The Annals of Statistics
2011, Vol. 39, No. 6, 2795–2819
DOI: 10.1214/11-AOS909

2010 RIETZ LECTURE

WHEN DOES THE SCREENING EFFECT HOLD?

BY MICHAEL L. STEIN[1]

University of Chicago

When using optimal linear prediction to interpolate point observations of a mean square continuous stationary spatial process, one often finds that the interpolant mostly depends on those observations located nearest to the predictand. This phenomenon is called the screening effect. However, there are situations in which a screening effect does not hold in a reasonable asymptotic sense, and theoretical support for the screening effect is limited to some rather specialized settings for the observation locations. This paper explores conditions on the observation locations and the process model under which an asymptotic screening effect holds. A series of examples shows the difficulty in formulating a general result, especially for processes with different degrees of smoothness in different directions, which can naturally occur for spatial-temporal processes. These examples lead to a general conjecture and two special cases of this conjecture are proven. The key condition on the process is that its spectral density should change slowly at high frequencies. Models not satisfying this condition of slow high-frequency change should be used with caution.

1. Introduction. The screening effect is the geostatistical term for the phenomenon of nearby observations tending to reduce the influence of more distant observations when using kriging (optimal linear prediction) for spatial interpolation [Journel and Huijbregts (1978), Chilès and Delfiner (1999)]. This phenomenon is often invoked as a justification for ignoring more distant observations when using kriging [Memarsadeghi and Mount (2007), Emery (2009)]. Only in some very limited special cases is the effect exact in the sense that the more distant observations make no contribution to the kriging predictor, so it is natural to use asymptotics as a way to study the screening effect.

Let us set some notation. Write $x \cdot y$ for the inner product of commensurate vectors x and y. Suppose Z is a mean square continuous, stationary, mean 0 Gaussian process on $\mathbb{R}^d$ with autocovariance function $K(x) = E\{Z(x)Z(0)\}$ and spectral density f, so that $K(x) = \int_{\mathbb{R}^d} e^{i\omega\cdot x} f(\omega)\,d\omega$. When the mean is assumed known to be 0, kriging is often called simple kriging. Throughout this work, we assume that the problem of interest is to predict $Z(0)$. For $S \subset \mathbb{R}^d$, write $Z(S)$ for the vector of

Received November 2010; revised April 2011.

[1]Supported by US Department of Energy Grant DE-SC0002557.

MSC2010 subject classifications. Primary 60G25; secondary 62M30, 62M15.

Key words and phrases. Space–time process, spectral analysis, kriging, fixed-domain asymptotics.

observations (in some order) of Z on S, and define $e(S)$ to be the error of the best linear predictor, or BLP, of $Z(0)$ based on $Z(S)$. Let N_ε and F_ε be two classes of sets indexed by the parameter $\varepsilon > 0$, with N_ε representing observations near 0 and F_ε more distant observations. We will say that N_ε asymptotically screens out the effect of F_ε if

$$\lim_{\varepsilon \downarrow 0} \frac{Ee(N_\varepsilon \cup F_\varepsilon)^2}{Ee(N_\varepsilon)^2} = 1. \tag{1}$$

Stein (2002) argues that a useful asymptotic approach is to let the smallest distance from the observations to the predictand tend to 0 as $\varepsilon \downarrow 0$. Specifically, Stein (2002) proves (1) when, essentially, for some $x_0 \in \mathbb{R}^d$ not in the integer lattice, F_ε is all points of the form $\varepsilon(x_0 + J)$ for J in the integer lattice, N_ε is the restriction of F_ε to some fixed region with 0 in its interior and f is regularly varying at infinity [Bingham, Goldie and Teugels (1987)] in every direction with a common index of variation. The methods used in Stein (2002) make strong use of the gridded nature of the observations and are not applicable here. Furthermore, requiring f to be regularly varying at infinity with common index of variation in all directions excludes models for spatial-temporal phenomena that exhibit a different degree of smoothness in space than in time. Section 4 provides further discussion of these issues. Ramm (2005), Chapter 5, takes a different approach to studying an asymptotic screening effect by considering a process observed with white noise everywhere in some domain and letting the variance of the white noise tend to 0. In this work, we take a closer look at how the set where Z is observed affects whether an asymptotic screening effect holds.

We will take the sets N_ε and F_ε to have a particular form that simplifies the asymptotic analysis. Suppose $x_1, \dots, x_n$ are distinct nonzero elements of $\mathbb{R}^d$, $y_1, \dots, y_m$ are distinct elements of $\mathbb{R}^d$ and $y_0 \in \mathbb{R}^d$ is nonzero. For the rest of this work, let $N_\varepsilon = \{\varepsilon x_1, \dots, \varepsilon x_n\}$ and $F_\varepsilon = \{y_0 + \varepsilon y_1, \dots, y_0 + \varepsilon y_m\}$. Section 2 explores when (1) holds through a series of examples leading to a broad conjecture under a key assumption on the spectral density f of the random field: for every $R < \infty$,

$$\lim_{\omega \to \infty} \sup_{|\nu| < R} \left| \frac{f(\omega + \nu)}{f(\omega)} - 1 \right| = 0. \tag{2}$$

The examples will demonstrate that one generally needs a further condition on N_ε depending on the mean square differentiability properties of the process. For nondifferentiable processes, no further assumptions on N_ε may be needed. Indeed, for nondifferentiable processes on $\mathbb{R}$, Theorem 1 in Section 3 has (1) as its conclusion under (2) and a mild additional condition on f. For nondifferentiable processes on $\mathbb{R}^2$, if one restricts the cardinality of N_ε to 1 and of F_ε to 2 (and sets $y_2 = 0$), then Theorem 2 proves (1) under (2) without any additional conditions on f.

Matérn models [Stein (1999a)] appear in both the examples and the proof of Theorem 1. Define $\mathcal{K}_\nu$ to be the modified Bessel function of the second kind of order ν [Olver et al. (2010)]. The Matérn model on $\mathbb{R}^d$ has autocovariance function $\phi(\alpha|x|)^\nu \mathcal{K}_\nu(\alpha|x|)$ for positive ϕ, α and ν. The parameter ν controls the smoothness of the process: Z has m mean square derivatives in any direction if and only if $\nu > m$. The corresponding spectral density equals $\phi(\alpha^2 + |\omega|^2)^{-\nu-d/2}$ times a constant depending on α, ν and d. All Matérn models satisfy (2).

2. Examples. This section studies a number of examples to gain some insight into the conditions on f and N_ε that are needed in order for (1) to hold. The derivations of these results are elementary but not necessarily easy. Rather than give detailed derivations of all of them, I will outline derivations in a few of the more difficult examples in Section 5.1.

To see why a condition like (2) is needed, let us first consider an example on $\mathbb{R}$ addressed in Stein and Handcock (1989) and Stein (1999a), pages 67–69. Suppose $n = 1$, $x_1 = 1$, $m = 2$, $y_0 = 1$, $y_1 = 0$ and $y_2 = 1$; see Figure 1. Consider $K(x) = e^{-|x|}$, a Matérn model with smoothness parameter $\frac{1}{2}$. The corresponding process is mean square continuous but is not mean square differentiable, and it is easy to show $Ee(N_\varepsilon)^2 \sim 2\varepsilon$ as $\varepsilon \downarrow 0$. This process is Markov, so that $Ee(N_\varepsilon \cup F_\varepsilon)^2 = Ee(N_\varepsilon)^2$ for all $\varepsilon < 1$ and (1) holds trivially. Next consider $K(x) = (1 - |x|)^+$ (where the superscipt $+$ indicates positive part), for which $f(\omega) = \frac{1-\cos\omega}{\pi\omega^2}$, which does not satisfy (2). Stein and Handcock (1989), page 180, give the BLP based on $Z(N_\varepsilon \cup F_\varepsilon)$, from which it is not difficult to show that $Ee(N_\varepsilon)^2 \sim 2\varepsilon$, just like for $K(x) = e^{-|x|}$, but $Ee(N_\varepsilon \cup F_\varepsilon)^2 \sim \frac{3}{2}\varepsilon$ as $\varepsilon \downarrow 0$ so that $Ee(N_\varepsilon \cup F_\varepsilon)^2 / Ee(N_\varepsilon)^2 \to \frac{3}{4}$ as $\varepsilon \downarrow 0$. The choice of $y_0 = 1$ is critical here: for $y_0 \neq 1$ but positive (keeping $x_1 = 1$, $y_1 = 0$, $y_2 = 1$), $Ee(N_\varepsilon \cup F_\varepsilon)^2 / Ee(N_\varepsilon)^2 \to 1$ as $\varepsilon \downarrow 0$. The anomaly for $y_0 = 1$ is related to the lack of differentiability of $K(x)$ at $x = 1$, which is in turn related to the oscillations at high frequencies in f. See Stein (2005) for further discussion on the relationship of the differentiability of K away from the origin and the high-frequency behavior of f.

Proposition 1 in Stein (2005) provides a second example showing why a condition like (2) is needed to have a screening effect. The following special case of this result suffices to illustrate the point. Suppose Z is a stationary process on $\mathbb{R}^2$ with autocovariance function $K(s,t) = e^{-|s|-|t|}$ for $s, t \in \mathbb{R}$. The corresponding spectral density $f(\omega_1, \omega_2)$ is proportional to $\frac{1}{(1+\omega_1)^2(1+\omega_2^2)}$, which does not satisfy (2).

FIG. 1. *Prediction problem for triangular autocovariance function. Prediction site* (+ *sign*), *nearby observation* (*solid circle*) *and distant observations* (*open circles*).

FIG. 2. *Prediction problem for autocovariance function* $K(s,t) = e^{-|s|-|t|}$. *Symbols as in Figure* 1.

Consider the situation pictured in Figure 2, for which $x_1 = (0,1)$, $y_0 = (1,0)$, $y_1 = (0,1)$ and $y_2 = (0,0)$. Then using either direct calculation or Proposition 1 in Stein (2005), $\lim_{\varepsilon \downarrow 0} Ee(N_\varepsilon \cup F_\varepsilon)^2 / Ee(N_\varepsilon)^2 = 1 - e^{-2}$.

The remaining examples all consider f satisfying (2). To see why an additional condition on N_ε is needed for (1) to hold for differentiable processes, consider a Matérn model with smoothness parameter $\frac{3}{2}: K(x) = e^{-|x|}(1+|x|)$, for which the corresponding process is exactly once mean square differentiable. For $N_\varepsilon = \{\varepsilon\}$, $F_\varepsilon = \{1\}$ (top plot in Figure 3), straightforward calculations yield $Ee(N_\varepsilon)^2 \sim \varepsilon^2$ and $Ee(N_\varepsilon \cup F_\varepsilon)^2 \sim \frac{e^2-5}{e^2-4}\varepsilon^2$ as $\varepsilon \downarrow 0$ so $Ee(N_\varepsilon \cup F_\varepsilon)^2 / Ee(N_\varepsilon)^2 \to \frac{e^2-5}{e^2-4}$ as $\varepsilon \downarrow 0$. Unlike the triangular case, there is nothing special about $y_0 = 1$ here and the more general result for $y_0 > 0$ is $Ee(N_\varepsilon \cup F_\varepsilon)^2 / Ee(N_\varepsilon)^2 \to 1 - y_0^2/(e^{2y_0} - 1 - 2y_0 - y_0^2)$. The reason the limit is less than 1 is not because there is anything unusual about f, but rather that N_ε is inadequate. Specifically, since $Z(0) = Z(\varepsilon) - \varepsilon Z'(0) + o_p(\varepsilon)$ and $\operatorname{cov}\{Z(\varepsilon), Z'(0)\} \to 0$ as $\varepsilon \downarrow 0$, it is apparent that having even a somewhat informative predictor for $Z'(0)$ would provide useful information about $Z(0)$ not contained in $Z(\varepsilon)$. In fact, as $\varepsilon \downarrow 0$, it is possible to show that $\widehat{Z'(0)} = \frac{e}{e^2-4}Z(\varepsilon) - \frac{2}{e^2-4}Z(1)$ is an asymptotically optimal predictor of $Z'(0)$ based on $(Z(\varepsilon), Z(1))$ and, in turn, that $Z(\varepsilon) - \varepsilon\widehat{Z'(0)}$ is an asymptotically optimal predictor of $Z(0)$ based on $(Z(\varepsilon), Z(1))$. A screening effect does hold if 2ε is added to N_ε (bottom plot of Figure 3). Then it is possible to show that $Ee(N_\varepsilon \cup F_\varepsilon)^2 \sim Ee(N_\varepsilon)^2 \sim \frac{8}{3}\varepsilon^3$ as $\varepsilon \downarrow 0$, so (1) is true. Furthermore, as $\varepsilon \downarrow 0$, $2Z(2\varepsilon) - Z(\varepsilon) = Z(\varepsilon) - \varepsilon[\{Z(2\varepsilon) - Z(\varepsilon)\}/\varepsilon]$ is an asymptotically optimal pre-

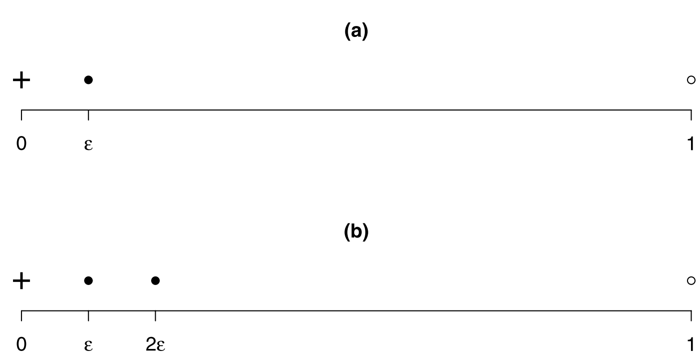

FIG. 3. *Prediction problems for Matérn model with* $\nu = \frac{3}{2}$ *on* $\mathbb{R}$. *Symbols as in Figure* 1.

dictor of $Z(0)$ based on $Z(N_\varepsilon \cup F_\varepsilon)$ and $\{Z(2\varepsilon) - Z(\varepsilon)\}/\varepsilon$ is a consistent predictor of $Z'(0)$. A reasonable conjecture for a process on $\mathbb{R}$ with exactly p mean square derivatives whose spectral density satisfies (2) is that any distinct $x_1, \dots, x_n$ with $n > p$ suffices to make (1) true.

It is helpful to consider this problem in the spectral domain. We need some further notation to proceed. For nonnegative-valued functions a and b defined on a common domain D, write $a(x) \ll b(x)$ if there exists finite C such that $a(x) \le Cb(x)$ for all $x \in D$ and, for $x \in \mathbb{R}$, $a(x) \ll b(x)$ as $x \downarrow x_0$ if, for some $c > 0$, $a(x) \ll b(x)$ for $D = (x_0, x_0 + c)$. Write $a(x) \asymp b(x)$ if $a(x) \ll b(x)$ and $b(x) \ll a(x)$ and define $a(x) \asymp b(x)$ as $x \downarrow 0$ if $a(x) \ll b(x)$ as $x \downarrow x_0$ and $b(x) \ll a(x)$ as $x \downarrow x_0$. For a complex-valued function g and a nonnegative function f defined on a domain D (always $\mathbb{R}^d$ here), define $\|g\|_f = \sqrt{\int_D |g(x)|^2 f(x)\,dx}$. To each random variable of the form $\sum_{j=1}^n \lambda_j Z(s_j)$ there is a corresponding function $\sum_{j=1}^n \lambda_j e^{i\omega\cdot s_j}$, and the mapping is an isometric isomorphism in the sense that $E\{\sum_{j=1}^n \lambda_j Z(s_j)\}^2 = \int_{\mathbb{R}^d} |\sum_{j=1}^n \lambda_j e^{i\omega\cdot s_j}|^2 f(\omega)\,d\omega$. Write $\sum_{j=1}^n \phi_{j\varepsilon} Z(\varepsilon x_j)$ for the BLP of $Z(0)$ based on $Z(N_\varepsilon)$ and $\phi_\varepsilon(\omega) = \sum_j \phi_{j\varepsilon} e^{i\varepsilon\omega\cdot x_j}$ for the corresponding function. If we set $\eta_\varepsilon(\omega) = 1 - \phi_\varepsilon(\omega)$, then $Ee(N_\varepsilon)^2 = \|\eta_\varepsilon\|_f^2$.

For any $A \subset \mathbb{R}^d$, call $\int_A |\eta_\varepsilon(\omega)|^2 f(\omega)\,d\omega / \|\eta_\varepsilon\|_f^2$ the fraction of $Ee(N_\varepsilon)^2$ attributable to the set of frequencies A. Write $b(r)$ for the ball of radius r centered at the origin. For the scenario in Figure 3(a), for any fixed $\omega_0 > 0$, as $\varepsilon \downarrow 0$,

$$\frac{\int_{b(\omega_0)} |\eta_\varepsilon(\omega)|^2 f(\omega)\,d\omega}{\|\eta_\varepsilon\|_f^2} \sim \frac{2}{\pi}\left\{\tan^{-1}\omega_0 - \frac{\omega_0}{1+\omega_0^2}\right\} > 0 \tag{3}$$

so that an asymptotically nonnegligible fraction of $Ee(N_\varepsilon)^2$ is attributable to a fixed range of frequencies. Similar to the definition of η_ε, let ψ_ε be the function corresponding to $e(N_\varepsilon \cup F_\varepsilon)$, so that $Ee(N_\varepsilon \cup F_\varepsilon)^2 = \|\psi_\varepsilon\|_f^2$. Then (3) allows $Z(1)$ to improve the prediction nonnegligibly by making $|\psi_\varepsilon(\omega)|^2/|\eta_\varepsilon(\omega)|^2$ substantially smaller than 1 in a neighborhood of the origin. In contrast, for the scenario in Figure 3(b), $\int_{b(\omega_0)} |\eta_\varepsilon(\omega)|^2 f(\omega)\,d\omega \ll \varepsilon^4$ as $\varepsilon \downarrow 0$ for any fixed ω_0, so that $\int_{b(\omega_0)} |\eta_\varepsilon(\omega)|^2 f(\omega)\,d\omega \ll \varepsilon\|\eta_\varepsilon\|_f^2$ as $\varepsilon \downarrow 0$. In this case, making $|\psi_\varepsilon(\omega)|^2/|\eta_\varepsilon(\omega)|^2$ substantially smaller than 1 in a neighborhood of the origin cannot yield a nonnegligible asymptotic impact on the mean squared prediction error. Thus, $\int_{b(\omega_0)^c} |\psi_\varepsilon(\omega)|^2 f(\omega)\,d\omega / \int_{b(\omega_0)^c} |\eta_\varepsilon(\omega)|^2 f(\omega)\,d\omega$ must be bounded by some constant less than 1 as $\varepsilon \downarrow 0$ for all ω_0 for (1) not to hold. The fact that f is well behaved at high frequencies [i.e., satisfies (2)] effectively precludes this possibility so that (1) holds. This line of reasoning forms the basis of the proof of Theorem 1; see Section 5.2.

It is interesting to reconsider the two cases pictured in Figure 3, for a process that is not quite mean square differentiable: $K(x) = |x|\mathcal{K}_1(|x|)$, a Matérn model with smoothness parameter 1, for which $K(x) = 1 + \frac{1}{2}x^2\log(\frac{1}{2}|x|) + \frac{1}{4}(2\gamma -$

$1)x^2 + O(x^4 \log|x|)$ as $x \to 0$ with γ being Euler's constant. The corresponding spectral density f is proportional to $(1+\omega^2)^{-3/2}$. Since the process has no mean square derivatives, I conjecture that (1) should hold for any nonempty N_ε. For the scenario in Figure 3(a), $Ee(N_\varepsilon)^2 \sim -\varepsilon^2 \log \varepsilon$ and, for fixed $\omega_0 > 0$,

$$\int_{b(\omega_0)} |\eta_\varepsilon(\omega)|^2 f(\omega)\, d\omega \asymp \int_{b(\omega_0)} \frac{|1 - e^{i\varepsilon\omega}|^2 + \{K(\varepsilon) - 1\}^2}{(1+\omega^2)^{3/2}}\, d\omega \asymp \varepsilon^2$$

as $\varepsilon \downarrow 0$. Thus, the fraction of the $Ee(N_e)^2$ attributable to $b(\omega_0)$ tends to 0 as $\varepsilon \downarrow 0$, although at only a logarithmic rate. Not coincidentally, direct calculation shows that for $F_\varepsilon = \{1\}$, (1) holds and I would expect it to hold for more general F_ε. In fact, Theorem 2 in Section 3 applies in this case and it follows that (1) holds when F_ε has two points (and $y_2 = 0$).

Next consider some settings for the Matérn model with $\nu = \frac{3}{2}$ on $\mathbb{R}^2$. Figure 4(a) shows a situation in which there are two nearby observations in the vertical direction from the origin and two distant observations in the horizontal direction. One

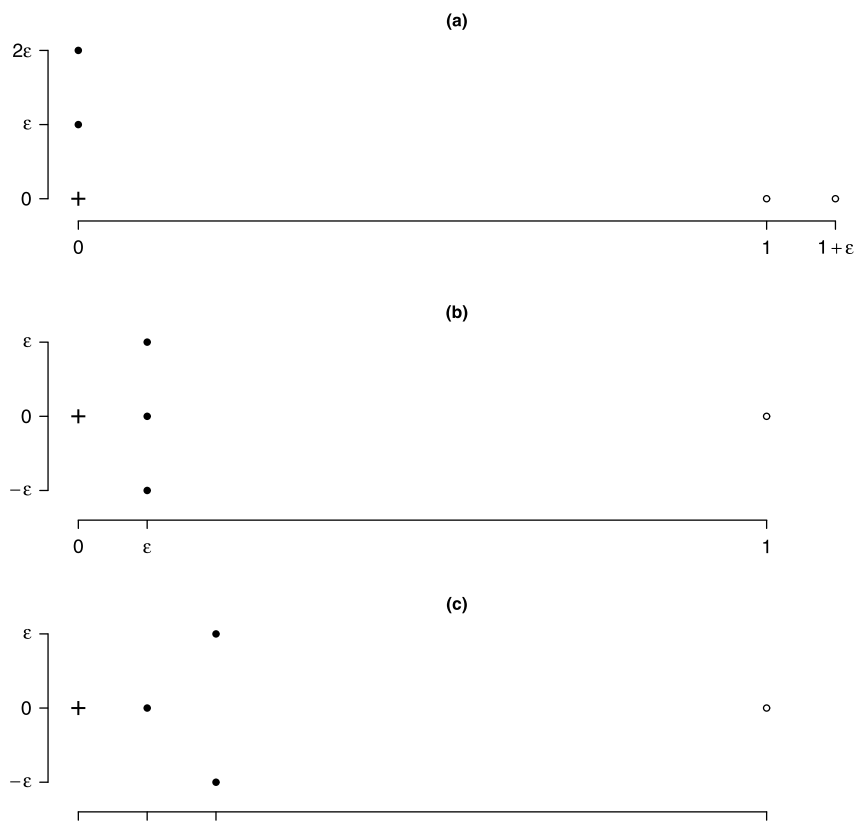

FIG. 4. *Prediction problems for isotropic Matérn model on $\mathbb{R}^2$ with $\nu = \frac{3}{2}$. Symbols as in Figure 1.*

might imagine that because the nearby observations provide no information about how the process varies in the horizontal direction, the distant observations might provide nonneglible new information about $Z(0)$. However, Section 5.1 demonstrates that (1) does hold in this case. The next two examples are related to the one-dimensional examples considered in Figure 3 for a Matérn model with $\nu = \frac{3}{2}$. Write $Z_{i,j}$ for the ijth partial derivative of Z. In Figure 4(b), N_ε has three observations, but they are collinear along a line that does not go through the origin and it is possible to show that the BLP of $Z_{1,0}(0,0)$ based on $Z(N_\varepsilon)$ has asymptotically negligible correlation with $Z_{1,0}(0,0)$ as $\varepsilon \downarrow 0$. As a consequence, the asymptotic results are identical to what we had in Figure 3(a): $Ee(N_\varepsilon)^2 \sim \varepsilon^2$ and $Ee(N_\varepsilon \cup F_\varepsilon)^2 \sim \frac{e^2-5}{e^2-4}\varepsilon^2$ as $\varepsilon \downarrow 0$. If N_ε has three points arranged as in Figure 4(c), then $\{Z(\varepsilon,0) - \frac{1}{2}Z(2\varepsilon,\varepsilon) - \frac{1}{2}Z(2\varepsilon,-\varepsilon)\}/\varepsilon$ is a consistent predictor of $Z_{1,0}(0,0)$ and (1) holds; see Section 5.1.

Now consider a model satisfying (2) for which the process is not equally differentiable in all directions. Stein (2005) gives an example of such a model. Specifically, consider a space–time model on $\mathbb{R}^3 \times \mathbb{R}$ with spectral density $\{(1+|\omega_1|^2)^2 + \omega_2^2\}^{-2}$, $(\omega_1,\omega_2) \in \mathbb{R}^3 \times \mathbb{R}$. Writing erfc for the complementary error function, the corresponding autocovariance function K is Stein (2005)

$$
\begin{aligned}
K(x,t) = {} & \frac{1}{16}\pi^2 e^{|x|}\operatorname{erfc}\biggl(|t|^{1/2} + \frac{|x|}{2|t|^{1/2}}\biggr)\biggl(1 - |x| + \frac{4t^2}{|x|}\biggr) \\
& + \frac{1}{16}\pi^2 e^{-|x|}\operatorname{erfc}\biggl(|t|^{1/2} - \frac{|x|}{2|t|^{1/2}}\biggr)\biggl(1 + |x| - \frac{4t^2}{|x|}\biggr) \\
& + \frac{1}{4}\pi^{3/2}|t|^{1/2}\exp\biggl(-|t| - \frac{|x|^2}{4|t|}\biggr)
\end{aligned}
\tag{4}
$$

for $x \neq 0$ and $t \neq 0$. For $x = 0$ or $t = 0$, we can define K by continuity. For $t = 0$, we get $K(x,0) = \frac{1}{8}\pi^2 e^{-|x|}(1+|x|)$, the Matérn model with $\nu = \frac{3}{2}$, so the corresponding process is exactly once mean square differentiable in any spatial direction. Stein (2005) shows that $K(0,t) = \frac{1}{8}\pi^2 - \frac{2}{3}\pi^{3/2}|t|^{3/2} + O(t^2)$ as $t \to 0$ so that $K(0,t)$ is not twice differentiable in t at $t = 0$, and the corresponding process is not mean square differentiable in time.

For (4), let us again consider the setting in Figure 4(c) with the horizontal axis corresponding to the first spatial coordinate and the vertical axis corresponding to time. It now turns out that the two points in N_ε off of the horizontal axis contribute negligibly to the BLP whether or not F_ε is included. The problem is that the lack of differentiability of Z in the vertical direction implies that the BLP of $Z_{1,0}(0,0)$ based on $Z(N_\varepsilon)$ has asymptotic correlation 0 with $Z_{1,0}(0,0)$. Consequently, the asymptotic results are the same as in Figure 3(a) for $K(x) = e^{-|x|}(1+|x|)$; that is, $Ee(N_\varepsilon \cup F_\varepsilon)^2/Ee(N_\varepsilon)^2 \to \frac{e^2-5}{e^2-4}$ as $\varepsilon \downarrow 0$ (Section 5.1).

Figure 5 displays two other settings we now consider for K as in (4). In Figure 5(a), we have $Ee(N_\varepsilon \cup F_\varepsilon)^2/Ee(N_\varepsilon)^2 \to 1$ as $\varepsilon \downarrow 0$ and, in Figure 5(b),

FIG. 5. *Prediction problems for autocovariance function given in* (4). *Horizontal axes are first spatial coordinate and vertical axes are time. Symbols as in Figure* 1.

$Ee(N_\varepsilon \cup F_\varepsilon)^2/Ee(N_\varepsilon)^2 \to \frac{e^2-5}{e^2-4}$ as $\varepsilon \downarrow 0$. These two cases show that it is possible to have sets $N_\varepsilon \subset \tilde{N}_\varepsilon$ yet have that (1) holds for the pair of sets $(N_\varepsilon, F_\varepsilon)$ but not $(\tilde{N}_\varepsilon, F_\varepsilon)$, further complicating any search for a general result that applies to processes that are not equally smooth in all directions.

These examples demonstrate that any general theorem that encompasses all of them will need a condition on N_ε that depends on f. The following conjecture is in accord with all of the examples presented here:

CONJECTURE 1. Suppose f satisfies (2) and the following assumption:

ASSUMPTION A. for $j = 1, \dots, n$, all mean square derivatives of Z at the origin in the direction x_j can be predicted based on $Z(N_\varepsilon)$ with mean squared error tending to 0 as $\varepsilon \downarrow 0$.

Then for all $r > 0$,

$$\lim_{\varepsilon \downarrow 0} \frac{Ee\{N_\varepsilon \cup b(r)^c\}^2}{Ee\{N_\varepsilon\}^2} = 1.$$

Note that here I have expanded the set of distant observations to include all locations more than r from the origin, which simplifies the statement of the result although undoubtedly complicates its proof (assuming it is true). It is somewhat unsatisfying to have the condition on N_ε given in terms of properties of predictors of derivatives of Z rather than some purely geometric condition, but I see no way to accommodate the examples treated here for K as in (4) without a condition something like Assumption A. Verifying whether Assumption A holds in any particular setting may require a fair amount of work, although for N_ε of fixed and finite cardinality as we consider here, it should generally be possible to make this determination. Note that if all mean square derivatives of Z at the origin can be consistently predicted based on $Z(N_\varepsilon)$ as $\varepsilon \downarrow 0$, then Assumption A holds for any $\tilde{N}_\varepsilon = \{\varepsilon s_1, \ldots, \varepsilon s_\ell\}$ with $\{x_1, \ldots, x_n\} \subset \{s_1, \ldots, s_\ell\}$.

In all of the examples for which (1) holds,

$$\lim_{\varepsilon \downarrow 0} \frac{\int_{b(\omega_0)} |\eta_\varepsilon(\omega)|^2 f(\omega)\, d\omega}{\|\eta_\varepsilon\|_f^2} = 0 \tag{5}$$

for all $\omega_0 > 0$, and I suspect that Assumption A is equivalent to (5). Examining the proof of Theorem 1 in Section 5.2 [see (19)], one sees that (5) is essential to making the proof work.

3. Theorems. I do not know how to prove Conjecture 1 in anything like its full generality. Assuming Z is not differentiable in any direction simplifies matters considerably, because Assumption A then holds for any nonempty N_ε. Theorem 1 considers nondifferentiable processes on $\mathbb{R}$ and Theorem 2 nondifferentiable processes on $\mathbb{R}^2$.

THEOREM 1. *Suppose, for $d = 1$ and some $\alpha \in (0, 2)$,*

$$f(\omega) \asymp (1 + |\omega|)^{-\alpha-1}, \tag{6}$$

and f satisfies (2). Then (1) holds.

Condition (6) is stronger than necessary to guarantee Z is not differentiable. Because part of the proof is to show that the low frequencies do not matter in the limit, (6) can likely be weakened to hold only for all ω sufficiently large. Removing (6) entirely would be more difficult.

The next theorem applies to nondifferentiable processes in $\mathbb{R}^2$ and does not require any conditions on f beyond (2). However, it does restrict N_ε to have only one point and F_ε to have two. The theorem also assumes $y_2 = 0$, but this restriction does not meaningfully detract from the content of the result and, in any case, could be removed at the cost of a somewhat messier proof. Extending the result to $\mathbb{R}^d$ is straightforward, but taking $d > 3$ is pointless in this setting because any 4 points in $\mathbb{R}^d$ fall on a three-dimensional hyperplane, and even taking $d = 3$ provides no new insight beyond what is learned from the two-dimensional setting.

THEOREM 2. *Suppose Z has spectral density f satisfying (2) and that Z is not mean square differentiable in any direction. In addition, suppose $N_\varepsilon = \{\varepsilon x_1\}$ and $F_\varepsilon = \{y_0, y_0 + \varepsilon y_1\}$, where x_1, y_0 and y_1 are all nonzero. Then (1) holds.*

Note that the example referred to in Figure 2 satisfies the conditions on N_ε and F_ε in Theorem 2, and the process is not mean square differentiable in any direction, but f does not satisfy (2). As we have seen, (1) does not hold in this setting, so that Theorem 2 would be false if we removed (2).

Throughout this work we assume that Z has a known mean 0. It is common in practice to assume that Z has an unknown constant mean μ and then predict $Z(0)$ by what is called the ordinary kriging predictor, which is just an example of the best linear unbiased predictor [Stein (1999a)]. In all of the examples considered in Section 2, for which (1) holds for simple kriging, it still holds for ordinary kriging. Furthermore, Theorems 1 and 2 can be easily shown to hold for ordinary kriging by proving that, under the conditions of the theorems, the ordinary kriging predictor based on N_ε is asymptotically optimal relative to the simple kriging predictor (see the ends of each proof in Section 5). Thus, if Conjecture 1 holds for simple kriging, then I would expect it also holds for ordinary kriging.

4. Discussion. The space–time process on $\mathbb{R}^3 \times \mathbb{R}$ considered in Section 3 with spectral density $\{(1+|\omega_1|^2)^2+\omega_2^2\}^{-2}$, $(\omega_1,\omega_2)\in\mathbb{R}^3\times\mathbb{R}$, is an example of a process with a different degree of differentiability in time than in space. It is a special case of the stochastic fractional heat equations studied by Kelbert, Leonenko and Ruiz-Medina (2005), which are in turn a special case of a class of space–time processes suggested in Stein (2005) whose spectral densities are of the form

$$f(\omega_1,\omega_2)=\{c_1(a_1^2+|\omega_1|^2)^{\alpha_1}+c_2(a_2^2+|\omega_2|^2)^{\alpha_2}\}^{-\nu} \tag{7}$$

for $\omega_1\in\mathbb{R}^{d_1}$, $\omega_2\in\mathbb{R}^{d_2}$, $\nu>\frac{d_1}{2\alpha_1}+\frac{d_2}{2\alpha_2}$ and $c_1,c_2,\alpha_1,\alpha_2$ and $a_1^2+a_2^2$ positive to ensure f is integrable. Because of the superficial similarity of this model to the Matérn model, we might call it doubly Matérn. All spectral densities of the form (7) satisfy (2) and thus, I conjecture, satisfy an asymptotic screening effect whenever Assumption A applies to N_ε. At the same time, by adjusting the parameters α_1,α_2 and ν, we can obtain processes with any desired degree of differentiability in time and any separate degree of differentiability in space [Stein (2005)]. Note that f of the form (7) satisfies the conditions of Theorem 2 when $d_1=d_2=1$, $2\nu\le\frac{3}{\alpha_1}+\frac{1}{\alpha_2}$ and $2\nu\le\frac{1}{\alpha_1}+\frac{3}{\alpha_2}$, the last two conditions being necessary and sufficient to make Z not mean square differentiable in any direction. Stein (2011) derives some results for the covariance structure when $a_1=a_2=0$ and $\alpha_2=1$.

Despite its flexibility, model (7) is still restrictive in some ways, in particular in exhibiting what Gneiting (2002) calls full symmetry, due to the fact that $f(\omega_1,\omega_2)=f(\omega_1,-\omega_2)$, and hence the corresponding process has the same covariance structure with time running backwards as it does with time running forward. Thus, for example, this model is unsuitable for processes with a dominant direction of advection. Stein (2005) discusses possible approaches to extending this model to allow for asymmetries.

As noted in Section 3, (5), which says that only an asymptotically negligible fraction of $Ee(N_\varepsilon)^2$ can be attributed to some fixed frequency range, is crucial to obtaining a screening effect. This same property was also the key idea in Stein (1999b) to obtaining explicit results on the asymptotic efficiency of predictors based on an incorrect spectral density having similar behavior to the correct spectral density at high frequencies. The high-frequency behavior of a Gaussian process is also crucial to estimation of the covariance structure [Stein (1999a)], and misspecification of this high-frequency behavior can lead to poor behavior of estimates, particularly if likelihood-based methods are used [Stein (1999a), Chapter 6, and Stein (2008)]. As statisticians strive to advance the statistical analysis of spatial-temporal processes, they should pay close attention to the spectral behavior of the models they use. In particular, models that do not satisfy (2) should be used with caution.

5. Proofs.

5.1. *Examples*. For a random vector Y, write $\operatorname{cov}(Y)$ for the covariance matrix of Y, write 0 for a column vector of zeroes whose length is apparent from context and denote transposes by primes. The following result simplifies the calculations for several of the examples.

LEMMA 1. *If there exists $a(\varepsilon) > 0$, $\delta_\varepsilon \in \mathbb{R}^{n+m}$ and Δ_ε an $(n+m)\times(n+m)$ matrix such that*

$$\lim_{\varepsilon\downarrow 0}\operatorname{cov}\begin{pmatrix} a(\varepsilon)\{Z(0)-\delta_\varepsilon\cdot Z(N_\varepsilon\cup F_\varepsilon)\} \\ \Delta_\varepsilon Z(N_\varepsilon\cup F_\varepsilon)\end{pmatrix} = \begin{pmatrix} k & 0' \\ 0 & K\end{pmatrix} \tag{8}$$

for some $k > 0$ and K positive definite, then

$$\lim_{\varepsilon\downarrow 0}\frac{E\{Z(0)-\delta_\varepsilon\cdot Z(\varepsilon)\}^2}{Ee(N_\varepsilon\cup F_\varepsilon)^2} = 1.$$

For (8) to hold, $\tilde e_\varepsilon = Z(0)-\delta_\varepsilon\cdot Z(N_\varepsilon\cup F_\varepsilon)$ must satisfy $Ee(N_\varepsilon\cup F_\varepsilon)^2/E\tilde e_\varepsilon^2 \to 1$ as $\varepsilon\downarrow 0$. To prove the lemma, note that (8) and K positive definite imply $\operatorname{cov}\{\Delta_\varepsilon Z(\varepsilon)\}$ is positive definite for all ε sufficiently small. Thus, for all ε sufficiently small, the BLP of $Z(0)$ based on $Z(N_\varepsilon\cup F_\varepsilon)$ is the same as the BLP of $Z(0)$ based on $\Delta_\varepsilon Z(N_\varepsilon\cup F_\varepsilon)$. Since matrix inverse is a continuous function in some neighborhood of K, using basic results on BLPs [e.g., Stein (1999a), Section 1.2],

$$\begin{aligned}
&a(\varepsilon)^2 Ee(N_\varepsilon\cup F_\varepsilon)^2\\
&\quad = a(\varepsilon)^2\bigl(\operatorname{var}\tilde e_\varepsilon - \operatorname{cov}\{\tilde e_\varepsilon, Z(N_\varepsilon\cup F_\varepsilon)'\Delta_\varepsilon'\}\\
&\qquad\qquad \times[\operatorname{cov}\{\Delta_\varepsilon Z(N_\varepsilon\cup F_\varepsilon)\}]^{-1}\operatorname{cov}\{\Delta_\varepsilon Z(N_\varepsilon\cup F_\varepsilon),\tilde e_\varepsilon\}\bigr)\\
&\quad = \operatorname{var}\{a(\varepsilon)\tilde e_\varepsilon\} - \operatorname{cov}\{a(\varepsilon)\tilde e_\varepsilon, Z(N_\varepsilon\cup F_\varepsilon)'\Delta_\varepsilon'\}\\
&\qquad \times[\operatorname{cov}\{\Delta_\varepsilon Z(N_\varepsilon\cup F_\varepsilon)\}]^{-1}\operatorname{cov}\{\Delta_\varepsilon Z(N_\varepsilon\cup F_\varepsilon), a(\varepsilon)\tilde e_\varepsilon\}\\
&\quad \to k - 0'K^{-1}0
\end{aligned}$$

as $\varepsilon \downarrow 0$, and the lemma follows.

To apply Lemma 1 to the setting in Figure 4(a) with $K(x) = e^{-|x|}(1 + |x|)$, it suffices to show

$$\lim_{\varepsilon \downarrow 0} \operatorname{cov} \begin{pmatrix} \varepsilon^{-3/2}\{Z(0,0) - 2Z(0,\varepsilon) + Z(0,2\varepsilon)\} \\ Z(0,\varepsilon) \\ \varepsilon^{-1}\{Z(0,2\varepsilon) - Z(0,\varepsilon)\} \\ Z(1,0) \\ \varepsilon^{-1}\{Z(1+\varepsilon,0) - Z(1,0)\} \end{pmatrix} = \begin{pmatrix} \frac{8}{3} & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2e^{-1} & -e^{-1} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 2e^{-1} & 0 & 1 & 0 \\ 0 & -e^{-1} & 0 & 0 & 1 \end{pmatrix}.$$

To show, for example, that $\operatorname{cov}[\varepsilon^{-1}\{Z(0,2\varepsilon) - Z(0,\varepsilon)\}, \varepsilon^{-1}\{Z(1+\varepsilon,0) - Z(1,0)\}] \to 0$, define the function $\tilde{K}$ on $[0,\infty)$ by $\tilde{K}(r) = e^{-r}(1+r)$, which has bounded derivatives of all orders on $[0,\infty)$. Then using a Taylor series,

$$\begin{aligned}
&\operatorname{cov}\{Z(0,2\varepsilon) - Z(0,\varepsilon), Z(1+\varepsilon,0) - Z(1,0)\} \\
&\quad = K\left(\sqrt{(1+\varepsilon)^2 + 4\varepsilon^2}\right) - K\left(\sqrt{(1+\varepsilon)^2 + \varepsilon^2}\right) \\
&\qquad - K\left(\sqrt{1+4\varepsilon^2}\right) + K\left(\sqrt{1+\varepsilon^2}\right) \\
&\quad = K'(1+\varepsilon)\left\{\sqrt{(1+\varepsilon)^2 + 4\varepsilon^2} - \sqrt{(1+\varepsilon)^2 + \varepsilon^2}\right\} \\
&\qquad - K'(1)\left\{\sqrt{1+4\varepsilon^2} - \sqrt{1+\varepsilon^2}\right\} + O(\varepsilon^4) \\
&\quad = K'(1+\varepsilon)(1+\varepsilon)\left\{\frac{2\varepsilon^2}{(1+\varepsilon)^2} - \frac{\varepsilon^2}{2(1+\varepsilon)^2}\right\} - K'(1)\left(2\varepsilon^2 - \frac{1}{2}\varepsilon^2\right) + O(\varepsilon^4) \\
&\quad \ll \varepsilon^3,
\end{aligned}$$

and $\operatorname{cov}[\varepsilon^{-1}\{Z(0,2\varepsilon) - Z(0,\varepsilon)\}, \varepsilon^{-1}\{Z(1+\varepsilon,0) - Z(1,0)\}] \to 0$ follows.

Lemma 1 can be applied to the setting in Figure 4(c) with $K(x) = e^{-|x|}(1+|x|)$ by showing

$$\lim_{\varepsilon \downarrow 0} \operatorname{cov} \begin{pmatrix} \varepsilon^{-3/2}\{Z(0,0) - 2Z(\varepsilon,0) + \frac{1}{2}Z(2\varepsilon,\varepsilon) + \frac{1}{2}Z(2\varepsilon,-\varepsilon)\} \\ Z(\varepsilon,0) \\ \varepsilon^{-1}\{2Z(\varepsilon,0) - Z(2\varepsilon,\varepsilon) - Z(2\varepsilon,-\varepsilon)\} \\ Z(1,0) \end{pmatrix} = \begin{pmatrix} \frac{1}{3}(10\sqrt{5} - 8\sqrt{2}) & 0 & 0 & 0 \\ 0 & 1 & 0 & 2e^{-1} \\ 0 & 0 & 4 & -2e^{-1} \\ 0 & 2e^{-1} & -2e^{-1} & 1 \end{pmatrix}.$$

Specifically, it is not necessary to consider $Z(2\varepsilon, \varepsilon)$ and $Z(2\varepsilon, -\varepsilon)$ separately: by symmetry, the BLP of $Z(0,0)$ based on $Z(N_\varepsilon \cup F_\varepsilon)$ depends on $Z(2\varepsilon, \varepsilon)$ and $Z(2\varepsilon, -\varepsilon)$ only through $Z(2\varepsilon, \varepsilon) + Z(2\varepsilon, -\varepsilon)$.

As a final example, let us apply Lemma 1 to the setting in Figure 4(c) with K given by (4). Again by symmetry, we can restrict to predictors that depend on $Z(2\varepsilon, \varepsilon)$ and $Z(2\varepsilon, -\varepsilon)$ only through $Z(2\varepsilon, \varepsilon) + Z(2\varepsilon, -\varepsilon)$. For a and b fixed and positive, using a Taylor series and

$$\operatorname{erfc}(x) = 1 - \frac{2}{\sqrt{\pi}}\left(x - \frac{1}{3}x^3\right) + O(|x|^5)$$

as $x \to 0$ [Olver et al. (2010), page 162], it is possible to show

$$K(a\varepsilon, b\varepsilon) = \tfrac{1}{8}\pi^2 - \tfrac{2}{3}(\pi b\varepsilon)^{3/2} + O(\varepsilon^2)$$

as $\varepsilon \downarrow 0$. This result also holds when a or b equals 0. It follows that

$$\begin{aligned}
&\lim_{\varepsilon\downarrow 0} \operatorname{cov}\begin{pmatrix} \varepsilon^{-1}\{Z(0,0) - Z(\varepsilon,0)\} \\ Z(\varepsilon,0) \\ \varepsilon^{-3/4}\{2Z(\varepsilon,0) - Z(2\varepsilon,\varepsilon) - Z(2\varepsilon,-\varepsilon)\} \end{pmatrix} \\
&\quad = \begin{pmatrix} \frac{1}{8}\pi^2 & 0 & 0 \\ 0 & \frac{1}{8}\pi^2 & 0 \\ 0 & 0 & \frac{8}{3}(2-\sqrt{2})\pi^{3/2} \end{pmatrix}
\end{aligned}$$

so that $Z(\varepsilon, 0)$ is an asymptotically optimal predictor of $Z(0,0)$ based on N_ε. Furthermore, for $c_1 = 2/(e^2 - 4)$ and $c_2 = -e/(e^2 - 4)$,

$$\begin{aligned}
&\lim_{\varepsilon\downarrow 0} \operatorname{cov}\begin{pmatrix} \varepsilon^{-1}\{Z(0,0) - (1 + c_1\varepsilon)Z(\varepsilon,0) - c_2\varepsilon Z(1,0)\} \\ Z(\varepsilon,0) \\ Z(1,0) \\ \varepsilon^{-3/4}\{2Z(\varepsilon,0) - Z(2\varepsilon,\varepsilon) - Z(2\varepsilon,-\varepsilon)\} \end{pmatrix} \\
&\quad = \begin{pmatrix} \dfrac{1}{8}\pi^2\dfrac{e^2-5}{e^2-4} & 0 & 0 & 0 \\ 0 & \dfrac{1}{8}\pi^2 & \dfrac{\pi^2}{4e} & 0 \\ 0 & \dfrac{\pi^2}{4e} & \dfrac{1}{8}\pi^2 & 0 \\ 0 & 0 & 0 & \dfrac{8}{3}(2-\sqrt{2})\pi^{3/2} \end{pmatrix},
\end{aligned}$$

and the conditions of Lemma 1 are satisfied.

5.2. *Proof of Theorem* 1. Theorem 3.1 in Xue and Xiao (2011) implies

$$\|\eta_\varepsilon\|_f^2 = Ee(N_\varepsilon)^2 \asymp \varepsilon^\alpha \tag{9}$$

as $\varepsilon \downarrow 0$. Let us use (9) to show that $\sum_{j=1}^n |\phi_{j\varepsilon}|$ is bounded in ε as $\varepsilon \downarrow 0$. If we define $M_\varepsilon = \max(1, \sum_{j=1}^n |\phi_{j\varepsilon}|)$ and $x_0 = 0$, we can write $\eta_\varepsilon(\omega)$ in the form $M_\varepsilon \sum_{j=0}^n \mu_{j\varepsilon} e^{i\varepsilon\omega x_j}$ for appropriate $\mu_{j\varepsilon}$'s, where, by construction, $|\mu_{j\varepsilon}| \le 1$ for all j and ε. Thus, if we can show M_ε bounded, then $\sum_{j=1}^n |\phi_{j\varepsilon}|$ is also bounded. By (6), there exists $0 < C_1 \le C_2 < \infty$ such that

$$\frac{C_1}{(1+|\omega|)^{\alpha+1}} \le f(\omega) \le \frac{C_2}{(1+|\omega|)^{\alpha+1}} \tag{10}$$

for all ω. Thus, making the change of variables $\nu = \varepsilon\omega$ in the second step,

$$\begin{aligned} Ee(N_\varepsilon)^2 &\ge C_1 M_\varepsilon^2 \int_{-\infty}^{\infty} \left| \sum_{j=0}^n \mu_{j\varepsilon} e^{i\varepsilon\omega x_j} \right|^2 (1+|\omega|)^{-\alpha-1}\, d\omega \\ &= C_1 M_\varepsilon^2 \varepsilon^\alpha \int_{-\infty}^{\infty} \left| \sum_{j=0}^n \mu_{j\varepsilon} e^{i\nu x_j} \right|^2 (\varepsilon+|\nu|)^{-\alpha-1}\, d\nu \qquad (11) \\ &\ge C_1 M_\varepsilon^2 \left(\frac{1}{2}\varepsilon\right)^\alpha \int_1^{\infty} \left| \sum_{j=0}^n \mu_{j\varepsilon} e^{i\nu x_j} \right|^2 \nu^{-\alpha-1}\, d\nu \end{aligned}$$

for all $\varepsilon < 1$. Suppose M_ε is unbounded. Then there exists a sequence $\{\varepsilon(k)\}$ tending to 0 such that $M_{\varepsilon(k)} \to \infty$. Because the $\mu_{j\varepsilon}$'s are bounded, there exists $(\mu_0, \dots, \mu_n) \in \mathbb{R}^{n+1}$ and a subsequence of $\{\varepsilon(k)\}$, call it $\{\varepsilon(k_\ell)\}$, along which $(\mu_{0\varepsilon(k_\ell)}, \dots, \mu_{n\varepsilon(k_\ell)}) \to (\mu_0, \dots, \mu_n)$ as $\ell \to \infty$. Since $\alpha > 0$, by dominated convergence, it follows that

$$\int_1^{\infty} \left| \sum_{j=0}^n \mu_{j\varepsilon(k_\ell)} e^{i\nu x_j} \right|^2 \nu^{-\alpha-1}\, d\nu \to \int_1^{\infty} \left| \sum_{j=0}^n \mu_j e^{i\nu x_j} \right|^2 \nu^{-\alpha-1}\, d\nu > 0$$

as $\ell \to \infty$, which, together with (11), contradicts (9), so M_ε and $\sum_{j=1}^n |\phi_{j\varepsilon}|$ must be bounded as $\varepsilon \downarrow 0$.

Now consider the behavior of η_ε at low frequencies. Define $p_\varepsilon = 1/\sum_j \phi_{j\varepsilon}$ and $\tilde{\eta}_\varepsilon(\omega) = 1 - p_\varepsilon \sum_{j=1}^n \phi_{j\varepsilon} e^{i\varepsilon\omega x_j}$. By (6) and (9), $\int_0^1 |\eta_\varepsilon(\omega)|^2\, d\omega \ll \varepsilon^\alpha$ and, writing Re for real part, $|\eta_\varepsilon(\omega)|^2 \ge \{\operatorname{Re} \eta_\varepsilon(\omega)\}^2 = (1-p_\varepsilon)^2 + O(\varepsilon^2)$ uniformly for $\omega \in [0,1]$. It follows that

$$(p_\varepsilon - 1)^2 \ll \varepsilon^\alpha \tag{12}$$

as $\varepsilon \downarrow 0$. Using $|e^{ix} - 1| \le |x|$ for all $x \in \mathbb{R}$, for $\beta \in [0,1]$ and $\alpha \in (0,2)$,

$$\int_{b(\varepsilon^{-\beta})} |\tilde{\eta}_\varepsilon(\omega)|^2 f(\omega)\, d\omega \ll \varepsilon^2 \int_0^{\varepsilon^{-\beta}} \frac{\omega^2}{1+\omega^{\alpha+1}}\, d\omega \ll \varepsilon^{2-\beta(2-\alpha)} \tag{13}$$

as $\varepsilon \downarrow 0$. Because $\sum_{j=1}^{n} \phi_{j\varepsilon} Z(\varepsilon x_j)$ is the BLP of $Z(0)$, $\|\eta_\varepsilon\|_f^2 \le \|\tilde{\eta}_\varepsilon\|_f^2$, so that

$$\begin{aligned}\int_{b(\varepsilon^{-\beta})} |\eta_\varepsilon(\omega)|^2 f(\omega)\, d\omega \le & \int_{b(\varepsilon^{-\beta})} |\tilde{\eta}_\varepsilon(\omega)|^2 f(\omega)\, d\omega \\ & + \int_{b(\varepsilon^{-\beta})^c} \{|\tilde{\eta}_\varepsilon(\omega)|^2 - |\eta_\varepsilon(\omega)|^2\} f(\omega)\, d\omega.\end{aligned} \tag{14}$$

Straightforward algebra shows

$$\begin{aligned}|\tilde{\eta}_\varepsilon(\omega)|^2 & - |\eta_\varepsilon(\omega)|^2 \\ & = (p_\varepsilon^2 - 1)|\phi_\varepsilon(\omega)|^2 - 2(p_\varepsilon - 1)\operatorname{Re}\phi_\varepsilon(\omega) \\ & = 2(p_\varepsilon - 1)^2 |\phi_\varepsilon(\omega)|^2 + 2(p_\varepsilon - 1)[|\phi_\varepsilon(\omega)|^2 - \operatorname{Re}\phi_\varepsilon(\omega)].\end{aligned} \tag{15}$$

The boundedness of the $\phi_{j\varepsilon}$'s in ε implies $||\phi_\varepsilon(\omega)|^2 - p_\varepsilon^{-2}|^2 \ll \min(1, \varepsilon^2\omega^2)$ and $|\operatorname{Re}\phi_\varepsilon(\omega) - p_\varepsilon^{-1}| \ll \min(1, \varepsilon^2\omega^2)$, and it follows that

$$\big||\phi_\varepsilon(\omega)|^2 - \operatorname{Re}\phi_\varepsilon(\omega)\big| \ll |p_\varepsilon - 1| + \min(1, \varepsilon^2\omega^2)$$

as $\varepsilon \downarrow 0$, which, together with (12) and (15), yields

$$|\tilde{\eta}_\varepsilon(\omega)|^2 - |\eta_\varepsilon(\omega)|^2 \ll \varepsilon^\alpha + \varepsilon^{\alpha/2} \min(1, \varepsilon^2\omega^2)$$

as $\varepsilon \downarrow 0$. Thus,

$$\begin{aligned}\int_{b(\varepsilon^{-\beta})^c} & \{|\tilde{\eta}_\varepsilon(\omega)|^2 - |\eta_\varepsilon(\omega)|^2\} f(\omega)\, d\omega \\ & \ll \int_{\varepsilon^{-\beta}}^{\varepsilon^{-1}} \frac{\varepsilon^\alpha + \varepsilon^{2+\alpha/2}\omega^2}{\omega^{\alpha+1}}\, d\omega + \int_{\varepsilon^{-1}}^{\infty} \frac{\varepsilon^{\alpha/2}}{\omega^{\alpha+1}}\, d\omega \\ & \ll \varepsilon^{3\alpha/2} + \varepsilon^{\alpha(\beta+1)}\end{aligned} \tag{16}$$

as $\varepsilon \downarrow 0$. Combining this bound with (13) and (14) implies that for all $\beta \in [0, 1]$,

$$\int_{b(\varepsilon^{-\beta})} |\eta_\varepsilon(\omega)|^2 f(\omega)\, d\omega \ll \varepsilon^{2-\beta(2-\alpha)} + \varepsilon^{3\alpha/2} + \varepsilon^{\alpha(\beta+1)} \tag{17}$$

as $\varepsilon \downarrow 0$. Note that the bound in (17) is $o(\varepsilon^\alpha)$ as $\varepsilon \downarrow 0$ for all $\alpha \in (0, 2)$ and $\beta \in (0, 1)$.

Let $\Lambda_\varepsilon = (\lambda_{1\varepsilon}, \dots, \lambda_{m\varepsilon})$, and assume $\Lambda_\varepsilon \neq 0$ hereafter, as the case $\Lambda_\varepsilon = 0$ is trivial to handle. We next show the correlation of $e(N_\varepsilon)$ and $\Lambda_\varepsilon \cdot Z(F_\varepsilon)$ is asymptotically negligible. Defining $\lambda_\varepsilon(\omega) = \sum_{j=1}^{m} \lambda_{j\varepsilon} e^{i\varepsilon\omega y_j}$,

$$\begin{aligned}\operatorname{corr}\{ & e(N_\varepsilon), \Lambda_\varepsilon \cdot Z(F_\varepsilon)\} \\ & = \frac{\int_{b(\varepsilon^{-\beta})} \eta_\varepsilon(\omega) e^{-i\omega y_0} \overline{\lambda_\varepsilon(\omega)} f(\omega)\, d\omega}{\|\eta_\varepsilon\|_f \|\lambda_\varepsilon\|_f} \\ & \quad + \frac{\int_{b(\varepsilon^{-\beta})^c} \eta_\varepsilon(\omega) e^{-i\omega y_0} \overline{\lambda_\varepsilon(\omega)} f(\omega)\, d\omega}{\|\eta_\varepsilon\|_f \|\lambda_\varepsilon\|_f} \\ & \triangleq I_1 + I_2.\end{aligned} \tag{18}$$

Using the Cauchy–Schwarz inequality and (17), for $\beta \in (0,1)$,

$$I_1 \le \frac{\sqrt{\int_{b(\varepsilon^{-\beta})} |\eta_\varepsilon(\omega)|^2 f(\omega)\, d\omega}}{\|\eta_\varepsilon\|_f} \to 0 \tag{19}$$

as $\varepsilon \downarrow 0$, uniformly in Λ_ε. Next, define $R_k = 2\pi k/y_0$ and $k_\varepsilon = \lfloor y_0 \varepsilon^{-\beta}/(2\pi) \rfloor$. Then $R_{k_\varepsilon} \le \varepsilon^{-\beta}$ and

$$\begin{aligned}
&\left| \int_{b(\varepsilon^{-\beta})^c} \eta_\varepsilon(\omega) e^{-i\omega y_0} \overline{\lambda_\varepsilon(\omega)} f(\omega)\, d\omega \right| \\
&\quad \le 2 \sum_{k=k_\varepsilon}^{\infty} \left| \int_{R_k}^{R_{k+1}} \eta_\varepsilon(\omega) e^{-i\omega y_0} \overline{\lambda_\varepsilon(\omega)} f(\omega)\, d\omega \right| \\
&\quad \le 2 \sum_{k=k_\varepsilon}^{\infty} f(R_k) \left| \int_{R_k}^{R_{k+1}} \eta_\varepsilon(\omega) e^{-i\omega y_0} \overline{\lambda_\varepsilon(\omega)}\, d\omega \right| \\
&\qquad + 2 \sum_{k=k_\varepsilon}^{\infty} \int_{R_k}^{R_{k+1}} |\eta_\varepsilon(\omega)\lambda_\varepsilon(\omega)| |f(\omega) - f(R_k)|\, d\omega \\
&\quad \triangleq I_3 + I_4.
\end{aligned} \tag{20}$$

For $\omega \in (R_k, R_{k+1}]$, by (2), there exist constants $c_k \to 0$ as $k \to \infty$ such that

$$|f(\omega) - f(R_k)| \le c_k \min\{f(R_k), f(\omega)\}, \tag{21}$$

so

$$\begin{aligned}
I_4 &\le 2 \sum_{k=k_\varepsilon}^{\infty} c_k \int_{R_k}^{R_{k+1}} |\eta_\varepsilon(\omega)\lambda_\varepsilon(\omega)| f(\omega)\, d\omega \\
&\le 2 \sup_{k \ge k_\varepsilon} c_k \int_{R_{k_\varepsilon}}^{\infty} |\eta_\varepsilon(\omega)\lambda_\varepsilon(\omega)| f(\omega)\, d\omega \\
&\le \sup_{k \ge k_\varepsilon} c_k \|\eta_\varepsilon\|_f \|\lambda_\varepsilon\|_f,
\end{aligned} \tag{22}$$

the last step by the Cauchy–Schwarz inequality. Now consider I_3 in (20). Defining $\theta_\varepsilon(\omega) = \eta_\varepsilon(\omega)\overline{\lambda_\varepsilon(\omega)}$, we can write θ_ε in the form $\sum_{j=1}^{m(n+1)} \theta_{j\varepsilon} e^{i\varepsilon\omega z_j}$. For $M_\varepsilon = \max(1, \sum_{j=1}^{n} |\phi_{j\varepsilon}|)$, let $M = \limsup_{\varepsilon \downarrow 0} M_\varepsilon$, which we showed is finite. Then, setting $L_\varepsilon = \sum_{j=1}^{m} |\lambda_{j\varepsilon}|$, it is easy to show that $\sum_{j=1}^{m(n+1)} |\theta_{j\varepsilon}| \le (2M+1)L_\varepsilon$ for all ε sufficiently small. Integrating by parts,

$$\begin{aligned}
\int_{R_k}^{R_{k+1}} \theta_\varepsilon(\omega) e^{-i\omega y_0}\, d\omega &= \frac{e^{-iR_k y_0}}{i y_0} \{\theta_\varepsilon(R_k) - \theta_\varepsilon(R_{k+1})\} \\
&\quad + \frac{1}{i y_0} \int_{R_k}^{R_{k+1}} e^{-i\omega y_0} \theta'_\varepsilon(\omega)\, d\omega.
\end{aligned}$$

Defining $\check{z} = \max_j |z_j|$, we have $|\theta_\varepsilon(R_k) - \theta_\varepsilon(R_{k+1})| \le \sum_j |\theta_{j\varepsilon}||1 - e^{i2\pi\varepsilon z_j/y_0}| \le 2\pi(2M+1)\check{z}L_\varepsilon\varepsilon/y_0$ and $|\theta'_\varepsilon(\omega)| \le (2M+1)\check{z}L_\varepsilon\varepsilon$ for all ε sufficiently small, so that

$$\left|\int_{R_k}^{R_{k+1}} \theta_\varepsilon(\omega)e^{-i\omega y_0}\,d\omega\right| \le \frac{4\pi}{y_0^2}(2M+1)\check{z}L_\varepsilon\varepsilon \tag{23}$$

for all ε sufficiently small. Setting $\beta = \frac{1}{2}$, inequalities (10) and (23) imply

$$\begin{aligned} I_3 &\le 2C_2(2M+1)\check{z}L_\varepsilon \sum_{k=k_\varepsilon}^{\infty} \frac{\varepsilon}{k^{\alpha+1}} \\ &\le 2\alpha^{-1}C_2(2M+1)\check{z}L_\varepsilon\left(\frac{8}{y_0}\right)^{\alpha}\varepsilon^{\alpha/2+1} \end{aligned} \tag{24}$$

for all ε sufficiently small. Similarly to (9), it is possible to show $L_\varepsilon\varepsilon^{\alpha/2} \ll \|\lambda_\varepsilon\|_f$ as $\varepsilon \downarrow 0$, so that by (9) and (24),

$$\frac{I_3}{\|\eta_\varepsilon\|_f\|\lambda_\varepsilon\|_f} \ll \varepsilon^{1-\alpha/2} \tag{25}$$

as $\varepsilon \downarrow 0$ uniformly in Λ_ε. Since $\alpha < 2$, this bound tends to 0 uniformly in Λ_ε. Applying (22) and (25) to (20) yields I_2 [defined in (18)] tending to 0 as $\varepsilon \downarrow 0$ uniformly in Λ_ε, which together with (18) and (19), implies

$$\lim_{\varepsilon\downarrow 0}\sup_{\Lambda_\varepsilon}|\operatorname{corr}\{e(N_\varepsilon), \Lambda_\varepsilon \cdot Z(F_\varepsilon)\}| = 0. \tag{26}$$

To finish the proof, it suffices to prove $e(N_\varepsilon)$ is asymptotically uncorrelated with all linear combinations of $Z(N_\varepsilon \cup F_\varepsilon)$. Specifically, defining $\Xi_\varepsilon = (\xi_{1\varepsilon}, \dots, \xi_{n\varepsilon})$, if we can show

$$\lim_{\varepsilon\downarrow 0}\sup_{\Lambda_\varepsilon, \Xi_\varepsilon}|\operatorname{corr}\{e(N_\varepsilon), \Lambda_\varepsilon \cdot Z(F_\varepsilon) - \Xi_\varepsilon \cdot Z(N_\varepsilon)\}| = 0, \tag{27}$$

then the theorem follows since

$$\frac{Ee(N_\varepsilon \cup F_\varepsilon)^2}{Ee(N_\varepsilon)^2} = 1 - \sup_{\Lambda_\varepsilon, \Xi_\varepsilon}\operatorname{corr}\{e(N_\varepsilon), \Lambda_\varepsilon \cdot Z(F_\varepsilon) - \Xi_\varepsilon \cdot Z(N_\varepsilon)\}^2.$$

Because $e(N_\varepsilon)$ is the error of a BLP based on N_ε, $\operatorname{corr}\{e(N_\varepsilon), \Xi_\varepsilon \cdot Z(N_\varepsilon)\} = 0$ for all Ξ_ε. Thus, (27) follows from (26) if

$$\operatorname{var}\{\Lambda_\varepsilon \cdot Z(F_\varepsilon)\} \ll \operatorname{var}\{\Lambda_\varepsilon \cdot Z(F_\varepsilon) - \Xi_\varepsilon \cdot Z(N_\varepsilon)\} \tag{28}$$

uniformly in Λ_ε and Ξ_ε. There is nothing to prove if $\Xi_\varepsilon = 0$, so assume $\Xi_\varepsilon \ne 0$ hereafter. Consider the Matérn spectral density $f_\alpha(\omega) = (1+\omega^2)^{-(\alpha+1)/2}$, for which the corresponding autocovariance function is $K_\alpha(x) = c_\alpha|x|^{\alpha/2}\mathcal{K}_{\alpha/2}(|x|)$, where $c_\alpha = \pi^{1/2}/\{2^{\alpha/2-1}\Gamma((\alpha+1)/2)\}$ [Stein (1999a), page 31]. I will write the subscript α to indicate quantities such as variances calculated under K_α. Since, by (6), $f(\omega) \asymp f_\alpha(\omega)$, (28) is equivalent to

$$\operatorname{var}_\alpha\{\Lambda_\varepsilon \cdot Z(F_\varepsilon)\} \ll \operatorname{var}_\alpha\{\Lambda_\varepsilon \cdot Z(F_\varepsilon) - \Xi_\varepsilon \cdot Z(N_\varepsilon)\} \tag{29}$$

uniformly in Λ_ε and Ξ_ε, which is in turn equivalent to

$$\limsup_{\varepsilon \downarrow 0} \sup_{\Lambda_\varepsilon, \Xi_\varepsilon} |\operatorname{corr}_\alpha\{\Lambda_\varepsilon \cdot Z(F_\varepsilon), \Xi_\varepsilon \cdot Z(N_\varepsilon)\}| < 1. \tag{30}$$

Define $\lambda_{\cdot\varepsilon} = \sum_{j=1}^m \lambda_{j\varepsilon}$, $\tilde{\lambda}_{j\varepsilon} = \lambda_{j\varepsilon} - \frac{1}{m}\lambda_{\cdot\varepsilon}$, $\tilde{L}_\varepsilon = \sum_{j=1}^m |\tilde{\lambda}_{j\varepsilon}|$ and $\tilde{\Lambda}_\varepsilon = (\tilde{\lambda}_{1\varepsilon}, \ldots, \tilde{\lambda}_{m\varepsilon})$. Using the series expansion for K_α [Stein (1999a), (15) page 32] and setting $b_\alpha = \pi/\{\Gamma(\alpha+1)\sin(\frac{1}{2}\pi\alpha)\}$ and $S_\alpha(\varepsilon) = -\sum_{j,k=1}^m \tilde{\lambda}_{j\varepsilon}\tilde{\lambda}_{k\varepsilon}|y_j - y_k|^\alpha$,

$$\operatorname{var}_\alpha\{\tilde{\Lambda}_\varepsilon \cdot Z(F_\varepsilon)\} - b_\alpha \varepsilon^\alpha S_\alpha(\varepsilon) \ll \varepsilon^2 \tilde{L}_\varepsilon^2.$$

Now $S_\alpha(\varepsilon)$ is nonnegative because $\sum_{j=1}^m \tilde{\lambda}_{j\varepsilon} = 0$, and $|x|^\alpha$ is a valid variogram for $\alpha \in (0,2)$ [Stein (1999a), page 37]. Furthermore, if $\tilde{L}_\varepsilon \neq 0$, $S_\alpha(\varepsilon)/\tilde{L}_\varepsilon^2$ is trivially bounded from above. It is also uniformly bounded from below: if $S_\alpha(\varepsilon)/\tilde{L}_\varepsilon^2$ tends to a limit along any sequence of ε values, then there is a further subsequence along which $\tilde{\Lambda}_\varepsilon/\tilde{L}_\varepsilon$ converges to some $\tilde{\Lambda} = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_m) \neq 0$ and, along this subsequence, by dominated convergence,

$$\frac{b_\alpha S_\alpha(\varepsilon)}{\tilde{L}_\varepsilon^2} \to \int_{-\infty}^{\infty} \left| \sum_{j=1}^m \tilde{\lambda}_j e^{i\omega x_j} \right|^2 |\omega|^{-\alpha-1}\, d\omega > 0.$$

Thus, no subsequence of $S_\alpha(\varepsilon)/\tilde{L}_\varepsilon^2$ can have 0 as its limit and

$$\operatorname{var}_\alpha\{\tilde{\Lambda}_\varepsilon \cdot Z(F_\varepsilon)\} \asymp \varepsilon^\alpha \tilde{L}_\varepsilon^2, \tag{31}$$

which holds even if $\tilde{L}_\varepsilon = 0$. Again using the series expansion for K_α, $|\operatorname{cov}_\alpha\{Z(y_0 + \varepsilon y_1), \tilde{\Lambda}_\varepsilon \cdot Z(F_\varepsilon)\}| \ll \varepsilon^\alpha \tilde{L}_\varepsilon$, so that $\operatorname{corr}_\alpha\{Z(y_0 + \varepsilon y_1), \tilde{\Lambda}_\varepsilon \cdot Z(F_\varepsilon)\} \to 0$ uniformly in $\tilde{\Lambda}_\varepsilon \neq 0$. Thus,

$$\operatorname{var}_\alpha\{\Lambda_\varepsilon \cdot Z(F_\varepsilon)\} \sim \lambda_{\cdot\varepsilon}^2 K_\alpha(0) + \operatorname{var}_\alpha\{\tilde{\Lambda}_\varepsilon \cdot Z(F_\varepsilon)\} \tag{32}$$

as $\varepsilon \downarrow 0$, uniformly in Λ_ε. Results similar to (31) and (32) apply to $\Xi_\varepsilon \cdot Z(N_\varepsilon)$.

Next, define $\xi_{\cdot\varepsilon} = \sum_{j=1}^n \xi_{j\varepsilon}$, $\tilde{\xi}_{j\varepsilon} = \xi_{j\varepsilon} - \frac{1}{n}\xi_{\cdot\varepsilon}$, $\tilde{X}_\varepsilon = \sum_{j=1}^n |\tilde{\xi}_{j\varepsilon}|$ and $\tilde{\Xi}_\varepsilon = (\tilde{\xi}_{1\varepsilon}, \ldots, \tilde{\xi}_{n\varepsilon})$ and consider

$$\begin{aligned} &\operatorname{cov}_\alpha\{\Lambda_\varepsilon \cdot Z(F_\varepsilon), \Xi_\varepsilon \cdot Z(N_\varepsilon)\} \\ &\quad = \lambda_{\cdot\varepsilon}\xi_{\cdot\varepsilon} K_\alpha(y_0 + \varepsilon y_1 - \varepsilon x_1) + \lambda_{\cdot\varepsilon} \operatorname{cov}_\alpha\{Z(y_0 + \varepsilon y_1), \tilde{\Xi}_\varepsilon \cdot Z(N_\varepsilon)\} \\ &\qquad + \xi_{\cdot\varepsilon} \operatorname{cov}_\alpha\{Z(\varepsilon x_1), \tilde{\Lambda}_\varepsilon \cdot Z(F_\varepsilon)\} + \operatorname{cov}_\alpha\{\tilde{\Lambda}_\varepsilon \cdot Z(F_\varepsilon), \tilde{\Xi}_\varepsilon \cdot Z(N_\varepsilon)\}. \end{aligned} \tag{33}$$

Since K_α has a bounded second derivative outside of a neighborhood of the origin, it is straightforward to obtain the following bounds:

$$|\lambda_{\cdot\varepsilon}\xi_{\cdot\varepsilon}K_\alpha(y_0+\varepsilon y_1-\varepsilon x_1)-\lambda_{\cdot\varepsilon}\xi_{\cdot\varepsilon}K_\alpha(y_0)| \ll \varepsilon|\lambda_{\cdot\varepsilon}||\xi_{\cdot\varepsilon}|,$$
$$|\lambda_{\cdot\varepsilon}\operatorname{cov}_\alpha\{Z(y_0+\varepsilon y_1),\tilde{\Xi}_\varepsilon\cdot Z(N_\varepsilon)\}| \ll \varepsilon|\lambda_{\cdot\varepsilon}|\tilde{X}_\varepsilon,$$
$$|\xi_{\cdot\varepsilon}\operatorname{cov}_\alpha\{Z(\varepsilon x_1),\tilde{\Lambda}_\varepsilon\cdot Z(F_\varepsilon)\}| \ll \varepsilon|\xi_{\cdot\varepsilon}|\tilde{L}_\varepsilon$$

and

$$|\operatorname{cov}_\alpha\{\tilde{\Lambda}_\varepsilon\cdot Z(F_\varepsilon),\tilde{\Xi}_\varepsilon\cdot Z(N_\varepsilon)\}| \ll \varepsilon^2\tilde{L}_\varepsilon\tilde{X}_\varepsilon$$

as $\varepsilon\downarrow 0$. Applying these bounds to (33) gives

$$\begin{aligned}(34)\qquad &|\operatorname{cov}_\alpha\{\Lambda_\varepsilon\cdot Z(F_\varepsilon),\Xi_\varepsilon\cdot Z(N_\varepsilon)\}-\lambda_{\cdot\varepsilon}\xi_{\cdot\varepsilon}K_\alpha(y_0)|\\ &\qquad \ll \varepsilon|\lambda_{\cdot\varepsilon}||\xi_{\cdot\varepsilon}|+\varepsilon|\lambda_{\cdot\varepsilon}|\tilde{X}_\varepsilon+\varepsilon|\xi_{\cdot\varepsilon}|\tilde{L}_\varepsilon+\varepsilon^2\tilde{L}_\varepsilon\tilde{X}_\varepsilon.\end{aligned}$$

Now, from (32),

$$\begin{aligned}&\frac{|\lambda_{\cdot\varepsilon}\xi_{\cdot\varepsilon}K_\alpha(y_0)|}{\sqrt{\operatorname{var}_\alpha\{\Lambda_\varepsilon\cdot Z(F_\varepsilon)\}\operatorname{var}_\alpha\{\Xi_\varepsilon\cdot Z(N_\varepsilon)\}}}\\ (35)\qquad &\sim\frac{|\lambda_{\cdot\varepsilon}\xi_{\cdot\varepsilon}K_\alpha(y_0)|}{\sqrt{\lambda_{\cdot\varepsilon}^2K_\alpha(0)+\operatorname{var}_\alpha\{\tilde{\Lambda}_\varepsilon\cdot Z(F_\varepsilon)\}}\sqrt{\xi_{\cdot\varepsilon}^2K_\alpha(0)+\operatorname{var}_\alpha\{\tilde{\Xi}_\varepsilon\cdot Z(N_\varepsilon)\}}}\\ &\le\frac{K_\alpha(y_0)}{K_\alpha(0)},\end{aligned}$$

which is in (0, 1) for all $y_0\neq 0$. And, since $\alpha<2$,

$$\frac{\varepsilon|\lambda_{\cdot\varepsilon}||\xi_{\cdot\varepsilon}|+\varepsilon|\lambda_{\cdot\varepsilon}|\tilde{X}_\varepsilon+\varepsilon|\xi_{\cdot\varepsilon}|\tilde{L}_\varepsilon+\varepsilon^2\tilde{L}_\varepsilon\tilde{X}_\varepsilon}{\sqrt{\lambda_{\cdot\varepsilon}^2+\varepsilon^\alpha\tilde{L}_\varepsilon^2}\sqrt{\xi_{\cdot\varepsilon}^2+\varepsilon^\alpha\tilde{X}_\varepsilon^2}}\to 0,$$

which, together with (32), (34) and (35), proves (30) and hence (27) and the theorem.

To prove that Theorem 1 also applies to ordinary kriging, note that by setting $\beta=\frac{1}{2}$, (13) and (16) together with (9) imply $\|\tilde{\eta}_\varepsilon\|_f^2\sim\|\eta_\varepsilon\|_f^2$ as $\varepsilon\downarrow 0$. Since $\tilde{\eta}_\varepsilon$ corresponds to the error of a linear unbiased predictor under the constant mean model, we have that the mean squared error of the ordinary kriging predictor based on $Z(N_\varepsilon)$ is at least $\|\eta_\varepsilon\|_f^2$ and at most $\|\tilde{\eta}_\varepsilon\|_f^2$, so that if (1) holds for the simple kriging predictor it also holds for the ordinary kriging predictor.

5.3. *Proof of Theorem* 2. Restricting N_ε to one point and F_ε to 2 allows us to make use of Lemma 1 to prove (1). Setting $y_2=0$ simplifies the calculations without changing any essential details. Specifically, defining $V(x)=K(0)-K(x)$, we will show that

$$
\begin{aligned}
W'_\varepsilon &= (W_{\varepsilon 1}, W_{\varepsilon 2}, W_{\varepsilon 3}, W_{\varepsilon 4}) \\
&= \left(\frac{Z(0) - Z(\varepsilon x_1)}{\sqrt{V(\varepsilon x_1)}}, Z(\varepsilon x_1), Z(y_0 + \varepsilon y_1), \frac{Z(y_0) - Z(y_0 + \varepsilon y_1)}{\sqrt{V(\varepsilon y_1)}} \right)'
\end{aligned} \tag{36}
$$

has limiting covariance matrix of the form given in (8), from which Theorem 2 readily follows.

Let us consider the easier parts of the proof first. Independent of ε, the variances of the elements of W_ε are 1, $K(0)$, $K(0)$ and 1, respectively. Since Z has a spectral density, K is continuous and $|K(y)| < K(0)$ for all $y \neq 0$. Thus, $\operatorname{cov}(W_{\varepsilon 2}, W_{\varepsilon 3}) \to K(y_0)$ as $\varepsilon \downarrow 0$, and the 2×2 matrix with $K(0)$ on the diagonals and $K(y_0)$ elsewhere is positive definite. Thus, it suffices to show that the other offdiagonal elements of the covariance matrix of W_ε tend to 0 as $\varepsilon \downarrow 0$. First, $\operatorname{cov}(W_{\varepsilon 1}, W_{\varepsilon 2}) = \frac{1}{2}\sqrt{V(\varepsilon x_1)/K(0)} \to 0$ as $\varepsilon \downarrow 0$. Similarly, $\operatorname{cov}(W_{\varepsilon 3}, W_{\varepsilon 4}) \to 0$ as $\varepsilon \downarrow 0$.

Now consider $\operatorname{cov}(W_{\varepsilon 1}, W_{\varepsilon 3})$. We have

$$
\operatorname{cov}\{Z(0) - Z(\varepsilon x_1), Z(y_0 + \varepsilon y_1)\} = \int_{\mathbb{R}^2} e^{-i\omega\cdot(y_0 + \varepsilon y_1)} (1 - e^{i\varepsilon\omega\cdot x_1}) f(\omega)\, d\omega,
$$

so that for $D(T) = \{\omega : |\omega \cdot x_1| \le T\}$,

$$
\begin{aligned}
\operatorname{cov}(W_{\varepsilon 1}, W_{\varepsilon 3})^2 &\le \frac{\{\int_{\mathbb{R}^2} |1 - e^{i\varepsilon\omega\cdot x_1}| f(\omega)\, d\omega\}^2}{K(0) \int_{\mathbb{R}^2} |1 - e^{i\varepsilon\omega\cdot x_1}|^2 f(\omega)\, d\omega} \\
&\le \frac{2\{\int_{D(T)} |1 - e^{i\varepsilon\omega\cdot x_1}| f(\omega)\, d\omega\}^2}{K(0) \int_{D(T)} |1 - e^{i\varepsilon\omega\cdot x_1}|^2 f(\omega)\, d\omega} \\
&\quad + \frac{2\{\int_{D(T)^c} |1 - e^{i\varepsilon\omega\cdot x_1}| f(\omega)\, d\omega\}^2}{K(0) \int_{D(T)^c} |1 - e^{i\varepsilon\omega\cdot x_1}|^2 f(\omega)\, d\omega}
\end{aligned} \tag{37}
$$

for all T sufficiently large (to guarantee $\int_{D(T)} |1 - e^{i\varepsilon\omega\cdot x_1}|^2 f(\omega)\, d\omega > 0$). Because $\varepsilon^{-1}|1 - e^{i\varepsilon\omega\cdot x_1}| \le |\omega \cdot x_1|$ and $\varepsilon^{-1}|1 - e^{i\varepsilon\omega\cdot x_1}| \to |\omega \cdot x_1|$ as $\varepsilon \downarrow 0$, by dominated convergence,

$$
\lim_{\varepsilon \downarrow 0} \frac{\{\int_{D(T)} |1 - e^{i\varepsilon\omega\cdot x_1}| f(\omega)\, d\omega\}^2}{\int_{D(T)} |1 - e^{i\varepsilon\omega\cdot x_1}|^2 f(\omega)\, d\omega} = \frac{\{\int_{D(T)} |\omega \cdot x_1| f(\omega)\, d\omega\}^2}{\int_{D(T)} |\omega \cdot x_1|^2 f(\omega)\, d\omega}. \tag{38}
$$

By the Cauchy–Schwarz inequality,

$$
\begin{aligned}
&\frac{\{\int_{D(T)^c} |1 - e^{i\varepsilon\omega\cdot x_1}| f(\omega)\, d\omega\}^2}{\int_{D(T)^c} |1 - e^{i\varepsilon\omega\cdot x_1}|^2 f(\omega)\, d\omega} \\
&\quad \le \frac{\int_{D(T)^c} |1 - e^{i\varepsilon\omega\cdot x_1}|^2 f(\omega)\, d\omega \int_{D(T)^c} f(\omega)\, d\omega}{\int_{D(T)^c} |1 - e^{i\varepsilon\omega\cdot x_1}|^2 f(\omega)\, d\omega} \\
&\quad = \int_{D(T)^c} f(\omega)\, d\omega.
\end{aligned} \tag{39}
$$

From (37)–(39), we will have $\operatorname{cov}(W_{\varepsilon 1}, W_{\varepsilon 3}) \to 0$ as $\varepsilon \downarrow 0$ if the right-hand sides of (38) and (39) tend to 0 as $T \to \infty$. The integrability of f implies $\int_{D(T)^c} f(\omega)\,d\omega \to 0$ as $T \to \infty$, so consider the right-hand side of (38). Let A be the 2×2 matrix with first row given by x_1, orthogonal rows and determinant of 1 and set $v = (v_1, v_2)' = A\omega$. Define $\bar{f}(v_1) = \int_{-\infty}^{\infty} f(A^{-1}v)\,dv_2$. Up to a linear rescaling, $\bar{f}$ is the spectral density of the process Z along the x_1 direction, so it is integrable. In addition, because Z is not mean square differentiable in any direction, $\int_0^\infty v_1^2 \bar{f}(v_1)\,dv_1 = \infty$. Then for any even function g, $\int_{D(T)} g(\omega \cdot x_1) f(\omega)\,d\omega = 2\int_0^T g(v_1)\bar{f}(v_1)\,dv_1$, so that for $0 < S < T$,

$$
\begin{aligned}
\frac{\{\int_{D(T)} |\omega \cdot x_1| f(\omega)\,d\omega\}^2}{\int_{D(T)} |\omega \cdot x_1|^2 f(\omega)\,d\omega} &= \frac{\{\int_0^T v_1 \bar{f}(v_1)\,dv\}^2}{\int_0^T v_1^2 \bar{f}(v_1)\,dv_1} \\
&= \frac{\{\int_0^S v_1 \bar{f}(v_1)\,dv_1 + \int_S^T v_1 \bar{f}(v_1)\,dv_1\}^2}{\int_0^T v_1^2 \bar{f}(v_1)\,dv_1}.
\end{aligned}
\tag{40}
$$

If we can show that

$$
\lim_{S\to\infty} \lim_{T\to\infty} \frac{\{\int_0^S v_1 \bar{f}(v_1)\,dv_1 + \int_S^T v_1 \bar{f}(v_1)\,dv_1\}^2}{\int_0^T v_1^2 \bar{f}(v_1)\,dv_1} = 0,
\tag{41}
$$

then the right-hand side of (38) will tend to 0 as $T \to \infty$. To prove (41), expand the square in the numerator and consider each term separately. First, by the Cauchy–Schwarz inequality,

$$
\lim_{T\to\infty} \frac{\{\int_0^S v_1 \bar{f}(v_1)\,dv_1\}^2}{\int_0^T v_1^2 \bar{f}(v_1)\,dv_1} = \lim_{T\to\infty} \frac{\int_0^S v_1^2 \bar{f}(v_1)\,dv_1 \int_0^S \bar{f}(v_1)\,dv_1}{\int_0^T v_1^2 \bar{f}(v_1)\,dv_1} = 0.
$$

Again by the Cauchy–Schwarz inequality,

$$
\frac{\{\int_S^T v_1 \bar{f}(v_1)\,dv_1\}^2}{\int_0^T v_1^2 \bar{f}(v_1)\,dv_1} \le \frac{\int_S^T v_1^2 \bar{f}(v_1)\,dv_1 \int_S^T \bar{f}(v_1)\,dv_1}{\int_0^T v_1^2 \bar{f}(v_1)\,dv_1} \le \int_S^T \bar{f}(v_1)\,dv_1,
$$

which tends to 0 when one takes $\lim_{S\to\infty} \lim_{T\to\infty}$ since $\bar{f}$ is integrable. Finally,

$$
\begin{aligned}
\frac{\int_0^S v_1 \bar{f}(v_1)\,dv_1 \int_S^T v_1 \bar{f}(v_1)\,dv_1}{\int_0^T v_1^2 \bar{f}(v_1)\,dv_1} &\le \frac{\int_0^S v_1 \bar{f}(v_1)\,dv_1 \int_S^T v_1 \bar{f}(v_1)\,dv_1}{\int_S^T v_1^2 \bar{f}(v_1)\,dv_1} \\
&\le \frac{\int_0^S v_1 \bar{f}(v_1)\,dv_1}{S} \cdot \frac{\int_S^T v_1 \bar{f}(v_1)\,dv_1}{\int_S^T v_1 \bar{f}(v_1)\,dv_1} \\
&\le \frac{1}{S} \int_0^{S^{1/2}} v_1 \bar{f}(v_1)\,dv_1 + \frac{1}{S} \int_{S^{1/2}}^S v_1 \bar{f}(v_1)\,dv_1 \\
&\le \frac{1}{S^{1/2}} \int_0^{S^{1/2}} \bar{f}(v_1)\,dv_1 + \int_{S^{1/2}}^S \bar{f}(v_1)\,dv_1,
\end{aligned}
$$

which tends to 0 as $S \to \infty$, and (41) follows. Thus, $\operatorname{cov}(W_{\varepsilon 1}, W_{\varepsilon 3}) \to 0$ as $\varepsilon \downarrow 0$. Similarly, $\operatorname{cov}(W_{\varepsilon 2}, W_{\varepsilon 4}) \to 0$ as $\varepsilon \downarrow 0$.

We will need the following lemma to handle $\operatorname{cov}(W_{\varepsilon 1}, W_{\varepsilon 4})$:

LEMMA 2. *If Z is not mean square differentiable in the direction x, then*

$$\lim_{\varepsilon \downarrow 0} \frac{\varepsilon^2}{V(\varepsilon x)} = 0.$$

To prove the lemma, first note that the assumption on Z is equivalent to

$$\int_{\mathbb{R}^2} |\omega \cdot x|^2 f(\omega)\, d\omega = \infty. \tag{42}$$

If $\limsup_{\varepsilon \downarrow 0} \frac{\varepsilon^2}{V(\varepsilon x)} > 0$, then there must exist some sequence $\varepsilon_n \downarrow 0$ along which $\lim_{n \to \infty} \frac{V(\varepsilon_n x)}{\varepsilon_n^2} = C$ for some finite C, or

$$\lim_{n \to \infty} \int_{\mathbb{R}^2} \frac{|1 - e^{i\varepsilon_n \omega \cdot x}|^2}{\varepsilon_n^2} f(\omega)\, d\omega = C.$$

But for any finite T, by dominated convergence,

$$\begin{aligned} C &= \lim_{n \to \infty} \int_{\mathbb{R}^2} \frac{|1 - e^{i\varepsilon_n \omega \cdot x}|^2}{\varepsilon_n^2} f(\omega)\, d\omega \\ &\geq \lim_{n \to \infty} \int_{|\omega| < T} \frac{|1 - e^{i\varepsilon_n \omega \cdot x}|^2}{\varepsilon_n^2} f(\omega)\, d\omega \\ &= \int_{|\omega| < T} |\omega \cdot x|^2 f(\omega)\, d\omega \end{aligned}$$

for all T, which contradicts (42), and the lemma is proven.

Consider

$$\begin{aligned} &\operatorname{cov}\{Z(0) - Z(\varepsilon x_1), Z(y_0) - Z(y_0 + \varepsilon y_1)\} \\ &\quad = \int_{\mathbb{R}^2} e^{-i\omega \cdot y_0} (1 - e^{i\varepsilon\omega \cdot x_1})(1 - e^{-i\varepsilon\omega \cdot y_1}) f(\omega)\, d\omega. \end{aligned}$$

Define $f_1(\omega) = \min(f(\omega), 1)$, and write cov_1 to indicate covariances calculated under the spectral density f_1. Then (2) and f integrable imply that $f(\omega) = f_1(\omega)$ outside some bounded set, and it easily follows that

$$\begin{aligned} &\operatorname{cov}\{Z(0) - Z(\varepsilon x_1), Z(y_0) - Z(y_0 + \varepsilon y_1)\} \\ &\quad = \operatorname{cov}_1\{Z(0) - Z(\varepsilon x_1), Z(y_0) - Z(y_0 + \varepsilon y_1)\} + O(\varepsilon^2). \end{aligned} \tag{43}$$

Because Z is not mean square differentiable in any direction, Lemma 2 implies the $O(\varepsilon^2)$ remainder in (43) makes no contribution to $\lim_{\varepsilon \downarrow 0} \operatorname{cov}(W_{\varepsilon 1}, W_{\varepsilon 4})$.

We proceed by rotating coordinates so that one of the frequency axes points in the direction of y_0. Specifically, let B be the 2×2 orthogonal matrix with determinant 1 and first row equal to y_0 and set $\tau = (\tau_1, \tau_2)' = B\omega$. Then, defining $H_\varepsilon(\tau_1, \tau_2) = (1 - e^{i\varepsilon(B^{-1}\tau)\cdot x_1})(1 - e^{-i\varepsilon(B^{-1}\tau)\cdot y_1})$,

$$
\begin{aligned}
&\operatorname{cov}_1\{Z(0) - Z(\varepsilon x_1), Z(y_0) - Z(y_0 + \varepsilon y_1)\} \\
&\quad = \int_{\mathbb{R}^2} e^{-i\tau_1} H_\varepsilon(\tau_1, \tau_2) f_1(B^{-1}\tau)\, d\tau \\
&\quad = \int_{\mathbb{R}} \sum_{k=-\infty}^{\infty} \int_{2\pi k}^{2\pi(k+1)} e^{-i\tau_1} H_\varepsilon(\tau_1, \tau_2) f_1(B^{-1}\tau)\, d\tau_1\, d\tau_2.
\end{aligned}
$$

Define the function g on $\mathbb{R}^2$ by, for $2\pi k \le \tau_1 < 2\pi(k+1)$, $g(B^{-1}\tau) = 1 - f_1(B^{-1}(2\pi k, \tau_2)')/f_1(B^{-1}\tau)$ if $f_1(B^{-1}\tau) > 0$ and 0 otherwise. We have

$$
\begin{aligned}
&\operatorname{cov}_1\{Z(0) - Z(\varepsilon x_1), Z(y_0) - Z(y_0 + \varepsilon y_1)\} \\
&\quad = \int_{\mathbb{R}} \sum_{k=-\infty}^{\infty} f_1\left(B^{-1}\begin{pmatrix} 2\pi k \\ \tau_2 \end{pmatrix}\right) \int_{2\pi k}^{2\pi(k+1)} e^{-i\tau_1} H_\varepsilon(\tau_1, \tau_2)\, d\tau_1\, d\tau_2 \\
&\qquad + \int_{\mathbb{R}^2} e^{-i\omega\cdot y_0}(1 - e^{i\varepsilon\omega\cdot x_1})(1 - e^{-i\varepsilon\omega\cdot y_1}) f_1(\omega) g(\omega)\, d\omega.
\end{aligned}
\tag{44}
$$

By (2), $g(\omega) \to 0$ as $\omega \to \infty$. Thus, given $\delta > 0$, we can find $T < \infty$ such that $g(\omega) < \delta$ for $|\omega| > T$. Then

$$
\begin{aligned}
&\left| \int_{\mathbb{R}^2} e^{-i\omega\cdot y_0}(1 - e^{i\varepsilon\omega\cdot x_1})(1 - e^{-i\varepsilon\omega\cdot y_1}) f_1(\omega) g(\omega)\, d\omega \right| \\
&\quad \le \varepsilon^2 \int_{|\omega|\le T} |\omega \cdot x_1||\omega \cdot y_1| f_1(\omega) |g(\omega)|\, d\omega \\
&\qquad + 4\delta \int_{|\omega|>T} |1 - e^{i\varepsilon\omega\cdot x_1}||1 - e^{-i\varepsilon\omega\cdot y_1}| f_1(\omega)\, d\omega.
\end{aligned}
$$

By the Cauchy–Schwarz inequality and $f_1 \le f$, $\int_{|\omega|>T} |1 - e^{i\varepsilon\omega\cdot x_1}||1 - e^{-i\varepsilon\omega\cdot y_1}| \times f_1(\omega)\, d\omega \le \sqrt{V(\varepsilon x_1) V(\varepsilon y_1)}$, which, together with Lemma 2, implies

$$
\limsup_{\varepsilon\downarrow 0} \frac{|\int_{\mathbb{R}^2} e^{-i\omega\cdot y_0}(1 - e^{i\varepsilon\omega\cdot x_1})(1 - e^{-i\varepsilon\omega\cdot y_1}) f_1(\omega) g(\omega)\, d\omega|}{\sqrt{V(\varepsilon x_1) V(\varepsilon y_1)}} \le 4\delta.
\tag{45}
$$

Since δ is arbitrary, this lim sup must in fact be 0.

Now return to the the first term on the right-hand side of (44). Integrating by parts,

$$
\begin{aligned}
\int_{2\pi k}^{2\pi(k+1)} e^{-i\tau_1} H_\varepsilon(\tau_1, \tau_2)\, d\tau_1 &= i H_\varepsilon\big(2\pi(k+1), \tau_2\big) - i H_\varepsilon(2\pi k, \tau_2) \\
&\quad - i \int_{2\pi k}^{2\pi(k+1)} e^{-i\tau_1} \frac{\partial}{\partial \tau_1} H_\varepsilon(\tau_1, \tau_2)\, d\tau_1.
\end{aligned}
$$

There exists finite C independent of ε and τ such that

$$\left|\frac{\partial}{\partial \tau_1} H_\varepsilon(\tau_1, \tau_2)\right| \le C\varepsilon\{|1 - e^{i\varepsilon(B^{-1}\tau)\cdot x_1}| + |1 - e^{-i\varepsilon(B^{-1}\tau)\cdot y_1}|\},$$

which implies

$$\begin{aligned}&\left|\int_{2\pi k}^{2\pi(k+1)} e^{-i\tau_1} H_\varepsilon(\tau_1, \tau_2)\, d\tau_1\right| \\ &\qquad \le 4\pi C\varepsilon\{|1 - e^{i\varepsilon(B^{-1}\tau)\cdot x_1}| + |1 - e^{-i\varepsilon(B^{-1}\tau)\cdot y_1}|\}.\end{aligned} \tag{46}$$

We can choose T finite so that if $2\pi k \le \tau_1 \le 2\pi(k+1)$, then $f_1(B^{-1}(2\pi k, \tau_2)') \le 2f(B^{-1}\tau)$ whenever $|\tau| > T$. Applying this result and (46) to the first term on the right-hand side of (44) and changing variables back to $\omega = B^{-1}\tau$, we get

$$\begin{aligned}&\left|\int_{\mathbb{R}} \sum_{k=-\infty}^{\infty} f_1\left(B^{-1}\begin{pmatrix}2\pi k\\ \tau_2\end{pmatrix}\right) \int_{2\pi k}^{2\pi(k+1)} e^{-i\tau_1} H_\varepsilon(\tau_1, \tau_2)\, d\tau_1\, d\tau_2\right| \\ &\qquad \le 8\pi C\varepsilon \int_{\mathbb{R}^2} f(\omega)\{|1 - e^{i\varepsilon\omega\cdot x_1}| + |1 - e^{-i\varepsilon\omega\cdot y_1}|\}\, d\omega + O(\varepsilon^2) \\ &\qquad \le 8\pi C\varepsilon\{\sqrt{V(\varepsilon x_1)} + \sqrt{V(\varepsilon y_1)}\}\sqrt{\int_{\mathbb{R}^2} f(\omega)\, d\omega} + O(\varepsilon^2),\end{aligned} \tag{47}$$

where the last step uses the Cauchy–Schwarz inequality. From Lemma 2 and (47), it follows that

$$\limsup_{\varepsilon\downarrow 0} \frac{|\int_{\mathbb{R}} \sum_{k=-\infty}^{\infty} f_1(B^{-1}\binom{2\pi k}{\tau_2}) \int_{2\pi k}^{2\pi(k+1)} e^{-i\tau_1} H_\varepsilon(\tau_1, \tau_2)\, d\tau_1\, d\tau_2|}{\sqrt{V(\varepsilon x_1)V(\varepsilon y_1)}} = 0.$$

Together with (44) and (45), this limit implies

$$\limsup_{\varepsilon\downarrow 0} \frac{\operatorname{cov}_1\{Z(0) - Z(\varepsilon x_1), Z(y_0) - Z(y_0 + \varepsilon y_1)\}}{\sqrt{V(\varepsilon x_1)V(\varepsilon y_1)}} = 0,$$

which together with (43) and Lemma 2, implies $\lim_{\varepsilon\downarrow 0} \operatorname{cov}\{W_{\varepsilon 1}, W_{\varepsilon 4}\} = 0$.

Theorem 2 applies to ordinary kriging as well. Specifically, $Z(\varepsilon x_1)$ is an asymptotically optimal linear predictor of $Z(0)$ based on $Z(N_\varepsilon \cup F_\varepsilon)$ when the mean of Z is assumed to be 0, so since it is a linear unbiased predictor when the mean is an unknown constant, $Z(\varepsilon x_1)$ must also be asymptotically optimal with respect to this more restricted class of predictors.

Acknowledgment. The author thanks Steven Lalley for help with the proof of Theorem 2.

REFERENCES

BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Regular Variation*. *Encyclopedia of Mathematics and Its Applications* **27**. Cambridge Univ. Press, Cambridge. MR0898871

CHILÈS, J.-P. and DELFINER, P. (1999). *Geostatistics*: *Modeling Spatial Uncertainty*. Wiley, New York. MR1679557

EMERY, X. (2009). The kriging update equations and their application to the selection of neighbouring data. *Comput*. *Geosci*. **13** 269–280.

GNEITING, T. (2002). Nonseparable, stationary covariance functions for space–time data. *J. Amer. Statist*. *Assoc*. **97** 590–600. MR1941475

JOURNEL, A. G. and HUIJBREGTS, C. J. (1978). *Mining Geostatistics*. Academic Press, New York.

KELBERT, M. Y., LEONENKO, N. N. and RUIZ-MEDINA, M. D. (2005). Fractional random fields associated with stochastic fractional heat equations. *Adv. in Appl*. *Probab*. **37** 108–133. MR2135156

MEMARSADEGHI, N. and MOUNT, D. (2007). Efficient implementation of an optimal interpolator for large spatial data sets. In *Computational Science—ICCS* 2007 (Y. Shi, G. van Albada, J. Dongarra and P. Sloot, eds.). *Lecture Notes in Computer Science* **4488** 503–510. Springer, Berlin.

OLVER, F. W. J., LOZIER, D. W., BOISVERT, R. F. and CLARK, C. W. (2010). *NIST Handbook of Mathematical Functions*. Cambridge Univ. Press, New York.

RAMM, A. G. (2005). *Random Fields Estimation*. World Scientific, Singapore.

STEIN, M. L. (1999a). *Interpolation of Spatial Data*: *Some Theory for Kriging*. Springer, New York. MR1697409

STEIN, M. L. (1999b). Predicting random fields with increasing dense observations. *Ann*. *Appl*. *Probab*. **9** 242–273. MR1682572

STEIN, M. L. (2002). The screening effect in kriging. *Ann*. *Statist*. **30** 298–323. MR1892665

STEIN, M. L. (2005). Space–time covariance functions. *J. Amer. Statist*. *Assoc*. **100** 310–321. MR2156840

STEIN, M. L. (2008). A modeling approach for large spatial datasets. *J. Korean Statist*. *Soc*. **37** 3–10. MR2420389

STEIN, M. L. (2011). On a class of space–time intrinsic random functions. *Bernoulli*. To appear.

STEIN, M. L. and HANDCOCK, M. S. (1989). Some asymptotic properties of kriging when the covariance function is misspecified. *Math*. *Geol*. **21** 171–190. MR0985969

XUE, Y. and XIAO, Y. (2011). Fractal and smoothness properties of space–time Gaussian models. *Frontiers Math. China*. To appear. DOI:10.1007/s11464-011-0126-9.

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