\delta(F, B(r)) \\ &\quad\leq \frac{C_1}{C_0}\exp(\alpha r) \sup_{\omega \in B(r)} \sum_J{}' \exp(-\alpha|\omega + 2\pi\delta^{-1}J|) \\ &\quad\leq \frac{C_1}{C_0}\exp(2\alpha r) \sum_J{}' \exp(-2\alpha\pi\delta^{-1}|J|) \\ &\quad\leq \frac{C_1}{C_0}\exp(2\alpha r)\Bigg\{(2d - 1)^d \exp(-2\alpha\pi\delta^{-1}) \\ &\qquad\qquad + 2^d d \sum_{J_1 = d}^{\infty} \sum_{J_2, \dots, J_d = 0}^{\infty} \exp\left(-2\alpha\pi\delta^{-1}d^{-1/2}\sum_{i=1}^d J_i\right)\Bigg\} \\ &\quad\leq \frac{C_1}{C_0}\exp\{-2\alpha(\pi\delta^{-1} - r)\} \\ &\qquad\times\left[(2d - 1)^d + \frac{2^d d}{\{1 - \exp(-2\alpha\pi\delta^{-1}d^{-1/2})\}^d}\right], \end{aligned}$$

where the third inequality follows by splitting up the sum over J into those terms for which all components of J have magnitude of at most d and all other terms. We see that, for r fixed and $\delta \downarrow 0$, $M_\delta(F, B(r))$ tends to 0 more quickly if f tends to 0 more quickly at high frequencies. While this behavior might initially appear to be unexpected, note that sampling theorems [Jerri (1977)], which give conditions under which perfect interpolation is possible, are an extreme example of this phenomenon. More specifically, suppose the process is bandlimited and B is the support of F. If B is contained in $A(\delta^{-1})$, then $M_\delta(F, B) = 0$, which implies $\int_B |V(\omega)|^2 F(d\omega) = 0$, so that $\|V\|_F^2 = 0$ since F has 0 mass outside B. We have just proven the standard sampling theorem for random fields observed on an infinite lattice with spacing δ, which says that perfect interpolation is possible if the support of F is a subset of $A(\delta^{-1})$. Thus, Theorem 1 can be viewed as an extension of sorts of the sampling theorem to nonbandlimited spectra.

Another way to think about Theorem 1 is in terms of the spectral representation of the prediction error process. Suppose $V(\omega; x, \delta)$ is the function in $\mathscr{V}(F)$ corresponding to $e(Z(x), \delta)$, the prediction error at x. Then $M_\delta(F, B)$ small implies that $\int_{B^c} V(\omega; x, \delta) Y(d\omega)$ will be similar to $e(Z(x), \delta) = \int_{\mathbb{R}^d} V(\omega; x, \delta) Y(d\omega)$ uniformly in x. Thinking of $e(Z(x), \delta)$ as a random field on $\mathbb{R}^d$, we have that the prediction error random field is only slightly distorted by filtering out the frequencies in B when $M_\delta(F, B)$ is small. Carr (1990) and Christakos [(1992), page 359] note that the prediction error random field can be viewed as a high-pass filter but they do not provide any quantitative assessment of this phenomenon.

Next consider applying Theorem 1 to the properties of linear predictions based on an incorrect spectral density. Suppose F_0 and F_1 have positive spectral densities f_0 and f_1 satisfying

$$0 < c_0 \le \frac{f_1(\omega)}{f_0(\omega)} \le c_1 < \infty \quad \text{for all } \omega, \tag{2.5}$$

so that $H(F_0) = H(F_1)$ as sets. Define E_i as expectation under F_i and $e_i(h, \delta)$ as the error of the best linear predictor of h under F_i. Let $H_{-\delta}(F)$ be the set of those h in $H(F)$ for which $Ee(h, \delta)^2 > 0$. Under (2.5), $H_{-\delta}(F_0) = H_{-\delta}(F_1)$. Think of F_0 as the true spectrum and F_1 as the presumed spectrum used to calculate linear predictors and evaluate their mean squared errors. Then

$$\frac{E_0\{e_1(h, \delta) - e_0(h, \delta)\}^2}{E_0 e_0(h, \delta)^2} = \frac{E_0 e_1(h, \delta)^2 - E_0 e_0(h, \delta)^2}{E_0 e_0(h, \delta)^2} \tag{2.6}$$

is the relative increase in mean squared error due to using F_1 instead of the correct F_0, where the equality follows from the fact that $e_0(h, \delta)$ is orthogonal to all elements of $H_\delta(F_0)$ under F_0 and hence is orthogonal to $e_1(h, \delta) - e_0(h, \delta)$. Another measure of the effect of presuming F_1 is the spectrum is

$$\frac{E_1 e_1(h, \delta)^2 - E_0 e_1(h, \delta)^2}{E_0 e_1(h, \delta)^2}, \tag{2.7}$$

the relative difference between the presumed mean squared error under F_1, given by $E_1 e_1(h, \delta)^2$, and the actual mean squared error of this predictor, given by $E_0 e_1(h, \delta)^2$. If both (2.6) and (2.7) are near 0, then little is lost in using F_1 both to predict h and to evaluate the mean squared error of the prediction. Daley [(1991), Section 4.9] reviews work in the atmospheric sciences that also takes the approach of considering the effects of using a fixed but wrong covariance structure on subsequent predictions. From a practical perspective, it would be more satisfactory to assess the effect of using an estimated spectral density on subsequent predictions. However, I maintain that it is essential to have a good understanding of the simpler problem of the effect of using a fixed but incorrect spectral density in order to know what one should mean by a good estimate of the spectral density when the ultimate goal is prediction. Handcock and Stein (1993) and Handcock and Wallis (1994) discuss the use of Bayesian methods to account for the uncertainty in the covariance structure on subsequent predictions.

For $h \in H(F_i)$, the element in $\mathscr{V}(F_i)$ corresponding to $e_i(h, \delta)$ is [Hannan (1970)]

$$T_\delta(\omega, h, f_i) = \frac{\sum_J f_i(\omega + 2\pi\delta^{-1}J)\hat{h}(\omega + 2\pi\delta^{-1}J)}{\sum_J f_i(\omega + 2\pi\delta^{-1}J)} - \hat{h}(\omega).$$

Define $\psi(\omega) = \{f_1(\omega) - f_0(\omega)\}/f_0(\omega)$, $u(r) = \sup_{\omega \in A(r)^c} \psi(\omega)$, $l(r) = \inf_{\omega \in A(r)^c} \psi(\omega)$ and $m(r) = \max(|u(r)|, |l(r)|)$. Then, for $h \in H(F_0)$,

$$\begin{aligned}
&\left|E_1 e_1(h, \delta)^2 - E_0 e_1(h, \delta)^2\right| \\
&\quad = \left|\int_{\mathbb{R}^d} f_0(\omega)\psi(\omega)|T_\delta(\omega, h, f_1)|^2\, d\omega\right| \\
&\quad \le c_0^{-1}\int_{A(\delta^{-1})} f_1(\omega)|\psi(\omega)|\,|T_\delta(\omega, h, f_1)|^2\, d\omega \\
&\qquad + m(\delta^{-1})\int_{A(\delta^{-1})^c} f_0(\omega)|T_\delta(\omega, h, f_1)|^2\, d\omega \\
&\quad \le c_0^{-1}\int_{A(\delta^{-1})} f_1(\omega)|\psi(\omega)|\,|T_\delta(\omega, h, f_1)|^2\, d\omega \\
&\qquad + m(\delta^{-1}) E_0 e_1(h, \delta)^2.
\end{aligned} \tag{2.8}$$

Next consider bounding $E_0\{e_1(h, \delta) - e_0(h, \delta)\}^2$. Define f_δ by $f_\delta(\omega) = f_0(\omega)$ for $\omega \in A(\delta^{-1})$ and $f_\delta(\omega) = f_1(\omega)$ elsewhere. Then

$$\begin{aligned}
E_0\{e_1(h, \delta) - e_0(h, \delta)\}^2 &\le 2E_0\{e_1(h, \delta) - e_\delta(h, \delta)\}^2 \\
&\quad + 2E_0\{e_\delta(h, \delta) - e_0(h, \delta)\}^2.
\end{aligned} \tag{2.9}$$

By equation (5) of Cleveland (1971),

$$E_0\{e_\delta(h,\delta)-e_0(h,\delta)\}^2 \le \frac{\{u(\delta^{-1})-l(\delta^{-1})\}^2}{4\{1+u(\delta^{-1})\}\{1+l(\delta^{-1})\}}E_0 e_0(h,\delta)^2. \tag{2.10}$$

The function $\Gamma(\omega)=T_\delta(\omega,h,f_\delta)-T_\delta(\omega,h,f_1)$ has period $2\pi\delta^{-1}$ in each coordinate and for $\omega\in A(\delta^{-1})$,

$$\Gamma(\omega)=\frac{f_0(\omega)\psi(\omega)T_\delta(\omega,h,f_1)}{f_0(\omega)+\Sigma'_J f_1(\omega+2\pi\delta^{-1}J)}.$$

Thus,

$$\begin{aligned}
&E_0\{e_1(h,\delta)-e_\delta(h,\delta)\}^2\\
&\quad=\int_{\mathbb{R}^d} f_0(\omega)|\Gamma(\omega)|^2\,d\omega\\
&\quad=\int_{A(\delta^{-1})}\sum_J f_0(\omega+2\pi\delta^{-1}J)\\
&\qquad\qquad\times\left|\frac{f_0(\omega)\psi(\omega)T_\delta(\omega,h,f_1)}{f_0(\omega)+\Sigma'_J f_1(\omega+2\pi\delta^{-1}J)}\right|^2 d\omega \qquad (2.11)\\
&\quad\le\int_{A(\delta^{-1})}\frac{\Sigma_J f_0(\omega+2\pi\delta^{-1}J)}{f_0(\omega)+c_0\Sigma'_J f_0(\omega+2\pi\delta^{-1}J)}\\
&\qquad\times f_0(\omega)\psi(\omega)^2|T_\delta(\omega,h,f_1)|^2\,d\omega\\
&\quad\le c_0^{-1}\max(1,c_0^{-1})\int_{A(\delta^{-1})} f_1(\omega)\psi(\omega)^2|T_\delta(\omega,h,f_1)|^2\,d\omega,
\end{aligned}$$

so that, for $h\in H_{-\delta}(F_0)$,

$$\begin{aligned}
&\frac{E_0\{e_1(h,\delta)-e_0(h,\delta)\}^2}{E_0e_0(h,\delta)^2}\\
&\quad\le\frac{\{u(\delta^{-1})-l(\delta^{-1})\}^2}{2\{1+u(\delta^{-1})\}\{1+l(\delta^{-1})\}} \qquad (2.12)\\
&\qquad+\frac{2\max(1,c_0^{-1})}{c_0E_0e_0(h,\delta)^2}\int_{A(\delta^{-1})}\psi(\omega)^2 f_1(\omega)|T_\delta(\omega,h,f_1)|^2\,d\omega
\end{aligned}$$

using (2.9)–(2.11). To make further progress in (2.8) or (2.12), we need to bound integrals of the form

$$\int_{A(\delta^{-1})}\sigma(\omega)f_1(\omega)|T_\delta(\omega,h,f_1)|^2\,d\omega,$$

where $\sigma = |\psi|$ in (2.8) and $\sigma = \psi^2$ in (2.12). The following lemma gives one example of how Theorem 1 can be used to give such a bound.

LEMMA 1. *Suppose* f_1 *satisfies* (2.3) *and there exist positive* D *and* β *such that* $0 \le \sigma(\omega) \le D(1+|\omega|)^{-\beta}$ *for all* ω. *Then, for any* $h \in H(F_1)$,

$$\int_{A(\delta^{-1})} \sigma(\omega) f_1(\omega) |T_\delta(\omega, h, f_1)|^2 \, d\omega$$
$$\le D E_1 e_1(h, \delta)^2 \left[2\left(\frac{\delta}{\pi}\right)^{\beta} + \frac{\beta C_1}{C_0}\left(\frac{\delta}{\pi}\right)^{\alpha} \xi_d(\alpha) \int_0^{\pi\delta^{-1}} (1+r)^{\alpha-\beta-1}\, dr \right].$$

PROOF. Let $\partial B(r)$ be the surface of $B(r)$. Then

$$\int_{A(\delta^{-1})} \sigma(\omega) f_1(\omega) |T_\delta(\omega, h, f_1)|^2 \, d\omega$$
$$\le \int_{B(\pi\delta^{-1})} \sigma(\omega) f_1(\omega) |T_\delta(\omega, h, f_1)|^2 \, d\omega$$
$$+ D\left(\frac{\delta}{\pi}\right)^{\beta} \int_{A(\delta^{-1}) \setminus B(\pi\delta^{-1})} f_1(\omega) |T_\delta(\omega, h, f_1)|^2 \, d\omega$$
$$\le D\int_0^{\pi\delta^{-1}} (1+r)^{-\beta} p(r)\, dr + D\left(\frac{\delta}{\pi}\right)^{\beta} E_1 e_1(h,\delta)^2,$$

where

$$p(r) = \int_{\partial B(r)} f_1(\nu) |T_\delta(\nu, h, f_1)|^2 \mu(d\nu)$$

and $\mu(d\nu)$ indicates the surface measure on $\partial B(r)$. Define $P(r) = \int_0^r p(s)\, ds$. By Theorem 1, $P(r) \le M_\delta(F_1, B(r)) E_1 e_1(h, \delta)^2$ and, by definition, $P(\pi\delta^{-1}) \le E_1 e_1(h,\delta)^2$. Integrating by parts,

$$\int_0^{\pi\delta^{-1}} (1+r)^{-\beta} p(r)\, dr$$
$$= \left(1 + \frac{\pi}{\delta}\right)^{-\beta} P(\pi\delta^{-1}) + \beta\int_0^{\pi\delta^{-1}} (1+r)^{-\beta-1} P(r)\, dr$$
$$\le \left\{ \left(\frac{\delta}{\pi}\right)^{\beta} + \beta\int_0^{\pi\delta^{-1}} (1+r)^{-\beta-1} M_\delta(F_1, B(r))\, dr \right\} E_1 e_1(h,\delta)^2.$$

Lemma 1 then follows from (2.4). □

Lemma 1 allows us to obtain explicit bounds on the effects of misspecifying the spectral density:

THEOREM 2. *If* f_1 *satisfies* (2.3), f_1/f_0 *satisfies* (2.5) *and* $|\psi(\omega)| \leq D(1 + |\omega|)^{-\gamma}$ *for all* ω, *then*

$$\sup_{h \in H_{-\delta}(F_1)} \frac{\left|E_1 e_1(h,\delta)^2 - E_0 e_1(h,\delta)^2\right|}{E_0 e_1(h,\delta)^2} = O\left(\delta^{\min(\alpha,\gamma)}(\log \delta^{-1})^{1\{\alpha=\gamma\}}\right) \tag{2.13}$$

and

$$\sup_{h \in H_{-\delta}(F_0)} \frac{E_0\{e_1(h,\delta) - e_0(h,\delta)\}^2}{E_0 e_0(h,\delta)^2} = O\left(\delta^{\min(\alpha,2\gamma)}(\log \delta^{-1})^{1\{\alpha=2\gamma\}}\right), \tag{2.14}$$

where $1\{\cdot\}$ *is an indicator function. Furthermore, except possibly in the case* $\alpha = \gamma$ *in* (2.13) *and* $\alpha = 2\gamma$ *in* (2.14), *these bounds are sharp in the sense that there exist* f_0 *and* f_1 *satisfying the stated conditions for which both conclusions are false if* $O(\cdot)$ *is replaced by* $o(\cdot)$.

Before proving this result, let us consider its interpretation. The larger the power of δ on the right-hand side of (2.13) or (2.14), the smaller the bound on the effect of using f_1 instead of the correct f_0. Figure 2 schematically illustrates the rates as a function of γ for a given α. We see that for $\gamma < \alpha/2$, the exponent of δ in (2.14) is twice that in (2.13). This result supports the empirical finding of geostatisticians that misspecifying the covariance structure has a greater effect on the evaluation of mean squared errors than on the predictions themselves [Starks and Sparks (1987)]. However, in neither case can the rate by faster than $O(\delta^{\alpha})$, so that, when $\gamma > \alpha$, the two rates are the same. Stein (1997) gives a result similar to (2.14) for periodic processes but the rate given there is not as good when $\gamma \leq (\alpha + d)/2$.

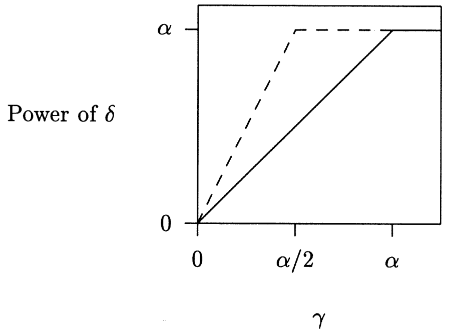

FIG. 2. *Schematic plot of rates of convergence in Theorem* 2: *solid line, power of* δ *in* (2.13); *dashed line, power of* δ *in* (2.14).

While Theorem 2 is stated in terms of a random field with known mean 0, the same result applies if the mean is an unknown constant μ. The point is that, whenever Z has a spectral density, μ can be recovered with probability 1 from the observations on the lattice for any $\delta > 0$, which follows from results of Yaglom [(1987), Section 16]. Thus, there is no meaningful difference between simple and ordinary kriging in this setting.

PROOF OF THEOREM 2. To prove (2.13), just apply Lemma 1 and the bound on $|\psi|$ to (2.8). To prove (2.14), apply Lemma 1 and the bound on ψ^2 to (2.12). To prove the sharpness of the bounds, suppose $f_0(\omega) = (1+|\omega|)^{-\alpha}$ and $\psi(\omega) = (1+|\omega|)^{-\gamma}$. First consider $\hat{h}(\omega) = \exp(-|\omega|^2) \in \mathscr{V}(F_1)$. Then

$$
\begin{aligned}
&E_0 e_0(h,\delta)^2 \\
&\quad = \int_{\mathbb{R}^d} (1+|\omega|)^{-\alpha} \left\{ \frac{\Sigma_J \exp(-|\omega + 2\pi\delta^{-1}J|^2)(1+|\omega+2\pi\delta^{-1}J|)^{-\alpha}}{\Sigma_J (1+|\omega+2\pi\delta^{-1}J|)^{-\alpha}} - \exp(-|\omega|^2) \right\}^2 d\omega \\
&\quad = \int_{A(\delta^{-1})} \sum_K (1+|\omega+2\pi\delta^{-1}K|)^{-\alpha} \\
&\qquad\qquad \times \left\{ \frac{\Sigma_J \exp(-|\omega + 2\pi\delta^{-1}J|^2)(1+|\omega+2\pi\delta^{-1}J|)^{-\alpha}}{\Sigma_J (1+|\omega+2\pi\delta^{-1}J|)^{-\alpha}} - \exp(-|\omega+2\pi\delta^{-1}K|^2) \right\}^2 d\omega \qquad (2.15) \\
&\quad = \int_{A(\delta^{-1})} (1+|\omega|)^{-\alpha} \exp(-2|\omega|^2) \\
&\qquad \times \left\{ \frac{\Sigma'_J (1+|\omega+2\pi\delta^{-1}J|)^{-\alpha}}{\Sigma_J (1+|\omega+2\pi\delta^{-1}J|)^{-\alpha}} \right\}^2 d\omega \\
&\qquad + \int_{A(\delta^{-1})} \sum_K{}' (1+|\omega+2\pi\delta^{-1}K|)^{-\alpha} \\
&\qquad\qquad \times \exp(-2|\omega|^2) \left\{ \frac{(1+|\omega|)^{-\alpha}}{\Sigma_J (1+|\omega+2\pi\delta^{-1}J|)^{-\alpha}} \right\}^2 d\omega \\
&\qquad + O(e^{-1/\delta}) \\
&\quad \asymp \delta^{\alpha}
\end{aligned}
$$

as $\delta \downarrow 0$, where $a(\delta) \asymp b(\delta)$ as $\delta \downarrow 0$ means $b(\delta)$ positive and $a(\delta)/b(\delta)$ is bounded away from 0 and ∞ for all sufficiently small δ. Since f_0/f_1 satisfies (2.5), we also have $E_0 e_1(h, \delta)^2 \asymp \delta^\alpha$ as $\delta \downarrow 0$. Next,

$$\begin{aligned}
&E_1 e_1(h, \delta)^2 - E_0 e_1(h, \delta)^2 \\
&\quad = \int_{A(\delta^{-1})} (1 + |\omega|)^{-\alpha-\gamma} \exp(-2|\omega|^2) \left\{ \frac{\Sigma'_J f_1(\omega + 2\pi\delta^{-1}J)}{\Sigma_J f_1(\omega + 2\pi\delta^{-1}J)} \right\}^2 d\omega \\
&\qquad + \int_{A(\delta^{-1})} \sum_K{}' (1 + |\omega + 2\pi\delta^{-1}K|)^{-\alpha-\gamma} \exp(-2|\omega|^2) \\
&\qquad\qquad \times \left\{ \frac{f_1(\omega)}{\Sigma_J f_1(\omega + 2\pi\delta^{-1}J)} \right\}^2 d\omega \\
&\qquad + O(e^{-1/\delta}) \\
&\quad \asymp \delta^{2\alpha} + \delta^{\alpha+\gamma}
\end{aligned}$$

as $\delta \downarrow 0$, which, together with (2.15), proves the sharpness of (2.13) for $\alpha \neq \gamma$. To prove the sharpness of (2.14) when $\alpha < 2\gamma$, as $\delta \downarrow 0$,

$$\begin{aligned}
&E_0\{e_0(h, \delta) - e_1(h, \delta)\}^2 \\
&\quad = \int_{A(\delta^{-1})} \sum_K f_0(\omega + 2\pi\delta^{-1}K) \\
&\qquad\qquad \times \left\{ \frac{\Sigma_J f_0(\omega + 2\pi\delta^{-1}J)\exp(-|\omega + 2\pi\delta^{-1}J|^2)}{\Sigma_J f_0(\omega + 2\pi\delta^{-1}J)} \right. \\
&\qquad\qquad\quad \left. - \frac{\Sigma_J f_1(\omega + 2\pi\delta^{-1}J)\exp(-|\omega + 2\pi\delta^{-1}J|^2)}{\Sigma_J f_1(\omega + 2\pi\delta^{-1}J)} \right\}^2 d\omega \\
&\quad \asymp \int_{A(\delta^{-1})} (1 + |\omega|)^{3\alpha} \left\{ \sum_J f_0(\omega + 2\pi\delta^{-1}J)\exp(-|\omega + 2\pi\delta^{-1}J|^2) \right. \\
&\qquad\qquad \times \sum_J f_1(\omega + 2\pi\delta^{-1}J) \\
&\qquad\qquad - \sum_J f_1(\omega + 2\pi\delta^{-1}J)\exp(-|\omega + 2\pi\delta^{-1}J|^2) \\
&\qquad\qquad\qquad \left. \times \sum_J f_0(\omega + 2\pi\delta^{-1}J) \right\}^2 d\omega \\
&\quad = \int_{A(\delta^{-1})} (1 + |\omega|)^{3\alpha} \left\{ f_0(\omega)\exp(-|\omega|^2) \sum_J{}' f_1(\omega + 2\pi\delta^{-1}J) \right. \\
&\qquad\qquad \left. - f_1(\omega)\exp(-|\omega|^2) \sum_J{}' f_0(\omega + 2\pi\delta^{-1}J) \right\}^2 d\omega \\
&\qquad + O(e^{-1/\delta})
\end{aligned}$$

$$= \int_{A(\delta^{-1})} (1+|\omega|)^{3\alpha} \exp(-2|\omega|^2) \times \left\{ \frac{1}{(1+|\omega|)^{\alpha}} \sum_J{}' \frac{1}{(1+|\omega+2\pi\delta^{-1}J|)^{\alpha+\gamma}} - \frac{1}{(1+|\omega|)^{\alpha+\gamma}} \sum_J{}' \frac{1}{(1+|\omega+2\pi\delta^{-1}J|)^{\alpha}} \right\}^2 d\omega + O(e^{-1/\delta})$$

$$\asymp \delta^{2\alpha},$$

which together with (2.15), proves the sharpness of (2.14) when $\alpha < 2\gamma$. To prove the sharpness of (2.14) when $\alpha > 2\gamma$, take f_0 and f_1 as before and define $\hat{h}_\delta = 1\{\omega \in A(\delta^{-1})\}$, which is in $\mathscr{V}(F_1)$. Then

$$\begin{aligned} E_0 e_0(h_\delta, \delta)^2 &= \int_{A(\delta^{-1})} f_0(\omega) \left\{ \frac{\Sigma'_J f_0(\omega + 2\pi\delta^{-1}J)}{\Sigma_J f_0(\omega + 2\pi\delta^{-1}J)} \right\}^2 d\omega \\ &\quad + \int_{A(\delta^{-1})} \sum_J{}' f_0(\omega + 2\pi\delta^{-1}J) \times \left\{ \frac{f_0(\omega)}{\Sigma_J f_0(\omega + 2\pi\delta^{-1}J)} \right\}^2 d\omega \\ &\asymp \int_{A(\delta^{-1})} (1+|\omega|)^{\alpha} \delta^{2\alpha}\, d\omega + \int_{A(\delta^{-1})} \delta^{\alpha}\, d\omega \\ &\asymp \delta^{\alpha-d} \end{aligned} \tag{2.16}$$

as $\delta \downarrow 0$; and,

$$\begin{aligned} &E_0\{e_0(h_\delta, \delta) - e_1(h_\delta, \delta)\}^2 \\ &\quad \geq \int_{A(\delta^{-1})} f_0(\omega) \left\{ \frac{f_0(\omega)}{\Sigma_J f_0(\omega + 2\pi\delta^{-1}J)} - \frac{f_1(\omega)}{\Sigma_J f_1(\omega + 2\pi\delta^{-1}J)} \right\}^2 d\omega \\ &\quad \asymp \int_{A(\delta^{-1})} (1+|\omega|)^{\alpha} \left\{ \sum_J{}' \frac{1}{(1+|\omega+2\pi\delta^{-1}J|)^{\alpha+\gamma}} - \frac{1}{(1+|\omega|)^{\gamma}} \sum_J{}' \frac{1}{(1+|\omega+2\pi\delta^{-1}J|)^{\alpha}} \right\}^2 d\omega. \end{aligned}$$

Now we can choose $\varepsilon > 0$ independent of δ such that

$$\frac{1}{(1+|\omega|)^{\gamma}} \sum_J{}' \frac{1}{(1+|\omega+2\pi\delta^{-1}J|)^{\alpha}} \geq 2 \sum_J{}' \frac{1}{(1+|\omega+2\pi\delta^{-1}J|)^{\alpha+\gamma}}$$

for all $\omega \in A(\varepsilon\delta^{-1})$, so there exists a constant C such that, for all δ sufficiently small,

$$\begin{aligned} &E_0\{e_0(h_\delta, \delta) - e_1(h_\delta, \delta)\}^2 \\ &\quad \geq C\int_{A(\varepsilon\delta^{-1})} (1 + |\omega|)^{\alpha - 2\gamma} \left\{ \sum_J{}' \frac{1}{(1 + |\omega + 2\pi\delta^{-1}J|)^{\alpha}} \right\}^2 d\omega \\ &\quad \asymp \delta^{\alpha + 2\gamma - d}, \end{aligned}$$

which, together with (2.16), proves the sharpness of (2.14) for $\alpha > 2\gamma$. □

3. Observations on a finite interval. The arguments in the previous section were critically dependent on the fact that the Hilbert space generated by the observations was isomorphic to an easily characterized space of periodic functions. When the observations do not have such a special structure, obtaining results analogous to Theorems 1 and 2 is much more difficult. Theorem 3 here considers processes observed perhaps unevenly on a finite interval and gives a bound analogous to the special case (2.4) of Theorem 1 when $\alpha = 2n$ for a positive integer n. This bound is then used to obtain some new results on the effect of using a misspecified spectral density on linear predictions.

Suppose Z is a mean 0 weakly stationary process on $\mathbb{R}$ with spectrum F and corresponding spectral density f. For a compact interval R, let $\mathscr{V}_R(F)$ be the subspace of $\mathscr{V}(F)$ generated by $e^{i\omega t}$, $t \in R$. For $N \geq 1$, let $\mathscr{V}_N(F)$ be the subspace of $\mathscr{V}_R(F)$ generated by $\varphi_{1N}, \dots, \varphi_{l_N N}$ where $\varphi_{jN} \in \mathscr{V}_R(F)$ for $1 \leq j \leq l_N$. Denote the orthogonal complement of $\mathscr{V}_N(F)$ relative to $\mathscr{V}_R(F)$ by $\mathscr{V}_N(F)^{\perp}$. Let $\mathscr{V}_{-N}(F)$ be all elements of $\mathscr{V}_R(F)$ that are not in $\mathscr{V}_N(F)$. Define $H_R(F)$, $H_N(F)$, $H_N(F)^{\perp}$ and $H_{-N}(F)$ to be the subsets of $H(F)$ corresponding to $\mathscr{V}_R(F)$, $\mathscr{V}_N(F)$, $\mathscr{V}_N(F)^{\perp}$ and $\mathscr{V}_{-N}(F)$. Finally, given a function φ: $\mathbb{R} \to \mathbb{C}$, define $\varphi_{(r)}$ by $\varphi_{(r)}(\omega) = \varphi(\omega)1\{|\omega| \leq r\}$. The key to obtaining results similar to those in the preceding section is to bound $\|\varphi_{(r)}\|_F$ for $\varphi \in \mathscr{V}_N(F)^{\perp}$, that is, to bound the fraction of the variance of prediction errors attributable to a set of low frequencies.

For $\varphi \in \mathscr{V}(F)$, define the operator M_φ on $\mathscr{V}(F)$ by $M_\varphi(\tau) = \langle \varphi, \tau \rangle_F$. Let P_R be the projection operator from $\mathscr{V}(F)$ to $\mathscr{V}_R(F)$: that is, for $\varphi \in \mathscr{V}(F)$, $\langle \varphi - P_R\varphi, \tau \rangle_F = 0$ for all $\tau \in \mathscr{V}_R(F)$, so that, in particular $M_\varphi(e^{i\omega t}) = M_{P_R\varphi}(e^{i\omega t})$ for $t \in R$. Define P_N to be the operator that projects elements of $\mathscr{V}(F)$ onto $\mathscr{V}_N(F)$. Then, for $\varphi \in \mathscr{V}_N(F)^{\perp} \subset \mathscr{V}_R(F)$,

$$\begin{aligned} \|\varphi_{(r)}\|_F^2 &= \langle \varphi_{(r)}, \varphi \rangle_F = \langle P_R\varphi_{(r)}, \varphi \rangle_F = \langle P_R\varphi_{(r)} - P_N P_R \varphi_{(r)}, \varphi \rangle_F \\ &\leq \|P_R\varphi_{(r)} - P_N P_R\varphi_{(r)}\|_F \|\varphi\|_F \leq \|P_R\varphi_{(r)} - \tau\|_F \|\varphi\|_F \end{aligned} \tag{3.1}$$

for any $\tau \in \mathscr{V}_N(F)$. Thus, we can bound $\|\varphi_{(r)}\|_F^2$ by approximating an element of $\mathscr{V}_R(F)$ with an element of $\mathscr{V}_N(F)$, which is exactly the problem addressed in Stein [(1990a), Section 4].

To apply the results from Stein (1990a), suppose from now on that $R = [0,1]$, C_0 and C_1 are positive constants and n is a positive integer such that

$$C_0 \le f(\omega)(1+\omega^2)^n \le C_1 \quad \text{for all } \omega \tag{3.2}$$

and set $K(t) = \int_{-\infty}^{\infty} f(\omega)e^{i\omega t}\,d\omega$, the covariance function corresponding to the spectral density f. For an interval Ω and a positive integer k, define $W^{k,2}(\Omega)$ to be the class of real-valued functions c on Ω such that $c^{(k-1)}$ exists and is absolutely continuous on Ω and the almost everywhere derivative of $c^{(k-1)}$ is square integrable on Ω. Set $n_0 = \lfloor (n-1)/2 \rfloor$. In the results to follow, assume that, for $|t| \le 1$,

$$K^{(2n-2)}(t) = \sum_{j=0}^{n_0} q_j |t|^{2j+1} + g(t), \qquad g \in W^{\max(3,n+1),2}([-1,1]). \tag{3.3}$$

Note that (3.2) implies $q_0 \ne 0$ so that $K^{(2n-1)}$ does not exist at 0. Thus, loosely speaking, (3.3) says that K has $\max(3, n+1)$ more derivatives on $(0,1]$, than it does at 0. For example, (3.3) holds for any rational spectral density f. For $n \ge 2$, (3.3) also holds if f can be written in the form $f_1 + f_2$, where f_1 is rational with $f_1(\omega) \asymp \omega^{-2n}$ as $\omega \to \infty$ and $\omega^{3n-1} f_2(\omega)$ is integrable. For a real-valued function c on $[0,1]$, define $\|c\| = \{\int_0^1 c(t)^2\,dt\}^{1/2}$ and set $Q = \max_{0 \le j \le n_0} |q_j|$. I will write $q_j(f)$ and $g(\cdot; f)$ when it is necessary to make clear the dependence on the spectral density.

THEOREM 3. *Suppose f satisfies* (3.2) *and that K satisfies* (3.3). *Furthermore, suppose $H_N(F)$ is made up of Z and its $n-1$ mean square derivatives at $0 = t_{0N} < t_{1N} < \cdots < t_{NN} = 1$ and set $S_N = \max_{1 \le i \le N}(t_{iN} - t_{i-1,N})$. Then there exists a universal constant A_n depending only on n such that, for all $\varphi \in \mathscr{V}_N(F)^{\perp}$,*

$$\|\varphi_{(r)}\|_F^2 \le A_n S_N^{2n} \left\{ \frac{C_1^2}{q_0^2} r^{2n} + b_n(f) \frac{C_1}{C_0} \right\} \|\varphi\|_F^2,$$

where, for $n \ge 2$,

$$b_n(f) = \left\{ \frac{Q + \sum_{j=2}^{n+1} \|g^{(j)}\|}{|q_0|} + \left| \frac{Q}{q_0} \right|^n + \left(|q_0|^{-1} \sum_{j=2}^{n+1} \|g^{(j)}\| \right)^{n/(n-1)} \right\}^2$$

and, for $n = 1$, $b_n(f) = \{(\|g''\| + \|g^{(3)}\|)/q_0\}^2$.

Theorem 3 is proven in Section 4. Although the bound looks rather complicated, it basically says that $\|\varphi_{(r)}\|_F^2 = O((1 + r^{2n})S_N^{2n})$, which is similar to (2.4) with S_N taking the role of δ and $\alpha = 2n$. The bound in Theorem 3 has the merit of holding for any S_N and is not just an asymptotic result.

Let us now consider applying Theorem 3 to derive analogs to the results in Theorem 2. For a spectrum F_i with density f_i, let K_i be the corresponding covariance function and $e_i(h, N)$ the error in predicting h based on $H_N(F_i)$. As in Section 2, let $\psi(\omega) = \{f_1(\omega) - f_0(\omega)\}/f_0(\omega)$. As a way of requiring that f_0 and f_1 have similar high-frequency behavior, assume there are positive constants D and γ such that

$$|\psi(\omega)| \le D(1 + |\omega|)^{-\gamma} \quad \text{for all } \omega. \tag{3.4}$$

Theorems 4 and 5 provide explicit bounds on the behavior of linear predictors under the wrong spectral density. While the bounds in Theorems 4 and 5 are rather messy, they again apply for all S_N and are not just asymptotic results. Under the assumption that $S_N = O(N^{-1})$, these bounds do yield rates of convergence, given in the corollaries. Proofs of Theorems 4 and 5 and Corollaries 1 and 2 are provided at the end of this section.

Define $\kappa(r, \gamma) = \gamma \int_0^r (1 + \omega)^{-\gamma - 1} \omega^{2n}\, d\omega$. For $r > 2$, $\kappa(r, \gamma) \asymp 1 + r^{2n - \gamma}$ unless $2n = \gamma$, in which case, $\kappa(r, \gamma) \asymp \log r$. Note that (3.2) and (3.4) imply $q_0(f_0) = q_0(f_1)$ so that q_0 is unambiguously defined in Theorems 4 and 5.

THEOREM 4. *Suppose f_0 and f_1 satisfy* (3.2), *ψ satisfies* (3.4) *and $H_N(F_1)$ is as in Theorem* 3. *If K_1 satisfies* (3.3),

$$\begin{aligned} &\left| E_1 e_1(h, N)^2 - E_0 e_1(h, N)^2 \right| \\ &\quad \le \frac{DC_1}{C_0}\Bigg[2N^{-\gamma} + A_n S_N^{2n} \\ &\qquad \times \left\{ \left(\frac{C_1}{q_0} \right)^2 \kappa(N, \gamma) + b_n(f_1) \frac{C_1}{C_0} \right\} \Bigg] E_1 e_1(h, N)^2, \end{aligned} \tag{3.5}$$

and if K_0 satisfies (3.3) *and $f_1(\omega) \ge f_0(\omega)$ for all ω,*

$$\begin{aligned} &\left| E_1 e_1(h, N)^2 - E_0 e_1(h, N)^2 \right| \\ &\quad \le D\Bigg[2N^{-\gamma} + A_n S_N^{2n} \\ &\qquad \times \left\{ \left(\frac{C_1}{q_0} \right)^2 \kappa(N, \gamma) + b_n(f_0) \frac{C_1}{C_0} \right\} \Bigg] E_0 e_0(h, N)^2. \end{aligned} \tag{3.6}$$

THEOREM 5. *Suppose* f_0 *and* f_1 *satisfy* (3.2), (3.4) *holds and* $H_N(F_1)$ *is as in Theorem* 3. *If both* K_0 *and* K_1 *satisfy* (3.3), *then, for all* $r > 0$,

$$\sup_{h \in H_{-N}(F_0)} \frac{E_0\{e_1(h,N) - e_0(h,N)\}^2}{E_0 e_0(h,N)^2} \tag{3.7}$$
$$\leq \frac{8C_1^6 D^2}{C_0^6 r^2 \gamma} + 4DA_n S_N^{2n}\left(\frac{C_1}{C_0}\right)^3 \times\left[\{b_n(f_0) + b_n(f_1)\}\frac{C_1}{C_0}(1 + r^{-\gamma}) + 2\left(\frac{C_1}{q_0}\right)^2\{r^{2n-\gamma} + \kappa(r,\gamma)\}\right].$$

If K_0 *satisfies* (3.3) *and* $f_1(\omega) \geq f_0(\omega)$ *for all* ω, *then, for all* $r > 0$,

$$\sup_{h \in H_{-N}(F_0)} \frac{E_0\{e_1(h,N) - e_0(h,N)\}^2}{E_0 e_0(h,N)^2} \tag{3.8}$$
$$\leq \frac{C_1 D^2}{C_0 r^{2\gamma}} + 2DA_n S_N^{2n}\left[b_n(f_0)\frac{C_1}{C_0}(1 + r^{-\gamma}) + \left(\frac{C_1}{q_0}\right)^2\{r^{2n-\gamma} + \kappa(r,\gamma)\}\right].$$

COROLLARY 1. *Suppose* f_0 *and* f_1 *satisfy* (3.2), ψ *satisfies* (3.4), $H_N(F_1)$ *is as in Theorem* 3 *and* $S_N = O(N^{-1})$. *If both* K_0 *and* K_1 *satisfy* (3.3) *or* K_0 *satisfies* (3.3) *and* $f_1(\omega) \geq f_0(\omega)$ *for all* ω *sufficiently large, then, for* $i = 0, 1$,

$$\sup_{h \in H_{-N}(F_i)} \frac{E_i\{e_i(h,N) - e_0(h,N)\}^2}{E_i e_i(h,N)^2} = O\left(N^{-\min(4\gamma n/(2n+\gamma),\, 2n)}(\log N)^{1\{\gamma = 2n\}}\right). \tag{3.9}$$

COROLLARY 2. *Suppose* f_0 *and* f_1 *satisfy* (3.2), ψ *satisfies* (3.4), $H_N(F_1)$ *is as in Theorem* 3 *and* $S_N = O(N^{-1})$. *If* K_1 *satisfies* (3.3) *or if* K_0 *satisfies* (3.3) *and* $f_1(\omega) \geq f_0(\omega)$ *for all* ω *sufficiently large, then*

$$\sup_{h \in H_{-N}(F_1)} \frac{\left|E_1 e_1(h,N)^2 - E_0 e_1(h,N)^2\right|}{E_0 e_1(h,N)^2} = O\left(N^{-\min(2n,\gamma)}(\log N)^{1\{\gamma = 2n\}}\right). \tag{3.10}$$

Figure 3 schematically indicates the rates of convergence in Corollaries 1 and 2. These results should be compared with those in Theorem 2 in the

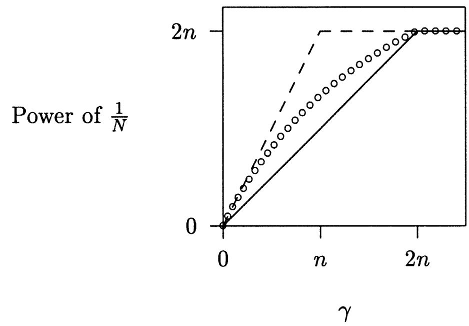

FIG. 3. *Schematic plot of rates of convergence in Corollaries* 1 *and* 2: *solid line*, *power of* $1/N$ *in* (3.10); *open circles*, *power of* $1/N$ *in* (3.9); *dashed line*, *conjectured best possible power of* $1/N$ *in* (3.9).

infinite lattice setting. I would conjecture that the rates obtained in that setting should also be attainable here if one sets $\alpha = 2n$ and $\delta = N^{-1}$. Thus, comparing Corollary 2 with (2.13) suggests that we have obtained the best possible rate of convergence except possibly when $\gamma = 2n$. However, comparing Corollary 1 with (2.14) suggests that the rates in this corollary can be improved to $N^{-\min(2n, 2\gamma)}$ when $\gamma < 2n$.

Proposition 2 of Stein (1990b) is comparable to Corollary 2 but only applies if $f_0(\omega) = (a^2 + \omega^2)^{-n}$ for some positive a and γ is an even integer. Even so, Proposition 2 of Stein (1990b) gives a slower rate of convergence than Corollary 2 when $\gamma > n + 1$ and the same rate when $\gamma \leq n + 1$, with the sole exception that if $\gamma = 2$ and $n = 1$ then Proposition 2 gives the rate N^{-2} whereas Corollary 2 gives $N^{-2} \log N$. Proposition 1 of Stein (1990b) gives results comparable to Corollary 1 but only in the special case $f_0(\omega) = (a^2 + \omega^2)^{-n}$ and $f_1(\omega) = (b^2 + \omega^2)^{-n}$, in which case, it gives the rate N^{-2} for $n = 1$ and N^{-3} for $n \geq 2$. For this special case, Corollary 1 implies the rate $N^{-2} \log N$ for $n = 1$ and $N^{-4n/(n+1)}$ for $n \geq 2$, so that Corollary 1 has a slower rate of convergence for $n = 1$ or 2, the same rate for $n = 3$ and a faster rate for $n \geq 4$.

There are many other possible bounds one could derive using the results in Stein (1990a) and (3.1). For example, Theorem 4.1 of Stein (1990a) can be used to bound the right-hand side of (3.1) under weaker conditions than (3.3). Theorem 4.3 of Stein (1990a) makes it possible to obtain bounds when only Z and not its mean square derivatives are observed. In both cases, the resulting rates of convergence will generally be slower than those given here. Finally, the Appendix shows that Corollaries 1 and 2 also apply to ordinary kriging predictors if K_0 and K_1 satisfy (3.3).

When considering processes on a finite interval, the low frequency behavior of the process should have very little impact on the behavior of predictions. Thus, I believe that the condition on the spectral densities in (3.2) requiring $f(\omega)(1 + \omega^2)^n$ to be bounded away from 0 and ∞ for all ω is unnecessarily strong. For example, I conjecture that Corollaries 1 and 2 still hold if the condition that f_0 and f_1 satisfy (3.2) is replaced by the weaker condition that, for $i = 0, 1$, $f_i(\omega)(1+\omega^2)^n$ is bounded away from 0 and ∞ for all $|\omega|$ sufficiently large. Furthermore, I believe that (3.9) holds for $i = 0$ without assuming f_1 is integrable in a neighborhood of the origin. In particular, when $f(\omega) = \omega^{-2n}$ and Z is Gaussian, then $Z^{(n-1)}$ is Brownian motion [see Yaglom (1987), Section 24.3]. Under this model and using the sampling scheme in Theorem 3, the best linear predictor of $Z(t)$ viewed as a function of $t \in [0,1]$ is uniquely defined as follows: it is a polynomial of degree $2n-1$ on each interval of the form $[t_{i-1,N}, t_{iN}]$ for $i = 1, \dots, N$ and it agrees with the available observations. It would be appealing to be able to give a specific rate at which a simple piecewise polynomial interpolant was an asymptotically optimal predictor in terms of conditions on the true spectral density f_0. Unfortunately, all of the proofs for the results in this section make use of (3.2) holding for all ω and I do not know how to modify the arguments to remove this condition.

PROOF OF THEOREM 4. Similar to the notation of Section 2, for $h \in H_R(F)$, define $T_N(\omega, h, f)$ to be the element in $\mathscr{V}_R(F)$ corresponding to $e(h,N)$. If K_1 satisfies (3.3), using Theorem 3 and integration by parts as in the proof of Lemma 1,

$$\begin{aligned}
&\left|E_1 e_1(h,N)^2 - E_0 e_1(h,N)^2\right| \\
&\quad\le \frac{C_1}{C_0}\int_{-\infty}^{\infty} |\psi(\omega)| f_1(\omega) \left|T_N(\omega,h,f_1)\right|^2 d\omega \\
&\quad\le \frac{DC_1}{C_0}\left[\int_0^N (1+\omega)^{-\gamma}\frac{d}{d\omega}\left\{\|T_N(\nu,h,f_1)1\{|\nu|\le\omega\}\|_{F_1}^2\right\} d\omega\right. \\
&\qquad\qquad\qquad\qquad\qquad\qquad \left. + N^{-\gamma}E_1 e_1(h,N)^2\right] \\
&\quad\le \frac{DC_1}{C_0}\left[2N^{-\gamma} + \gamma\int_0^N (1+\omega)^{-\gamma-1} A_n S_N^{2n}\left\{\left(\frac{C_1}{q_0}\right)^2 \omega^{2n} + b_n(f_1)\frac{C_1}{C_0}\right\} d\omega\right] \\
&\qquad \times E_1 e_1(h,N)^2
\end{aligned}$$

and (3.5) follows.

If K_0 satisfies (3.3) and $f_1 \ge f_0$, then using the shorthand e_i for $e_i(h,N)$,

$$\begin{aligned}
0 &\le E_1 e_1^2 - E_0 e_1^2 \\
&= E_1 e_0^2 - E_0 e_0^2 + E_0 e_0^2 - E_0 e_1^2 + E_1 e_1^2 - E_1 e_0^2 \\
&= E_1 e_0^2 - E_0 e_0^2 - E_0(e_1 - e_0)^2 - E_1(e_1 - e_0)^2 \\
&\le E_1 e_0^2 - E_0 e_0^2
\end{aligned}$$

and (3.6) follows using essentially the same argument as above. □

PROOF OF THEOREM 5. Define $f_2(\omega) = \max(f_0(\omega), f_1(\omega))$ and $f_{ijr}(\omega)$ to equal $f_i(\omega)$ for $|\omega| \le r$ and $f_j(\omega)$ for $|\omega| > r$. We have

$$E_0(e_1 - e_0)^2 \le 2E_0(e_2 - e_0)^2 + 2\frac{C_1}{C_0}E_1(e_2 - e_1)^2 \tag{3.11}$$

and

$$E_0(e_2 - e_0)^2 \le 2\frac{C_1}{C_0}E_2(e_2 - e_{20r})^2 + 2E_0(e_{20r} - e_0)^2. \tag{3.12}$$

Then

$$E_2(e_2 - e_{20r})^2 \le \frac{D^2}{r^{2\gamma}}E_2 e_2^2 \le \frac{D^2}{r^{2\gamma}}E_2 e_0^2 \le \frac{C_1 D^2}{C_0 r^{2\gamma}}E_0 e_0^2, \tag{3.13}$$

where the first inequality is due to equation (5) of Cleveland (1971). Since $f_{20r} \ge f_0$,

$$\begin{aligned} E_0(e_{20r} - e_0)^2 &\le E_0(e_{20r} - e_0)^2 + E_{20r}(e_{20r} - e_0)^2 \\ &= E_0 e_{20r}^2 - E_{20r}e_{20r}^2 + E_{20r}e_0^2 - E_0 e_0^2 \\ &\le E_{20r}e_0^2 - E_0 e_0^2. \end{aligned} \tag{3.14}$$

Thus, if K_0 satisfies (3.3) and $h \in H_{-N}(F_0)$,

$$\begin{aligned} E_0(e_{20r} - e_0)^2 &\le \int_{-r}^{r} |f_2(\omega) - f_0(\omega)|\,|T_N(\omega, h, f_0)|^2\, d\omega \\ &\le D\int_0^r (1+\omega)^{-\gamma}\,\frac{d}{d\omega}\left\{\|T_N(\nu, h, f_0)1\{|\nu| \le \omega\}\|_{F_0}^2\right\} d\omega \\ &\le DA_n S_N^{2n}\left[b_n(f_0)\frac{C_1}{C_0}(1 + r^{-\gamma}) + \left(\frac{C_1}{q_0}\right)^2\{r^{2n-\gamma} + \kappa(r,\gamma)\}\right]. \end{aligned}$$

which together with (3.12) and (3.13) implies

$$\begin{aligned} \frac{E_0(e_2 - e_0)^2}{E_0 e_0^2} \le{}& 2\left(\frac{C_1 D}{C_0 r^{\gamma}}\right)^2 + 2DA_n S_N^{2n} \\ &\times\left[b_n(f_0)\frac{C_1}{C_0}(1 + r^{-\gamma}) + \left(\frac{C_1}{q_0}\right)^2\{r^{2n-\gamma} + \kappa(r,\gamma)\}\right] \end{aligned}$$

for $h \in H_{-N}(F_0)$. Similarly,

$$\begin{aligned} \frac{E_1(e_2 - e_1)^2}{E_0 e_0^2} \le{}& 2\frac{C_1^5 D^2}{C_0^5 r^{2\gamma}} + 2DA_n S_N^{2n}\left(\frac{C_1}{C_0}\right)^2 \\ &\times\left[b_n(f_1)\frac{C_1}{C_0}(1 + r^{-\gamma}) + \left(\frac{C_1}{q_0}\right)^2\{r^{2n-\gamma} + \kappa(r,\gamma)\}\right] \end{aligned}$$

for $h \in H_{-N}(F_0)$. Then (3.7) follows from these last two inequalities and (3.11). Finally, (3.8) follows by noting that $f_1 \geq f_0$ and (3.14) imply

$$\begin{aligned} E_0(e_1 - e_0)^2 &\leq 2E_0(e_0 - e_{10r})^2 + 2E_0(e_{10r} - e_1)^2 \\ &\leq 2\big(E_{10r}e_0^2 - E_0 e_0^2\big) + 2E_1(e_{10r} - e_1)^2 \end{aligned}$$

and proceeding as in the proof of (3.7). $\square$

PROOF OF COROLLARIES 1 AND 2. When the conditions of Theorem 5 are satisfied, Corollary 1 follows by setting $r = N^{n/(n+\gamma+1)}$. Thus, we only need to consider proving Corollary 1 when K_0 satisfies (3.3) and $f_1(\omega) \geq f_0(\omega)$ for all ω sufficiently large. Define $f_3(\omega) = \min(f_0(\omega), f_1(\omega))$ and let K_3 be the corresponding covariance function. Since K_0 satisfies (3.3) and $f_3 - f_0$ has bounded support, K_3 also satisfies (3.3). Next,

$$E_0(e_1 - e_0)^2 \leq \frac{2C_1}{C_0} E_3(e_1 - e_3)^2 + 2E_0(e_3 - e_0)^2. \tag{3.15}$$

The corollary follows under the stated conditions by applying (3.8) with $r = N^{n/(n+\gamma+1)}$ to the first term on the right-hand side of (3.15) and applying (3.7) with $r = N^{n/(n+\gamma+1)}$ to the second term on the right-hand side of (3.15).

To prove Corollary 2, note that it follows immediately from Theorem 4 when the conditions of that theorem are satisfied. Thus, we are again left with proving the result when K_0 satisfies (3.3) and $f_1(\omega) \geq f_0(\omega)$ for all ω sufficiently large. We have

$$\begin{aligned} E_1 e_1^2 - E_0 e_1^2 = {} & E_1 e_1^2 - E_3 e_1^2 + E_3 e_1^2 - E_3 e_3^2 + E_3 e_3^2 - E_0 e_3^2 \\ & + E_0 e_3^2 - E_0 e_0^2 + E_0 e_0^2 - E_0 e_1^2, \end{aligned}$$

so that

$$\begin{aligned} |E_1 e_1^2 - E_0 e_1^2| \leq {} & |E_1 e_1^2 - E_3 e_1^2| + E_3(e_1 - e_3)^2 + |E_3 e_3^2 - E_0 e_3^2| \\ & + E_0(e_3 - e_0)^2 + E_0(e_1 - e_0)^2. \end{aligned}$$

The result follows by applying (3.6) to the first term on the right-hand side of this inequality, (3.5) to the third term and Corollary 1 to the other three terms. $\square$

4. Proof of Theorem 3.

Define the operator $\Phi: \mathbb{R}^n \times L^2([0,1]) \to \mathscr{V}_R(F)$ by taking $\varphi = \Phi(\mathbf{a}, c)$ to mean

$$\varphi(\omega) = \sum_{j=0}^{n-1} a_j (i\omega)^j + (1 + i\omega)^n \int_0^1 c(t) e^{i\omega t}\, dt, \tag{4.1}$$

where $\mathbf{a} = (a_0, \dots, a_{n-1})^T$. Ibragimov and Rozanov [(1978), page 30], show this mapping is onto $\mathscr{V}_R(F)$ under (3.2). Hence $P_R \varphi_{(r)}$ can be represented as in (4.1). If the function c in this representations is sufficiently smooth, we can bound the right-hand side of (3.1) using the results of Stein [(1990a), Section 4]. Lemma 2 below gives the critical needed bound on the smoothness of c as a function of r.

We first note some properties of the operator Φ. Define the Hilbert space $G = \mathbb{R}^n \times L^2([0,1])$ with inner product $\langle(\mathbf{a}_0, c_0), (\mathbf{a}_1, c_1)\rangle_G = \mathbf{a}_0^T\mathbf{a}_1 + \int_0^1 c_0(t)c_1(t)\,dt$. Then the linear operator $\Phi: G \to \mathscr{V}_R(F)$ is onto, bounded and invertible with bounded inverse. To prove this, first note that if we set $f_n(\omega) = (1+\omega^2)^{-n}$ and let F_n be the corresponding spectrum, $\|\Phi(\mathbf{a}, c)\|_{F_n}^2 = \|\Phi(\mathbf{a}, 0)\|_{F_n}^2 + \|\Phi(\mathbf{0}, c)\|_{F_n}^2$, since the cross term is 0 due to the analyticity in the upper half complex plane of $(i\omega)^j/(1-i\omega)^n$. By Parseval's relation, $\|\Phi(\mathbf{0}, c)\|_{F_n}^2 = (2\pi)^{-1}\|c\|^2$. Furthermore, $\|\Phi(\mathbf{a}, 0)\|_{F_n}^2/\mathbf{a}^T\mathbf{a}$ is uniformly bounded away from 0 and ∞. It follows that there exist positive finite constants u_n and v_n independent of f, $\mathbf{a}$ and c such that

$$C_0 u_n \le \frac{\|\Phi(\mathbf{a}, c)\|_F^2}{\mathbf{a}^T\mathbf{a} + \|c\|^2} \le C_1 v_n \tag{4.2}$$

as long as the denominator is positive. Define $\Phi^{-1}: \mathscr{V}_R(F) \to G$ to be the inverse of Φ and $\Phi^{-1} = (\Phi_1^{-1}, \Phi_2^{-1})$, so that $\Phi(\mathbf{a}, c) = \varphi$ implies $\Phi_1^{-1}(\varphi) = \mathbf{a}$ and $\Phi_2^{-1}(\varphi) = c$.

From (3.1) and Theorem 4.1 of Stein (1990a) we get

$$\begin{aligned}
\|\varphi_{(r)}\|_F^4 &\le P_R\varphi_{(r)} - P_N P_R \varphi_{(r)}\|_F^2 \|\varphi\|_F^2 \\
&\le \frac{2\pi C_1 e^2}{\{(n-1)!\}^2} \sum_{k=1}^{N} \left(\frac{t_{kN} - t_{k-1,N}}{2}\right)^2 \int_{t_{k-1,N}}^{t_{kN}} \left[\frac{d^n}{dt^n}\{c_r(t)e^{-t}\}\right]^2 dt \|\varphi\|_F^2 \\
&\ll_n C_1 S_N^{2n} \|\varphi\|_F^2 \sum_{j=0}^{n} \|c_r^{(j)}\|^2,
\end{aligned}$$

where $\ll_n$ is used to indicate that there exists a universal constant depending only on n such that the right-hand side times this constant is greater than or equal to the left-hand side. Theorem 3 results from combining the preceding bound and the following lemma.

Lemma 2. *If f satisfies (3.2) and K satisfies (3.3), then $c_r = \Phi_2^{-1}(P_R\varphi_{(r)}) \in W^{n,2}([0,1])$ for all $\varphi \in \mathscr{V}_R(F)$ and all $r > 0$. Furthermore, there exists a universal constant U_n depending only on n such that*

$$\sum_{j=0}^{n} \|c_r^{(j)}\| \le U_n \|\varphi_{(r)}\|_F \left\{ C_1^{1/2} \frac{r^n}{|q_0|} + \left(\frac{b_n(f)}{C_0}\right)^{1/2} \right\}. \tag{4.3}$$

Proof. For a function $\varphi \in \mathscr{V}(F)$ and $t \in \mathbb{R}$, define

$$m(t;\varphi) = \int_{-\infty}^{\infty} f(\omega) e^{-i\omega t} \varphi(\omega)\,d\omega.$$

The key step of the proof is to relate the smoothness of $m(\cdot; P_R\varphi_{(r)})$ to that of c_r.

Our first task is to show $c_r \in W^{n,2}([0,1])$. For this part of the proof, the specific value of r is irrelevant, so I will write $m(t)$ for $m(t; P_R \varphi_{(r)})$, which equals $m(t; \varphi_{(r)})$ on $[0,1]$. There exists $(\mathbf{a}, c) \in G$ such that, for $0 \le t \le 1$,

$$m(t) = \int_{-\infty}^{\infty} f(\omega) e^{-i\omega t} \left[\sum_{j=0}^{n-1} a_j (i\omega)^j + (1 + i\omega)^n \int_0^1 c(s) e^{i\omega s}\, ds \right] d\omega. \tag{4.4}$$

Consider first proving $c \in W^{n,2}([0,1])$ for $n \ge 2$. Then

$$m(t) = \sum_{j=0}^{n-1} (-1)^j a_j K^{(j)}(t) + \sum_{j=0}^{n} (-1)^j \binom{n}{j} \int_0^1 c(s) K^{(j)}(t-s)\, ds. \tag{4.5}$$

Define

$$\mu(t) = m^{(n-1)}(t) - \sum_{j=0}^{n-1} (-1)^j a_j K^{(j+n-1)}(t),$$

which is in $W^{n+1,2}([0,1])$ due to the analyticity of m and (3.3). Differentiating both sides of (4.5) and $n-1$ times and rearranging terms yields

$$\begin{aligned} \mu(t) &- \sum_{j=0}^{n-1} (-1)^j \binom{n}{j} \int_0^1 c(s) K^{(j+n-1)}(t-s)\, ds \\ &- (-1)^n \sum_{j=1}^{n_0} q_j (2j+1) \int_0^1 c(s)(t-s)^{2j} \operatorname{sgn}(t-s)\, ds \\ &- (-1)^n \int_0^1 c(s) g'(t-s)\, ds \\ &\quad = (-1)^n q_0 \frac{d}{dt} \int_0^1 c(s)|t-s|\, ds \\ &\quad = (-1)^n q_0 \int_0^1 c(s) \operatorname{sgn}(t-s)\, ds, \end{aligned}$$

where a sum whose upper limit is lower than its lower limit is taken to be 0. Since c is integrable on $[0,1]$, $\int_0^1 c(s)\operatorname{sgn}(t-s)\, ds = 2\int_0^t c(s)\, ds - \int_0^1 c(s)\, ds$ is absolutely continuous on $[0,1]$ with almost everywhere derivative $2c(t)$. Hence,

$$\begin{aligned} 2q_0 c(t) = (-1)^n \mu'(t) &- \sum_{j=0}^{n-1} \binom{n}{j} (-1)^{n-j} \int_0^1 c(s) K^{(j+n)}(t-s)\, ds \\ &- \sum_{j=1}^{n_0} q_j (2j+1) 2j \int_0^1 c(s) |t-s|^{2j-1}\, ds \qquad (4.6) \\ &- \int_0^1 c(s) g''(t-s)\, ds \end{aligned}$$

almost everywhere on $[0, 1]$, where $K^{(2n-1)}$ is the almost everywhere derivative of the absolutely continuous $K^{(2n-2)}$. The right-hand side of (4.6) is absolutely continuous on $[0, 1]$, so we may take c to be as well, in which case (4.6) holds everywhere on $[0, 1]$.

To complete the proof when $n = 2$, note that we now have

$$2q_0 c(t) = \mu'(t) - \int_0^1 c(s) K''(t-s)\,ds + 2\int_0^1 c(s) K^{(3)}(t-s)\,ds$$
$$-\int_0^1 c(s) g''(t+s)\,ds$$

with c and g'' absolutely continuous. Integration by parts yields

$$\int_0^1 c(s) g''(t-s)\,ds = -c(1) g'(t-1) + c(0) g'(t) + \int_0^1 c'(s) g'(t-s)\,ds$$

and hence

$$2q_0 c'(t) = \mu''(t) + c(1) g''(t-1) - c(0) g''(t) - \int_0^1 c'(s) g''(t-s)\,ds$$
$$-\int_0^1 c(s) K^{(3)}(t-s)\,ds + 4q_0 c(t) + 2\int_0^1 c(s) g''(t-s)\,ds.$$

so that c' is absolutely continuous on $[0, 1]$. Since c' is bounded on $[0, 1]$ and g'' is absolutely continuous on $[-1, 1]$, for almost every t, we can differentiate $\int_0^1 c'(s) g''(t-s)\,ds$ inside the integral [Graves (1956), page 215] to obtain

$$2q_0 c''(t) = \mu^{(3)}(t) + c(1) g^{(3)}(t-1) - c(0) g^{(3)}(t) - \int_0^1 c'(s) g^{(3)}(t-s)\,ds$$
$$- 2q_0 c(t) - \int_0^1 c(s) g''(t-s)\,ds$$
$$+ 4q_0 c'(t) + 2\int_0^1 c(s) g^{(3)}(t-s)\,ds$$

is the almost everywhere derivative of c' on $[0, 1]$. The right-hand side of this expression is square integrable, so $c \in W^{2,2}([0, 1])$ as required.

To complete the proof when $n = 3$, from (4.6),

$$2q_0 c(t) = -\mu'(t) + \sum_{j=0}^{2} (-1)^j \binom{3}{j} \int_0^1 c(s) K^{(j+3)}(t-s)\,ds$$
$$- 6q_1 \int_0^1 c(s) |t-s|\,ds - \int_0^1 c(s) g''(t-s)\,ds$$

with c absolutely continuous on $[0,1]$. It follows that

$$2q_0c''(t) = -\mu^{(3)}(t) + \int_0^1 c(s)K^{(5)}(t-s)\,ds + 6q_0c(t)$$
$$- 18q_1\int_0^1 c(s)|t-s|\,ds$$
$$- 3\int_0^1 c(s)g''(t-s)\,ds + 6q_0c'(t) + 18q_1\int_0^1 c(s)\operatorname{sgn}(t-s)\,ds$$
$$+ 3\int_0^1 c(s)g^{(3)}(t-s)\,ds - 12q_1c(t) - \int_0^1 c(s)g^{(4)}(t-s)\,ds.$$

After applying integration by parts to the last term on the right-hand side, the rest of the proof that $c \in W^{3,2}([0,1])$ proceeds as in the $n = 2$ case. The proof for $n \geq 4$ can be obtained by repeatedly applying the above arguments and is omitted.

When $n = 1$, more care is needed because K is not differentiable at 0. Define $\hat{c}(\omega) = \int_0^1 c(s)e^{i\omega s}\,ds$ and $m_1(t) = \int_0^t m(s)\,ds$. Then

$$m_1(t) = a_0\int_0^t K(s)\,ds + \int_0^t\int_{-\infty}^{\infty}(1+i\omega)f(\omega)e^{-i\omega s}\hat{c}(\omega)\,d\omega\,ds. \tag{4.7}$$

Now $(1+i\omega)f(\omega)$ and $\hat{c}(\omega)$ are both in $L^2(\mathbb{R})$, which implies $(1+i\omega)f(\omega)\hat{c}(\omega)$ is integrable, so the last integral in (4.7) equals

$$\int_{-\infty}^{\infty}(1+i\omega)f(\omega)\hat{c}(\omega)\left\{\int_0^t e^{-i\omega s}\,ds\right\}d\omega$$
$$= \int_{-\infty}^{\infty}\frac{1+i\omega}{i\omega}f(\omega)(1-e^{-i\omega t})\left\{\int_0^1 c(s)e^{i\omega s}\,ds\right\}d\omega$$
$$= \int_0^1 c(s)\left[\int_{-\infty}^{\infty}f(\omega)\left\{e^{i\omega s} - e^{i\omega(s-t)} + \int_{s-t}^{s}e^{i\omega u}\,du\right\}d\omega\right]ds$$
$$= \int_0^1 c(s)K(s)\,ds - \int_0^1 c(s)K(t-s)\,ds + \int_0^1 c(s)\int_{s-t}^{s}K(u)\,du\,ds.$$

Thus,

$$m_1(t) - a_0\int_0^t K(s)\,ds - \int_0^1 c(s)K(s)\,ds - \int_0^1 c(s)\int_{s-t}^{s}K(u)\,du\,ds$$
$$= -\int_0^1 c(s)K(t-s)\,ds.$$

Differentating both sides with respect to t and using $K(t) = q_0|t| + g(t)$ yields

$$m(t) - a_0K(t) - \int_0^1 c(s)K(t-s)\,ds$$
$$= -q_0\int_0^1 c(s)\operatorname{sgn}(t-s)\,ds - \int_0^1 c(s)g'(t-s)\,ds.$$

Thus, for almost every $t \in [0,1]$,

$$
\begin{aligned}
(4.8)\qquad 2q_0 c(t) &= -m'(t) + a_0\{q_0 + g'(t)\} - q_0 \int_0^1 c(s)\operatorname{sgn}(t-s)\,ds \\
&\quad + \int_0^1 c(s) g'(t-s)\,ds - \int_0^1 c(s) g''(t-s)\,ds.
\end{aligned}
$$

Since the right-hand side of (4.8) is continuous, we can set $2q_0 c(t)$ equal to the right-hand side of (4.8) for all $t \in [0,1]$ and hence take c continuous. However, c continuous on $[0,1]$ and g'' absolutely continuous on $[-1,1]$ imply the right-hand side of (4.8) is absolutely continuous with almost everywhere derivative

$$
\begin{aligned}
2q_0 c'(t) &= -m''(t) + a_0 g''(t) - 2q_0 c(t) \\
&\quad + \int_0^1 c(s) g''(t-s)\,ds - \int_0^1 c(s) g^{(3)}(t-s)\,ds,
\end{aligned}
$$

which is square integrable on $[0,1]$.

Now that we have $c \in W^{n,2}([0,1])$, using the representation for c in (4.6), I next prove (4.3) for $n \geq 2$; the case $n = 1$ is simpler and the proof is omitted. Since the dependence of various quantities on r is now critical, I now use the subscript r to indicate this dependence wherever it exists and set $\mathbf{a}_r = (a_{0r}, \dots, a_{n-1,r})^T$. From the definition of μ, for $0 \leq t \leq 1$,

$$
\mu_r^{(n+1)}(t) = m^{(2n)}(t; \varphi_{(r)}) - \sum_{j=0}^{n-1} (-1)^j a_{jr} K^{(j+2n)}(t),
$$

which is well defined due to (3.3), and

$$
\begin{aligned}
\left\| m^{(2n)}(\cdot\,; \varphi_{(r)}) \right\|^2 &= \int_0^1 \left| \int_{-r}^{r} \varphi(\omega)(i\omega)^{2n} f(\omega) e^{-i\omega t}\,d\omega \right|^2 dt \\
&\leq 2\pi \int_{-r}^{r} \left| \varphi(\omega)\,\omega^{2n} f(\omega) \right|^2 d\omega \\
&\leq 2\pi C_1 r^{2n} \|\varphi_{(r)}\|_F^2,
\end{aligned}
$$

where the first inequality uses Parseval's theorem. From (4.2),

$$
\mathbf{a}_r^T \mathbf{a}_r + \|c_r\|^2 \leq (C_0 u_n)^{-1} \|P_R \varphi_{(r)}\|_F^2 \leq (C_0 u_n)^{-1} \|\varphi_{(r)}\|_F^2,
$$

so that, by the definition of μ,

$$
(4.9)\qquad \|\mu_r^{(n+1)}\| \ll_n C_1^{1/2} r^n \|\varphi_{(r)}\|_F + C_0^{-1/2} \|\varphi_{(r)}\|_F \left(Q + \sum_{j=2}^{n+1} \|g^{(j)}\| \right).
$$

Next,

$$\left\| \frac{d^n}{dt^n} \left\{ \sum_{j=0}^{n-1} \binom{n}{j} (-1)^j \int_0^1 c_r(s) K^{(j+n)}(t-s)\, ds \right\} \right\| \ll_n Q \sum_{j=0}^{n-1} \|c_r^{(j)}\| + \|c_r\| \sum_{j=2}^{n+1} \|g^{(j)}\| \tag{4.10}$$

and

$$\left\| \frac{d^n}{dt^n} \left\{ \sum_{j=1}^{n_0} q_j (2j+1) 2j \int_0^1 c_r(s) |t-s|^{2j-1}\, ds \right\} \right\| \ll_n Q \sum_{j=0}^{n-2} \|c_r^{(j)}\|. \tag{4.11}$$

From Lemma 5.42 of Adams (1978) with $\delta = 2$ and $T = 1$, the Cauchy–Schwarz inequality and $(a^2 + b^2)^{1/2} \le |a| + |b|$,

$$\sup_{0 \le t \le 1} |u(t)| \le 2^{1/2} \|u\| + 2^{1/2} \|u\|^{1/2} \|u'\|^{1/2} \tag{4.12}$$

for any $u \in W^{1,2}([0,1])$. So,

$$\begin{aligned} \left\| \frac{d^n}{dt^n} \int_0^1 c_r(s) g''(t-s)\, ds \right\| &= \left\| \frac{d}{dt} \{ c_r(0) g^{(n)}(t) - c_r(1) g^{(n)}(t-1) \right. \\ &\qquad \left. + \int_0^1 c_r'(s) g^{(n)}(t-s)\, ds \} \right\| \\ &\ll_n \left(|c_r(0)| + |c_r(1)| + \|c_r'\| \right) \|g^{(n+1)}\| \\ &\ll_n \left(\|c_r\| + \|c_r'\| \right) \|g^{(n+1)}\|, \end{aligned} \tag{4.13}$$

where the last step makes use of (4.12). Applying (4.9)–(4.11) and (4.13) to the nth derivative of (4.6) yields

$$\begin{aligned} |q_0| \sum_{j=0}^{n} \|c_r^{(j)}\| &\ll_n \|\varphi_{(r)}\|_F \left\{ C_1^{1/2} r^n + C_0^{-1/2} \left(Q + \sum_{j=2}^{n+1} \|g^{(j)}\| \right) \right\} \\ &\quad + Q \sum_{j=0}^{n-1} \|c_r^{(j)}\| + \left(\|c_r\| + \|c_r'\| \right) \sum_{j=2}^{n+1} \|g^{(j)}\|. \end{aligned} \tag{4.14}$$

From (4.12), for any $u \in W^{2,2}([0,1])$,

$$\begin{aligned} \|u'\|^2 &= u'(1)u(1) - u'(0)u(0) - \int_0^1 u(t) u''(t)\, dt \\ &\le 2^{3/2} \{ \|u\| + \|u\|^{1/2} \|u'\|^{1/2} \} \{ \|u'\| + \|u'\|^{1/2} \|u''\|^{1/2} \} + \|u\|\, \|u''\|, \end{aligned}$$

which can be used to show that, for example,

$$\|u'\| / \|u\| \ge 100 \quad \text{implies} \quad \|u''\| / \|u'\| \ge 10 \|u'\| / \|u\|. \tag{4.15}$$

We next bound $\sum_{j=0}^{n-1}\|c_r^{(j)}\|$ in terms of $\|c_r\|$ and $\|c_r^{(n)}\|$. If $\|c_r\| = 0$, then $\sum_{j=0}^{n-1}\|c_r^{(j)}\| = 0$, so suppose otherwise. Define $A = \max(10^{n+2}, (\|c_r^{(n)}\|/\|c_r\|)^{1/n})$ and $B = \max_{1\le j\le n-1}(\|c_r^{(j)}\|/\|c_r\|)^{1/j}$. If $B \le A$, $\sum_{j=0}^{n-1}\|c_r^{(j)}\| \le nA^{n-1}\|c_r\|$. If $B > A$, then define j_0 to be the smallest value of j for which $(\|c_r^{(j)}\|/\|c_r\|)^{1/j} \ge A$, with j_0 necessarily satisfying $1 \le j_0 \le n-1$. From (4.15), for $j_0 \le j \le n-1$, $\|c_r^{(j)}\| \le \|c_r^{(n-1)}\| \le 10^{n-j_0}A^{-1}\|c_r^{(n)}\|$, so that, for $B > A$,

$$\sum_{j=0}^{n-1}\|c_r^{(j)}\| = \sum_{j=0}^{j_0-1}\|c_r^{(j)}\| + \sum_{j=j_0}^{n-1}\|c_r^{(j)}\| \le j_0A^{j_0-1} + (n-j_0)10^{n-j_0}A^{-1}\|c_r^{(n)}\|.$$

Considering the two cases $B \le A$ and $B > A$ separately yields

$$\sum_{j=0}^{n-1}\|c_r^{(j)}\| \ll_n A^{n-1}\|c_r\| + A^{-1}\|c_r^{(n)}\| \ll_n \|c_r\| + \|c_r\|^{1/n}\|c_r^{(n)}\|^{(n-1)/n}. \tag{4.16}$$

Using a similar argument to bound $\|c_r'\|$ gives

$$\|c_r'\| \ll_n \|c_r\| + \|c_r\|^{(n-1)/n}\|c_r^{(n)}\|^{1/n} \tag{4.17}$$

and (4.2) implies

$$\|c_r\| \ll_n C_0^{-1/2}\|\varphi_{(r)}\|_F. \tag{4.18}$$

Applying (4.16)–(4.18) to (4.14) yields

$$\begin{aligned}|q_0|\sum_{j=0}^{n}\|c_r^{(j)}\| \ll_n\ & C_1^{1/2}\|\varphi_{(r)}\|_F r^n + \|\varphi_{(r)}\|_F C_0^{-1/2}\left(Q + \sum_{j=2}^{n+1}\|g^{(j)}\|\right)\\ &+ QC_0^{-1/(2n)}\|\varphi_{(r)}\|_F^{1/n}\|c_r^{(n)}\|^{(n-1)/n}\\ &+ C_0^{-(n-1)/(2n)}\|\varphi_{(r)}\|_F^{(n-1)/n}\|c_r^{(n)}\|^{1/n}\sum_{j=2}^{n+1}\|g^{(j)}\|.\end{aligned}$$

The left-hand side is at most 4 times the maximum of the four summands on the right-hand side. If the third summand is the maximum, $\sum_{j=0}^{n}\|c_r^{(j)}\| \ll_n |Q/q_0|^n C_0^{-1/2}\|\varphi_{(r)}\|_F$; and if the fourth summand is the maximum, then

$$\sum_{j=0}^{n}\|c_r^{(j)}\| \ll_n C_0^{-1/2}\|\varphi_{(R)}\|_F\left\{|q_0|^{-1}\sum_{j=0}^{n+1}\|g^{(j)}\|\right\}^{n/(n-1)}.$$

Equation (4.3) follows for $n \ge 2$. □

APPENDIX

Sacks and Ylvisaker (1966, 1968, 1970, 1971), Cambanis (1985), Wahba (1974), Eubank, Smith and Smith (1982) and Ritter (1995, 1996) studied the problem of choosing locations for observing a continuous stochastic process on $[0, 1]$ to obtain asymptotically optimal estimates of regression coefficients. When there is a single regression coefficient, the setup is to consider a process of the form $Y(t) = \beta x(t) + Z(t)$, where x is a known function, β an unknown scalar and Z a mean 0 stochastic process with $\operatorname{cov}(Z(s), Z(t)) = K(s, t)$. The basic assumption used about x in all of these works is that it can be written in the form

$$x(t) = \int_0^1 K(s,t)\eta(s)\,ds \tag{A.1}$$

for $t \in [0,1]$, under some condition on η, often continuity. Cambanis (1985) notes that this representation also arises when studying a problem in signal detection. When Z is m-fold integrated Brownian motion, then if $x \in W^{2m,2}([0,1])$ and $x^{(j)}(1) = 0$ for $j = m+1, \dots, 2m+1$, we can take $\eta(t) = (-1)^{m+1}x^{(2m-2)}(t)$ in (A.1) [Eubank, Smith and Smith (1982)]. Except for some minor extensions to this case described by Eubank, Smith and Smith (1982), there are no results showing when x satisfies (A.1) in these works. The proof of Lemma 2 provides a method of demonstrating that a representation such as in (A.1) exists, at least for stationary Z.

Suppose Z is stationary and $K(s, t) = K(s - t)$ satisfies (3.3). To avoid the problem of boundary conditions, consider the slightly more general representation

$$x(t) = \int_0^1 K(s-t)\eta(s)\,ds + \sum_{j=0}^{n-1}\left\{a_j K^{(j)}(t) + b_j K^{(j)}(1-t)\right\}. \tag{A.2}$$

As noted by Sacks and Ylvisaker (1971) and Cambanis (1985), including these extra terms in the representation will not change any of the results on asymptotically optimal designs. Examining the proof of Lemma 2, we see that any function $x \in W^{2n,2}([0,1])$ has a representation as in (4.4) with $c \in W^{n,2}([0,1])$. Defining $\bar{c}(t) = c(t)e^{-t}$, it follows by straightforward calculations using integration by parts that

$$\begin{aligned} x(t) = {} & \int_0^1 K(t-s)(-1)^n \bar{c}^{(n)}(s)e^s\,ds \\ & + \sum_{k=0}^{n-1}\sum_{j=0}^{n-1-k}(-1)^j\binom{n-1-j}{k} \\ & \qquad \times\left\{e\bar{c}^{(j)}(1)K^{(k)}(1-t) - (-1)^k\bar{c}^{(j)}(0)K^{(k)}(t)\right\}, \end{aligned}$$

which is of the form (A.2) with $\eta(t) = (-1)^n\bar{c}^{(n)}(t)e^t$. If, as in Sacks and Ylvisaker (1970) or Cambanis (1985), we want η to be continuous, then it suffices to assume that $x^{(2n)}$ exists and is continuous on $[0,1]$ and that K is of the form given in (3.3) where $g^{(n+1)}$ exists and is continuous on $[-1,1]$.

These results can be used to show that Corollaries 1 and 2 apply to ordinary kriging predictors when K_0 and K_1 satisfy (3.3), f_0 and f_1 satisfy (3.2) and the observations are as in Theorem 3 with $S_N = O(N^{-1})$. Let $\tilde{e}_i(h, N)$ be the error of the ordinary kriging predictor of h under K_i. Since K_1 satisfies (3.3), the results of the preceding paragraph imply the function $x(t) = 1$ on $[0,1]$ can be written as in (4.4) with $c \in W^{n,2}([0,1])$. Considering the discussion in the paragraph preceding Theorem 4.1 of Stein (1990a), Theorems 4.1 and 5.2 of Stein (1990a) imply

$$\sup_{h \in H_{-N}(F_i)} \frac{E_i\{\tilde{e}_i(h, N) - e_i(h, N)\}^2}{E_i e_i(h, N)^2} = O(N^{-2n}) \tag{A.3}$$

for $i = 0, 1$. Since f_0 and f_1 satisfy (3.2),

$$E_0(\tilde{e}_1 - \tilde{e}_0) \le 3\frac{C_1}{C_0}E_1(\tilde{e}_1 - e_1)^2 + 3E_0(e_1 - e_0)^2 + 3E_0(\tilde{e}_0 - e_0)^2$$

and

$$\begin{aligned} \sup_{h \in H_{-N}(F_0)} & \frac{E_0\{\tilde{e}_1(h, N) - \tilde{e}_0(h, N)\}^2}{E_0 \tilde{e}_0(h, N)^2} \\ &= O\left(N^{-\min(4\gamma n/(2n+\gamma), 2n)}(\log N)^{1\{\gamma = 2n\}}\right) \end{aligned} \tag{A.4}$$

follows from (3.9) and (A.3). next,

$$\begin{aligned} |E_1\tilde{e}_1^2 - E_0\tilde{e}_1^2| &\le |E_1\tilde{e}_1^2 - E_1 e_1^2| + |E_1 e_1^2 - E_0 e_1^2| \\ &\quad + |E_0 e_1^2 - E_0 e_0^2| + |E_0 e_0^2 - E_0\tilde{e}_0^2| + |E_0\tilde{e}_0^2 - E_0\tilde{e}_1^2| \\ &\le E_1(\tilde{e}_1 - e_1)^2 + |E_1 e_1^2 - E_0 e_1^2| \\ &\quad + E_0(e_1 - e_0)^2 + E_0(\tilde{e}_0 - e_0)^2 + E_0(\tilde{e}_1 - \tilde{e}_0)^2 \end{aligned}$$

and

$$\sup_{h \in H_{-N}(F_1)} \frac{\left|E_1\tilde{e}_1(h, N)^2 - E_0\tilde{e}_1(h, N)^2\right|}{E_0\tilde{e}_1(h, N)^2} = O\left(N^{-\min(2n, \gamma)}(\log N)^{1\{\gamma = 2n\}}\right)$$

follows from (3.9), (3.10), (A.3) and (A.4).

REFERENCES

ADAMS, R. A. (1978). *Sobolev Spaces*. Academic Press, San Diego.

CAMBANIS, S. (1985). Sampling designs for time series. In *Time Series in the Time Domain. Handbook of Statistics* (E. J. Hannan, P. R. Krishnaiah and M. M. Rao, eds.) **5** 337–362. North-Holland, Amsterdam.

CARR, J. R. (1990). Application of spatial filter theory to kriging. *Math. Geol.* **22** 1063–1079.

CHRISTAKOS, G. (1992). *Random Field Models in Earth Sciences*. Academic Press, San Diego.

CHRISTENSEN, R. (1991). *Linear Models for Multivariate, Time Series, and Spatial Data*. Springer, New York.

CLEVELAND, W. S. (1971). Projection with the wrong inner product and its application to regression with correlated errors and linear filtering of time series. *Ann. Math. Statist.* **42** 616–624.

CRESSIE, N. A. C. (1993). *Statistics for Spatial Data*, rev. ed. Wiley, New York.

DALEY, R. (1991). *Atmospheric Data Analysis*. Cambridge Univ. Press.

EUBANK, R. L. SMITH, P. L. and SMITH, P. W. (1982). A note on optimal and asymptotically optimal designs for certain time series models. *Ann. Statist.* **10** 1295–1301.

GRAVES, L. M. (1956). *The Theory of Functions of Real Variables*, 2nd ed. McGraw-Hill, New York.

HANDCOCK, M. S. and STEIN, M. L. (1993). A Bayesian analysis of kriging. *Technometrics* **35** 403–410.

HANDCOCK, M. S. and WALLIS, J. R. (1994). An approach to statistical spatial–temporal modeling of meteorological fields (with discussion). *J. Amer. Statist. Assoc.* **89** 368–390.

HANNAN, E. J. (1970). *Multiple Time Series*. Wiley, New York.

IBRAGIMOV, I. A. and ROZANOV, Y. A. (1978). *Gaussian Random Processes*. Springer, New York.

JERRI, A. J. (1977). The Shannon sampling theorem—its various extensions and applications: a tutorial review. *Proc. IEEE* **65** 1565–1596.

RITTER, K. (1995). Average case analysis of numerical problems. Thesis, Univ. Erlangen, Germany.

RITTER, K. (1996). Asymptotic optimality of regular sequence designs. *Ann. Statist.* **24** 2081–2096.

ROBINSON, G. K. (1991). That BLUP is a good thing: the estimation of random effects (with discussion). *Statist. Sci.* **6** 15–51.

SACKS, J. and YLVISAKER, D. (1966). Designs for regression problems with correlated errors. *Ann. Math. Statist.* **37** 66–89.

SACKS, J. and YLVISAKER, D. (1968). Designs for regression problems with correlated errors: many parameters. *Ann. Math. Statist.* **39** 49–69.

SACKS, J. and YLVISAKER, D. (1970). Designs for regression problems with correlated errors III. *Ann. Math. Statist.* **41** 2057–2074.

SACKS, J. and YLVISAKER, D. (1971). Statistical designs and integral approximation. In *Proceedings of the Twelfth Biennial Canadian Math. Society Seminar* (R. Pyke, ed.) 115–136. Canadian Math. Congress, Montreal.

STARKS, T. H. and SPARKS, A. R. (1987). Rejoinder to "Comment on 'Estimation of the generalized covariance function. II. A response surface approach' by Thomas H. Starks and Allen R. Sparks." *Math. Geol.* **19** 789–792.

STEIN, M. L. (1990a). Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure. *Ann. Statist.* **18** 850–872.

STEIN, M. L. (1990b). Bounds on the efficiency of linear predictions using an incorrect covariance function. *Ann. Statist.* **18** 1116–1138.

STEIN, M. L. (1993). A simple condition for asymptotic optimality of linear predictions of random fields. *Statist. Probab. Lett.* **17** 399–404.

STEIN, M. L. (1997). Efficiency of linear predictors for periodic processes using an incorrect covariance function. *J. Statist. Plann. Inference* **58** 321–331.

WAHBA, G. (1974). Regression design for some equivalence classes of kernels. *Ann. Statist.* **2** 925–934.

YAGLOM, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions* **1**. Springer, New York.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
5734 UNIVERSITY AVE.
CHICAGO, ILLINOIS 60637
E-MAIL: stein@galton.uchicago.edu