The Annals of Statistics
1995, Vol. 23, No. 6, 1991–2012

LOCALLY LATTICE SAMPLING DESIGNS FOR ISOTROPIC RANDOM FIELDS[1]

BY MICHAEL L. STEIN

University of Chicago

For predicting $\int_G v(x)Z(x)\,dx$, where v is a fixed known function and Z is a stationary random field, a good sampling design should have a greater density of observations where v is relatively large in absolute value. Designs using this idea when $G = [0, 1]$ have been studied for some time. For G a region in two dimensions, very little is known about the statistical properties of cubature rules based on designs with varying density. This work proposes a class of designs that are locally parallelogram lattices but whose densities can vary. The asymptotic variance of the cubature error for these designs is obtained for a class of isotropic random fields and an asymptotically optimal sequence of cubature rules within this class is found. I conjecture that this sequence of cubature rules is asymptotically optimal with respect to all cubature rules.

1. Introduction. This work studies sampling designs for predicting integrals of a stationary, isotropic random field Z in two dimensions. Specifically, consider predicting $\int_G v(x)Z(x)\,dx$ for v a fixed known smooth function and G a bounded region of integration. Lattice-based designs have certain desirable properties, not the least of which is that their properties can be studied using spectral means. However, designs using a fixed lattice throughout G are inefficient when v is not constant, since it is better to have a greater density of observations where $|v(x)|$ is relatively large. Here, I introduce designs and corresponding cubature rules that are locally parallelogram lattices but whose densities vary with x. By a parallelogram lattice I mean any lattice that can be obtained by linearly transforming a square lattice. The equilateral triangular lattice is a special case of a parallelogram lattice and will be of particular importance here. For Z with spectral density f satisfying $g(|\omega|) = f(\omega)$, where g is regularly varying with exponent $-p$, $2 < p < 4$, I obtain the asymptotic mean squared error for these cubature rules and give asymptotically optimal cubature rules within this class of locally lattice designs. Furthermore, I conjecture that these rules, which are based on designs that are locally an equilateral triangular lattice, are asymptotically optimal with respect to all cubature rules based on point evaluations of Z.

Received October 1993; revised November 1994.

[1]This research was supported in part by NSF Grant DMS-92-04504. This manuscript was prepared using computer facilities supported in part by NSF Grants DMS-89-05292, DMS-87-03942 and DMS-86-01732 awarded to the Department of Statistics at the University of Chicago, and by the University of Chicago Block Fund.

AMS 1991 *subject classifications*. Primary 62M40; secondary 65D32.

Key words and phrases. Cubature, Epstein zeta-function, regular variation, spatial statistics.

Since useful finite sample results are difficult to obtain, nearly all theoretical work on stochastic evaluation of cubature errors is asymptotic. Some early work on this problem was done by Quenouille (1949), Matérn (1960) and Dalenius, Hájek and Zubrzycki (1961). All of these authors considered predicting $\int_G Z(x)\,dx$ and letting G grow with the number of observations, what Cressie (1993) calls increasing-domain asymptotics. Matérn (1960) found that equilateral triangular lattice designs work well for many isotropic processes. Dalenius, Hájek and Zubrzycki (1961) showed that there are isotropic processes for which other lattices perform better asymptotically than the equilateral triangular lattice.

A different asymptotic approach and the one used here is to fix the region of integration G and to increase the density of observations as their number increases. Cressie (1993) calls this infill asymptotics, although fixed-domain asymptotics is perhaps a good alternative name to highlight the contrast with increasing-domain asymptotics. Note that while statisticians first studied increasing-domain asymptotics, this approach is quite unnatural for a numerical analyst who thinks in terms of some fixed integral to approximate. Indeed, I am unaware of any work in the numerical analysis literature that uses increasing-domain asymptotics. Tubilla (1975) appears to be the first to have studied the error variance of various sampling designs for predicting $\int_G Z(x)\,dx$ using fixed-domain asymptotics. Schoenfelder and Cambanis (1982) and Schoenfelder (1982) considered the more general problem of predicting $\int_B v(x)Z(x)\,dx$. These works show that in two or more dimensions, stratified designs can yield better predictions than systematic designs if Z is sufficiently smooth. Stein (1993, 1995a) showed that the problem with centered systematic sampling designs for Z smooth and G the unit cube is that the coefficients used in simple integration rules do not handle boundary effects well and that by adjusting the coefficients near the boundary, systematic designs do at least as well as stratified designs.

Section 2 provides some background on deterministic and stochastic approaches to evaluating cubature rules. Section 3 defines a class of cubature rules based on locally lattice designs and obtains their asymptotic mean squared error. Section 4 gives an asymptotically optimal sequence of rules within this class and discusses the evidence for believing this sequence is asymptotically optimal with respect to all rules. Section 5 discusses some alternative systematic designs.

2. Background. This section reviews some basic concepts of deterministic and stochastic approaches to defining optimal and asymptotically optimal cubature formulae. The standard deterministic approach [Levin and Girshovich (1979)] is to choose some class of functions H on some set G and then consider for $h \in H$, an n-point cubature rule of the form

$$L(h; X, C) = \sum_{k=1}^{n} c_k h(x_k),$$

where $X = (x_1, \dots, x_n)$, each $x_i \in G$ and $C = (c_1, \dots, c_n)$. The cubature error is

$$e(h; X, C) = \int_G h(x)\,dx - L(h; X, C).$$

Note that X and C are not allowed to depend on the particular h in H selected. The quality of a cubature rule is measured by

$$e(H; X, C) = \sup_{h \in H} |e(h; X, C)|.$$

An n-point cubature rule (X^*, C^*) is called optimal among all n-point rules if

$$e(H; X^*, C^*) = \inf_{(X, C) \in A_n} e(H; X, C),$$

where $A_n = G^n \times \mathbb{R}^n$. In some circumstances, it may be desirable to allow observations outside of G, in which case, we would take $A_n = F^n \times \mathbb{R}^n$, where F is some set containing G. We can also consider optimizing just over C, the coefficients, taking X as fixed. Following Davis and Rabinowitz (1984), I will call C^* relatively optimal if, for given $X \in G^n$,

$$e(H; X, C^*) = \inf_{C \in \mathbb{R}^n} e(H; X, C).$$

To define asymptotic optimality of cubature rules, we need to consider sequences of rules. In some constructions of sequences of rules, it may be difficult to obtain rules based on exactly n points for any given n. Thus, I will consider a sequence of rules $\{(X_n, C_n)\}$, $n = 1, 2, \dots,$ such that N_n, the number of points in rule n, tends to infinity as $n \to \infty$. This sequence of rules is called asymptotically optimal if

$$\lim_{n \to \infty} \frac{e(H; X_n, C_n)}{\inf_{(X, C) \in A_{N_n}} e(H; X, C)} = 1.$$

For a fixed sequence of designs $X_1, X_2, \dots$ with N_n points in X_n, I will call $C_1, C_2, \dots$ asymptotically relatively optimal if

$$\inf_{n \to \infty} \frac{e(H; X_n, C_n)}{\inf_{C \in \mathbb{R}^{N_n}} e(H; X_n, C)} = 1.$$

In comparing the asymptotic performance of different sampling designs, it is important to keep in mind whether or not the corresponding coefficients are asymptotically relatively optimal. For example, consider a sequence of partitions of G where the largest diameter of the elements of the nth partition tends to 0 as n tends to infinity. Define the nth in a sequence of systematic designs by placing an observation at the center of mass of each element of the nth partition. Define the corresponding nth stratified design by placing one observation randomly in each element of the nth partition. Tubilla (1975) and Schoenfelder (1982) show that for partitions of the unit square into a lattice of squares (Schoenfelder also considers more general partitions) and using naively chosen coefficients, systematic designs can have a slower rate of convergence than the corresponding stratified designs. Stein (1993, 1995a) demonstrates that this apparent superiority of stratified designs disappears if asymptotically relatively optimal coefficients are used with the systematic designs. More generally, I conjecture that under mild conditions on Z and the sequence of partitions, systematic designs will do at least as well asymptotically as the corresponding stratified designs if asymptotically relatively optimal coefficients are used.

The choice of H has a profound effect on which cubature rules are judged good; see Section 5 for some further discussion. Because the focus in this paper is on isotropic processes, I will only mention here classes H that have some sort of invariance under rotations. For example, Babenko (1977) obtained asymptotically optimal cubature rules for two-dimensional integrals with H the class of functions h satisfying $|h(x) - h(y)| \le \psi(|x - y|)$, where ψ is a modulus of continuity and $|x - y|$ indicates Euclidean distance between x and y. Due to their close relationship to the stochastic measures of good rules, classes of functions defined by bounds on Sobolev norms have a closer connection to this work [Wahba (1990)]. An example of such a space for $G \subset \mathbb{R}^2$ is all functions satisfying

$$\left[\int_G \sum_{\alpha=1}^{m} \binom{m}{\alpha} \left| \frac{\partial^m}{\partial x_1^{\alpha}\, \partial x_2^{m-\alpha}} h(x_1, x_2) \right|^q dx \right]^{1/q} \le 1, \tag{2.1}$$

where $x = (x_1, x_2)$. In statistical settings, only the case $q = 2$ has been considered. A generalization of particular relevance here is classes of functions of the form $v(x)h(x)$, where v is fixed and known everywhere on G and h satisfies a constraint such as (2.1). Polovinkin (1989) studies classes of functions of this form and derives certain properties that asymptotically optimal rules must possess without actually finding such rules.

The stochastic approach used here considers the function being integrated a random field with finite second moments. The quality of a cubature rule (X,C) is measured by its mean squared error and (relative) optimality and asymptotic (relative) optimality are defined using obvious analogs to their definitions in the deterministic case. Sacks and Ylvisaker (1971) note the close connection between the stochastic and deterministic approaches when the space of deterministic functions is a Hilbert space with a reproducing kernel [Wahba (1990)]. Specifically, if Z has mean identically 0,

$$E\left[e(Z; X, C)^2\right] = e(H; X, C)^2, \tag{2.2}$$

where H is made up of the elements of a Hilbert space whose norms are at most 1 and the reproducing kernel of the Hilbert space is the covariance function of Z. For example, if $G = [0, 1]$ and H is the set of absolutely continuous functions h on $[0, 1]$ for which $\int_0^1 [h'(t)]^2\, dt \le 1$, then the reproducing kernel is proportional to $\min(s, t)$ so that Z in (2.2) can be taken to be Brownian motion. For the space of functions given by (2.1) with $q = 2$, the corresponding reproducing kernel has spectral “density” proportional to

$|\omega|^{-2m}$. Note that this function is not integrable in a neighborhood of the origin, which is related to the fact that (2.1) only defines a seminorm and not a norm [see Wahba (1990) for further details]. For our purposes, only the high frequency behavior of the spectral density matters, so I will ignore this point here. Since m is an integer, we see that there is no overlap between (2.1) and the class of spectral densities studied here, for which $g(|\omega|) = f(\omega)$ is regularly varying with exponent $-p$, $2 < p < 4$. Even if the restriction to m an integer is removed in (2.1), which is most conveniently done by considering the reproducing kernel or its spectral density directly, the classes of functions one obtains is still in some ways more narrow than I consider here. In particular, only requiring g, as in Stein (1993), to be regularly varying allows consideration of many processes (or, equivalently, many reproducing kernel Hilbert spaces) that have not, to my knowledge, been studied in the cubature literature using deterministic error bounds.

3. Main results. Consider predicting $\int_G v(x)Z(x)\,dx$ for v and G sufficiently well behaved and Z a stationary isotropic random field. It is plausible that a good design should place more observations in areas of G for which $|v(x)|$ is relatively large. Sacks and Ylvisaker (1971), Eubank, Smith and Smith (1981), Cambanis (1985), Benhenni and Cambanis (1992), Stein (1995b) and Pitt, Robeva and Wang (1995) studied this problem for one-dimensional integrals. In particular, Sacks and Ylvisaker (1971) and Eubank, Smith and Smith (1981) showed that systematic designs of varying density yield asymptotically optimal quadrature rules in some circumstances. Thus, one might expect an appropriate systematic design of varying density to provide asymptotically optimal cubature rules for multidimensional integrals. Other than some work by Schoenfelder (1982) extending results of Tubilla (1975), there does not appear to be any theoretical work on the mean squared error of cubature rules based on systematic designs of varying density for isotropic processes. Ylvisaker (1975) obtained some intriguing results on systematic designs for a class of anisotropic processes, but I will argue in Section 5 that these results depend critically on the nature of the anisotropy and so are not helpful for studying isotropic processes. Using deterministic error bounds, Babenko (1976) obtained asymptotically optimal cubature formulae for integrals of the form $\int_G v(x)h(x)\,dx$, where v is a fixed function and h is in a class of functions defined using a modulus of continuity constraint. While Babenko's results cannot be applied here, the basic approach can be. Specifically, partition G into regions on which v is nearly constant and then use a lattice design within each of these regions, the density of the lattice depending on the value of v in the region. The basic obstacle to obtaining an analog to Babenko's results here is that an asymptotically optimal design in the case $v(x) \equiv 1$ is not known. Still, a reasonable conjecture is that for a large class of isotropic processes cubature formulae based on equilateral triangular lattices are asymptotically optimal for the constant v case, which, combined with Babenko's (1976) approach, yields in turn a reasonable conjecture for asymptotically optimal procedures for v not constant.

To describe these cubature formulae, I will first consider the case where the design is locally a square lattice; results for other parallelogram lattices follow by a linear transformation of the coordinates. Let ϕ be a continuous positive function bounded away from 0 on G with $\int_G \phi(x)\,dx = 1$; this function will give the local density of observations. Define a sequence of designs indexed by n as follows. Let w_n be a sequence of positive reals satisfying $w_n \to 0$ and $w_n n^{1/2} \to \infty$ as $n \to \infty$. For design n, partition $\mathbb{R}^2$ into a lattice of squares of side w_n with corners at $w_n J$, $J \in \mathbb{Z}^2$. Define S_J to be the square with upper right corner at $w_n J$, $R_J = S_J \cap G$ and $r_J = \int_{R_J} dx$. For $r_J > 0$, place points in R_J by dividing S_J into an $m_J \times m_J$ lattice of squares, where

$$m_J = \left\lfloor w_n \left\{ \frac{r_J n}{\int_{R_J} \phi(x)\,dx} \right\}^{1/2} \right\rfloor. \tag{3.1}$$

Note that $m_J > 0$, for all n sufficiently large, since $w_n n^{1/2} \to \infty$. Let R_{JP} be the intersection of G with the square of side $w_n m_J^{-1}$ in R_J with upper right corner at $w_n\{J - (1,1) + m_J^{-1}P\}$, for $P \in \{1, \dots, m_J\}^2$. Furthermore, let c_{JP} be the center of mass of R_{JP} and r_{JP} its area. Then design n is the set of c_{JP}'s for which the corresponding r_{JP}'s are positive. See Figure 1 for an example of such a design. Consider the cubature formula

$$\hat{Z}_n = \sum_J \hat{Z}_n(J), \tag{3.2}$$

where

$$\hat{Z}_n(J) = \sum_P v(c_{JP}) r_{JP} Z(c_{JP}).$$

The number of points in design n, denoted by N_n, satisfies $n^{-1}N_n \to 1$. Note that some design points may not be in G, which could be undesirable in some circumstances. However, if G is convex, then all design points are always in G.

Next consider conditions on the random field Z. If v and the mean function of Z have bounded second partial derivatives on $\mathbb{R}^2$, then it is easy to show that the squared bias of $\hat{Z}_n$ is $O(n^{-2})$. Under the conditions of Theorem 1, the squared bias will then be asymptotically negligible relative to the variance. Thus, from now on I will only consider the variance of the cubature error and ignore the bias. Assume Z is isotropic with spectral density f of the form $f(\omega) = g(|\omega|)$. The covariance function for Z is $C(|x|)$, where C is given by the Bessel transform

$$C(t) = 2\pi \int_0^\infty J_0(t\nu)\,\nu g(\nu)\,d\nu.$$

The critical condition needed here to obtain the asymptotic variance of the cubature error is

$$g \text{ is regularly varying at } \infty \text{ with exponent } -p,\ 2 < p < 4. \tag{3.3}$$

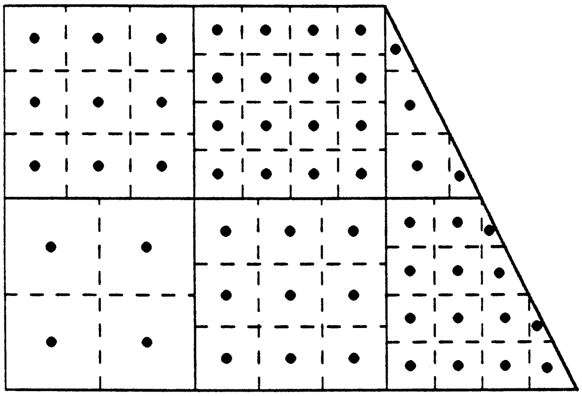

FIG. 1. *Example of locally square lattice design with G a trapezoid with vertices at* $[0,0]$, $[1.5,0]$, $[0,1]$ *and* $[1,1]$, $w_n = \frac{1}{2}$ *and* $m_{(1,1)} = 2$, $m_{(2,1)} = 3$, $m_{(3,1)} = 4$, $m_{(1,2)} = 3$, $m_{(2,2)} = 4$ *and* $m_{(3,2)} = 3$.

It follows [see Theorem 1 of Bingham (1972), where his $\alpha = p - 2$] that

$$C(0) - C(t) = O\big(t^{-2} g(t^{-1})\big) \quad \text{as } t \downarrow 0. \tag{3.4}$$

THEOREM 1. *For G a compact Jordan-measurable subset of $\mathbb{R}^2$, ϕ continuous and bounded away from 0 on G and $\int_G \phi(x)\,dx = 1$, v having bounded partial derivatives through order 3 on $\mathbb{R}^2$ and Z a stationary random field on $\mathbb{R}^2$ with covariance structure satisfying* (3.3),

$$\begin{aligned} g\big(N_n^{1/2}\big)^{-1} \operatorname{var}\bigg(\int_G v(x) Z(x)\,dx - \hat{Z}_n\bigg) \\ \to (2\pi)^{2-p} \int_G v(x)^2 \phi(x)^{-p}\,dx \sum_K{}' |K|^{-p}, \end{aligned}$$

where Σ' means to sum over all elements of $\mathbb{Z}^2$ except the origin.

PROOF. The following result, proven in Appendix A, is helpful:

LEMMA 1. *For g regularly varying with exponent $-p$, $2 < p < 4$, there exists a function $\tilde{g}$ that is ultimately monotone with $\tilde{g}(t)/g(t) \to 1$ as $t \to \infty$ and that, for any given positive integer q and finite T, there exists a constant β with*

$$|\tilde{C}^{(q)}(t)| \le \beta t^{-2-q} \tilde{g}(t^{-1}) \quad \textit{for } 0 < t \le T, \tag{3.5}$$

where $\tilde{C}(t) = 2\pi \int_0^\infty J_0(t\nu)\nu \tilde{g}(\nu)\,d\nu$.

For now, assume we can take $\tilde{g} = g$ in (3.5). Define

$$\operatorname{cov}_{JK} = \operatorname{cov}\left(\int_{R_J} v(x)Z(x)\,dx - \hat{Z}_n(J), \int_{R_K} v(x)Z(x)\,dx - \hat{Z}_n(K)\right).$$

The basic outline of the proof is to show

$$g(n^{1/2}) \sum_{J \neq K} \operatorname{cov}_{JK} \to 0 \quad \text{as } n \to \infty \tag{3.6}$$

and then to handle $\sum \operatorname{var}_J$ where $\operatorname{var}_J = \operatorname{cov}_{JJ}$, using methods in Stein (1995a). The following result will be helpful: for any two Jordan-measurable subsets A and B of G with centers of mass a and b, $\varepsilon = \max(\operatorname{diam}(A), \operatorname{diam}(B))$ and $\delta = d(A, B) = \inf\{|x - y|: x \in A, y \in B\}$ as defined, there exists a constant γ independent of A and B such that

$$\begin{aligned} &\left|\operatorname{cov}\left(\int_A \{v(x)Z(x) - v(a)Z(a)\}\,dx,\right.\right.\\ &\qquad\left.\left.\int_B \{v(x)Z(x) - v(b)Z(b)\}\,dx\right)\right| \\ &\quad\leq \gamma\varepsilon^8\{\max(\varepsilon, \delta)\}^{-6} g\left(\{\max(\varepsilon, \delta)\}^{-1}\right). \end{aligned} \tag{3.7}$$

Appendix B outlines a proof of (3.7). To establish (3.6), first consider $|J - K| \geq 2$. Then for all possible P and Q, $d(R_{JP}, R_{KQ}) \geq 5^{-1/2} w_n |J - K|$, where the lower bound can be achieved if, for example, $J - K = (2, 1)$. Since $\operatorname{diam}(R_{JP}) = O(n^{-1/2})$ uniformly in J and P, using (3.7) and $m_J = O(n^{1/2} w_n)$ uniformly in J,

$$\begin{aligned} \operatorname{cov}_{JK} &= \sum_{P:\, r_{JP} > 0} \sum_{Q:\, r_{KQ} > 0} \operatorname{cov}\left(\int_{R_{JP}} \{v(x)Z(x) - v(c_{JP})Z(c_{JP})\}\,dx,\right.\\ &\qquad\qquad\left.\int_{R_{KQ}} \{v(x)Z(x) - v(c_{KQ})Z(c_{KQ})\}\,dx\right) \\ &\ll n^2 w_n^4 (n^{1/2})^{-8} \{w_n |J - K|\}^{-6} g\left(\{w_n |J - K|\}^{-1}\right), \end{aligned}$$

where $\alpha \ll \beta$ means $|\alpha| \leq c\beta$ for some constant c independent of n and any other index such as J or K that can implicitly depend on n. Thus,

$$\begin{aligned} \sum_{|J-K| \geq 2} \operatorname{cov}_{JK} &\ll n^{-2} w_n^{-2} \sum_{|J-K| \geq 2} |J - K|^{-6} g\left(\{w_n |J - K|\}^{-1}\right) \\ &\ll n^{-2} w_n^{-4} \sum_{0 < |J| < w_n^{-1}} |J|^{-6} g\left(\{w_n |J|\}^{-1}\right) \end{aligned}$$

$$\begin{aligned}
(3.8) \qquad & \ll n^{-2} w_n^{-4} \int_1^{w_n^{-1}} t^{-5} g\big((w_n t)^{-1}\big)\, dt \\
&= n^{-2} \int_1^{w_n^{-1}} s^3 g(s)\, ds \\
&\ll n^{-2} w_n^{-4} g\big(w_n^{-1}\big) \\
&= o\big(g(n^{1/2})\big)
\end{aligned}$$

using Karamata's theorem [Bingham, Goldie and Teugels (1987), page 28] to obtain the penultimate step and $p < 4$ and $n^{1/2} w_n \to \infty$ for the last step. Next, consider those cases for which $0 < |J - K| < 2$, such as $J - K = (1,0)$ or $(1,1)$. There are $O(w_n^{-2})$ such cases. Suppose $J - K = (1,0)$; other cases are similar. Then using (3.7),

$$\begin{aligned}
\mathrm{cov}_{JK} &\ll n^{-4} \sum_{j,k=1}^{m_J} \sum_{l,m=1}^{m_K} \left(\frac{(j-l)^2 + (k+m)^2}{n} \right)^{-3} \\
&\qquad \times g\left(\frac{n^{1/2}}{\{(j-l)^2 + (k+m)^2\}^{1/2}} \right) \\
&\ll n^{-1} \int_{[1, w_n n^{1/2}]^4} \{(s-t)^2 + (u+v)^2\}^{-3} \\
&\qquad \times g\left(\frac{n^{1/2}}{\{(s-t)^2 + (u+v)^2\}^{1/2}} \right) ds\, dt\, du\, dv \\
&\ll n^{-1/2} w_n \int_{[1, w_n n^{1/2}]^3} \{s^2 + (u+v)^2\}^{-3} \\
&\qquad \times g\left(\frac{n^{1/2}}{\{s^2 + (u+v)^2\}^{1/2}} \right) ds\, du\, dv \\
&\ll n^{-1/2} w_n \int_1^{w_n n^{1/2}} \int_s^{w_n n^{1/2}} r^{-2} g\left(\frac{n^{1/2}}{r} \right) dr\, ds \\
&\ll n^{-1/2} w_n \int_1^{w_n n^{1/2}} s^{-1} g\left(\frac{n^{1/2}}{s} \right) ds \\
&\ll n^{-1/2} w_n g(n^{1/2}),
\end{aligned}$$

so that

$$(3.9) \qquad \sum_{0 < |J-K| < 2} \mathrm{cov}_{JK} \ll n^{-1/2} w_n^{-1} g(n^{1/2}) = o\big(g(n^{1/2})\big).$$

Then (3.6) follows from (3.8) and (3.9).

Now consider var_J. By an argument similar to the one leading to (3.9),

$$\begin{aligned}\mathrm{var}_J &\ll n^{-1}\int_{[1,\,w_n n^{1/2}]^4}\left\{(s-t)^2+(u-v)^2\right\}^{-6}\\ &\qquad\times g\left(\frac{n^{1/2}}{\left\{(s-t)^2+(u-v)^2\right\}^{1/2}}\right)ds\,dt\,du\,dv\\ &\ll w_n^2\int_1^{w_n n^{1/2}} r^{-5} g\left(\frac{n^{1/2}}{r}\right)dr\\ &\ll w_n^2 g(n^{1/2}).\end{aligned}$$

Since G is Jordan measurable, the number of S_J's partially in G is $o(w_n^{-2})$, so that

$$\sum_{R_J\neq S_J}\mathrm{var}_J = o\big(g(n^{1/2})\big). \tag{3.10}$$

The terms for which $S_J\subset G$ can be handled by extending results in Stein (1993, 1995a) to yield (see Appendix C)

$$\begin{aligned}\mathrm{var}_J &= (2\pi)^{2-p}\int_{S_J} v(x)^2\,dx\left(\frac{1}{r_J}\int_{R_J}\phi(x)\,dx\right)^p\\ &\qquad\times\sum_K{}'|K|^{-p}g(n^{1/2})+o\big(w_n^2 g(n^{1/2})\big),\end{aligned}\tag{3.11}$$

where the remainder is small uniformly in J. Then (3.10), (3.11), $w_n\to 0$ and the conditions on ϕ imply

$$g(n^{1/2})^{-1}\sum\mathrm{var}_J\to(2\pi)^{2-p}\int_G v(x)^2\phi(x)^{-p}\,dx\sum_K{}'|K|^{-p}$$

as $n\to\infty$, which in conjunction with (3.6) and $n^{-1}N_n\to 1$ yields the theorem when we can take $\tilde g = g$ in (3.5).

The proof for more general g follows using an idea of Pitt, Robeva and Wang (1995). Specifically, given $\varepsilon>0$, we can choose T such that $\tilde g(t)(1-\varepsilon)\le g(t)\le \tilde g(t)(1+\varepsilon)$, for $t\ge T$. Define $g_T(t)$ to be $g(t)$ on $0<t<T$ and 0 otherwise. Define $\tilde g_T$ similarly. Then $\tilde g(t)(1-\varepsilon)-\tilde g_T(t)\le g(t)\le(1+\varepsilon)\tilde g(t)+g_T(t)$, for all t. Using a subscript on a variance to indicate the isotropic spectral density under which the variance is computed and setting $e_n=\int_G v(x)Z(x)\,dx-\hat Z_n$,

$$(1-\varepsilon)\mathrm{var}_{\tilde g}(e_n)-\mathrm{var}_{g_T}(e_n)\le\mathrm{var}_g(e_n)\le(1+\varepsilon)\mathrm{var}_{\tilde g}(e_n)+\mathrm{var}_{g_T}(e_n).$$

Because g_T has bounded support, its Bessel transform is analytic on $[0,\infty)$

and it follows by a straightforward argument using Taylor series that

$$\operatorname{var}_{g_T}\left(\int_{R_{JP}} \{v(x)Z(x) - v(c_{JP})Z(c_{JP})\}\,dx\right) = O(n^{-4}),$$

which implies $\operatorname{var}_{g_T}(e_n) = O(n^{-2})$, which is $o(g(n^{1/2}))$ since $p < 4$. Similarly, $\operatorname{var}_{\tilde{g}_T}(e_n) = O(n^{-2})$. Thus,

$$1 - \varepsilon \le \liminf_{n\to\infty} \frac{\operatorname{var}_g(e_n)}{\operatorname{var}_{\tilde{g}}(e_n)} \le \limsup_{n\to\infty} \frac{\operatorname{var}_g(e_n)}{\operatorname{var}_{\tilde{g}}(e_n)} \le 1 + \varepsilon.$$

Since ε is arbitrary Theorem 1 holds for all g satisfying (3.3). □

We can obtain a similar result for nonsquare lattices by linear transformation. For a 2×2 matrix B with determinant 1, let the rows of B determine the lattice structure. Specifically, let the corners of a lattice of parallelograms be given by $w_n BJ$, $J \in \mathbb{Z}^2$. Call S_J the parallelogram with upper right corner at $w_n BJ$ and set $R_J = S_J \cap G$ and $r_J = \int_{R_J} dx$ as before. For $r_J > 0$, divide S_J into an $m_J \times m_J$ lattice of parallelograms, where m_J is defined as in (3.1). Place observations at the centers of mass of the nonempty intersections of these parallelograms with G and define $\hat{Z}_n$ as in (3.2). Under the same conditions on g, G, v, ϕ and w_n as in Theorem 1, a straightforward extension of that result gives

$$g\left(N_n^{1/2}\right)^{-1} \operatorname{var}(e_n) \to (2\pi)^{2-p} \int_G v(x)^2 \phi(x)^{-p}\,dx \sum_K{}' |(B')^{-1}K|^{-p}. \tag{3.12}$$

4. Possibly asymptotically optimal cubature formulae. The three elements of the cubature formulae described in Section 3 that we need to choose are w_n, ϕ and B. Equation (3.12) provides no guidance on choosing w_n; any sequence satisfying $w_n \to 0$ and $w_n n^{1/2} \to \infty$ yields the same asymptotic result. Regarding ϕ and B, note that (3.12) factors into two terms: one depending only on ϕ; the other only on B. Thus, to minimize the right-hand side of (3.12), it suffices to minimize each of these two terms separately. The minimizer of $\int_G v(x)^2 \phi(x)^{-p}\,dx$ subject to $\int_G \phi(x)\,dx = 1$ is given by Hölder's inequality:

$$\phi_0(x) = |v(x)|^{2/(p+1)} \Big/ \int_G |v(y)|^{2/(p+1)}\,dy. \tag{4.1}$$

If we want Theorem 1 to hold for ϕ_0, we need to require that v is bounded away from 0 on G. Now consider $\Sigma' |(B')^{-1}K|^{-p}$. The function

$$Z_A(s) = \sum{}' |J'AJ|^{-s},$$

where A is a symmetric nonsingular 2×2 matrix, is known as the Epstein zeta-function [Rankin (1953)]. This sum converges for all $s > 1$. For any $s > 1$, the global minimum of $Z_A(s)$ with respect to symmetric A with determinant 1 is achieved by

$$A = \frac{2}{3^{1/2}}\begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix},$$

which was proven for $s > 1.035$ by Rankin (1953) and by Cassels (1959), Ennola (1964a) and Diananda (1964) for $1 < s < 1.035$. It follows that for $p > 2$, to minimize $\Sigma'|(B')^{-1}K|^{-p}$ with respect to B with determinant 1, we can use an equilateral triangular lattice

$$B_0 = \frac{2^{1/2}}{3^{1/4}}\begin{pmatrix} 1/(2\cdot 3^{1/2}) & -1/2 \\ 0 & 1 \end{pmatrix}, \tag{4.2}$$

or any orthogonal rotation of this. Sobolev (1974, 1992) was the first to recognize the relevance of the Epstein zeta-function to the error analysis of cubature formulae based on lattice designs, although he did not consider lattices of varying density. I conjecture that the sequence of cubature rules $\hat{Z}_n$ defined as in Section 3 with ϕ_0 given by (4.1) and B_0 given by (4.2) is asymptotically optimal in the following sense. Suppose $\{\tilde{Z}_n\}$ is a sequence of cubature rules where rule n is based on $\tilde{N}_n$ observations with $\tilde{N}_n \le N_n$, for all n sufficiently large. Furthermore, suppose v, g and G satisfy the conditions of Theorem 1 and v is bounded away from 0. Then I conjecture

$$\liminf_{n\to\infty} \frac{\operatorname{var}\left(\int_G v(x)Z(x)\,dx - \tilde{Z}_n\right)}{\operatorname{var}\left(\int_G v(x)Z(x)\,dx - \hat{Z}_n\right)} \ge 1. \tag{4.3}$$

What I have proven is that among all rules of the form described in Section 3, ϕ_0 and B_0 yield an asymptotically optimal sequence of rules. Furthermore, considering Proposition 4.2 of Stein (1995a), I would expect the coefficients for this sequence of rules to be asymptotically relatively optimal, although I cannot prove this.

The fact that the asymptotic variance factors into a term depending on the density and another on the lattice makes it easy to study each term's impact. From Theorem 1 and (4.1) we have that the ratio of the asymptotic variances using constant ϕ to the optimal ϕ is

$$\frac{\int_G v(x)^2\,dx}{\left\{\int_G |v(x)|^{2/(p+1)}\,dx\right\}^{p+1}}.$$

Thus, the improvement in using the optimal ϕ increases with, roughly speaking, increasing variation in v and decreases with increasing p. Similar results hold for one-dimensional integrals [Pitt, Robeva and Wang (1995) and Stein (1995b)]. The asymptotic improvement using the optimal triangular lattice rather than the square lattice is quite small. For example, if $p = 3$, which holds if $C(t) = e^{-|t|}$, then the reduction in the asymptotic variance due to using a triangular lattice rather than a square lattice is less than 2%.

Thus, the advantage to using triangular rather than square lattices is mainly of theoretical interest.

As noted in Section 3, the main obstacle to proving the asymptotic optimality of $\hat{Z}_n$ is that in the case $v(x) \equiv 1$, it is only known that the equilateral triangular lattice is asymptotically optimal with respect to parallelogram lattice designs and may not be asymptotically optimal with respect to all possible designs. Babenko (1977) showed that if the class of functions being considered is those h satisfying $|h(x) - h(y)| \leq \psi(|x - y|)$, where ψ is a modulus of continuity, then a sequence of rules based on equilateral triangular lattices is asymptotically optimal with respect to all possible rules. However, this result does not appear to be of any help here. The basic problem is that bounds on cubature errors for classes of functions satisfying a modulus of continuity constraint can be obtained by bounding the difference between $h(x)$ and h at its nearest observed location to x, which roughly speaking, eliminates the difficult problem of computing covariances between interpolation errors at different points.

5. Discussion. Even if the conjecture given in (4.3) is true, there is still something unsatisfying about the sequence of proposed quadrature rules. First, there is the issue of how to choose w_n. However, even if a sharper analysis yielded information on this question, there is still a problem. What we would like to do is to vary the density of the lattice continuously, without having to make sharp breaks in the lattice structure as I have done at the boundaries of the S_J's. Indeed, in one dimension with $G = [0, 1]$, this is easily accomplished using regular sequences of designs [Sacks and Ylvisaker (1971)]: to obtain the design of size n, place points at $F(i/n)$, for $i = 1, \dots, n$, where F is a smooth cumulative distribution function with $F(0) = 0$ and $F(1) = 1$.

It is possible in limited circumstances to do something similar in two dimensions. Think of G as a set of points in the complex plane and suppose Ψ is a conformal mapping from some set B onto G. If we place an equilateral triangular lattice on B and map it into G using Ψ, the resulting set of points will look locally like an equilateral triangular lattice. Since every smooth conformal mapping (or its conjugate) is holomorphic in G [Hille (1959), page 95], we can restrict attention to holomorphic Ψ. As the density of the lattice on B increases, the density of points in a neighborhood of z in G becomes proportional to $|\eta'(z)|^2$, where $\eta = \Psi^{-1}$, which is well defined and holomorphic on G. Thus, for a given real-valued positive function $\phi(s, t)$, where s and t are real, we want to find holomorphic η such that $|\eta'(z)|^2 = \phi(\mathrm{Re}(z), \mathrm{Im}(z))$. This will be possible if and only if $\mu = \log \phi$ is harmonic on G, where the only if part is noted by Titchmarsh [(1939), page 120] and the if part follows by defining $\eta'(x - iy) = \exp(\mu(x, y) + i\lambda(x, y))$, where λ is the harmonic function conjugate to μ [Carrier, Krook and Pearson (1966), page 45]. For example, for G the unit square and $\phi(s, t) = e^s/(e - 1)$, we can take $\Psi(z) = \log((e - 1)z^2/4)$, where log takes its principal value. Figure 2 shows an example of such a design where the distance between neighbors in the

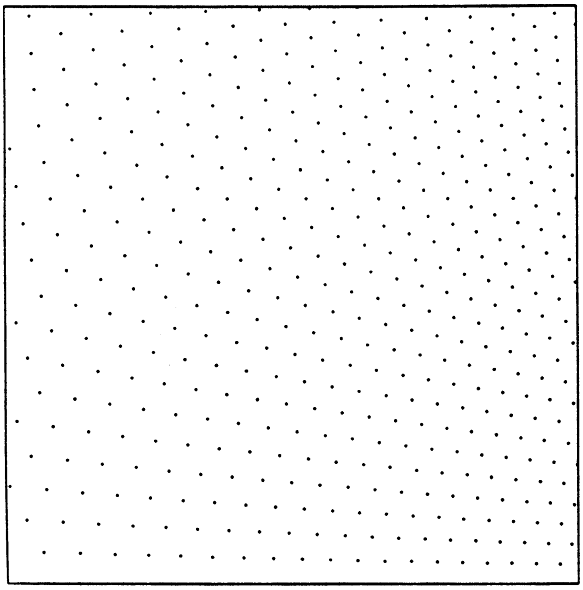

FIG. 2. *Locally equilateral triangular lattice on the unit square with local density proportional to* $\phi(s,t) = e^s/(e-1)$.

equilateral triangular lattice in B, the inverse image of G, is 0.05. To use this design as a basis of a good cubature rule would require adjusting the locations or the coefficients of the points near the boundary. Since the class of harmonic functions is quite limited, this method will not be broadly applicable. Furthermore, in three or higher dimensions, the class of conformal mappings is extremely limited [Carman (1974), page 20], so that unlike the designs defined in Section 3, extensions to higher dimensions are not in general possible.

An important limitation of Theorem 1 is the restriction to two dimensions and to $p < 4$. The results in Stein [(1995a), Section 4] suggest that Theorem 1 can be extended to $p \geq 4$ if the coefficients used in the cubature rule are modified, although how to do this explicitly for irregular G is not so obvious. As far as extensions to three or more dimensions, Theorem 1 does easily extend to three dimensions if we make the restriction $w_n \to 0$ and $w_n n^{1/3} \to \infty$. However, we still need $p < 4$, but since we must also have $p > 3$ to have a density, the class of models covered by such a result is not very interesting. If we go to four or higher dimensions, then requiring $p < 4$ leaves us with nothing, as p must be greater than the number of dimensions. Thus, to obtain a meaningful extensions to more than two dimensions, we need to be able to weaken the restriction on p.

There is also an obstacle to extending the optimality result in Section 4 on the Epstein zeta-function to more than two dimensions. Specifically, while the Epstein zeta-function extends in the obvious way to higher dimensions, the minimizing lattice is not known. Even in three dimensions, while Ennola (1964b) shows that the well-known body-centered cubic lattice gives a local minimum for all s, whether it gives a global minimum is not known. Shushbaev (1989) studied the Epstein zeta-function in higher dimensions and shows there can be more than one local minimum, but gives no results for global minima.

Finally, it is important to note that parallelogram lattice designs will not work well for all stationary random fields. For example, Ylvisaker (1975) showed that if $x = (x_1, x_2)$ and $\operatorname{cov}(Z(0), Z(x)) = C_1(x_1)C_2(x_2)$, then lattice designs can be highly inefficient relative to tensor product designs. Delvos (1989) considered related designs from a deterministic perspective. Quasirandom numbers and good lattice points [Niederreiter (1992)] can also be much more efficient than parallelogram lattice designs for certain classes of functions. However, the classes of functions considered in all of these works share the common feature that their behavior depends crucially on the choice of coordinate axes, a fact that perhaps has not been adequately appreciated. For classes of functions that are isotropic in some sense, or whose anisotropy can be removed by a linear transformation of coordinates, I suspect it will be hard to improve on the locally lattice designs studied in Section 3. For spatial data, where the choice of coordinate axes is often somewhat or completely arbitrary, models whose behavior depends strongly on such a choice, as in Ylvisaker (1975), Delvos (1989) or Levin and Girshovich [(1979), Chapter 4], are difficult to justify.

APPENDIX A

PROOF OF LEMMA 1. If $g(t)t^p \to \alpha$ as $t \to \infty$, we can take $\tilde{g}(t) = \alpha(1 + t^2)^{-p/2}$ and obtain the result. The more general case can be handled using smoothly varying functions [Bingham, Goldie and Teugels (1987), Section 1.8]. A positive function f defined on some neighborhood of infinity varies smoothly with index α if $h(x) = \log f(e^x)$ is C^∞ and as $x \to \infty$, $h'(x) \to \alpha$, $h^{(n)}(x) \to 0$, for $n \geq 2$. We will write $f \in SR_\alpha$ for such a function. By the smooth variation theorem [Bingham, Goldie and Teugels (1987), Theorem 1.8.2], there exists $\tilde{g} \in SR_{-p}$ such that $\tilde{g}(t)/g(t) \to 1$ as $t \to \infty$ and $\tilde{g}(t)$ is monotone for t sufficiently large [see the remark before Theorem 1.8.3 of Bingham, Goldie and Teugels (1987)]. So, it suffices to show $\tilde{g}$ satisfies (3.5).

Let $r(\nu) = 2\pi\nu\tilde{g}(\nu)$. Then it is trivial to show $r \in SR_{-p+1}$. Formally, we have $C'(t) = \int_0^\infty \nu J_1(t\nu)r(\nu)\,d\nu$. To justify differentiating inside the integral, let $f_n(t) = \int_0^n J_0(t\nu)r(\nu)\,d\nu$, so that $f_n(t) \to C(t)$ and $f_n'(t) = \int_0^n J_1(t\nu)\nu r(\nu)\,d\nu$, for all $t \geq 0$. For any given $0 < a < b < \infty$, it suffices to show that $f_n'(t)$ converges uniformly to $\int_0^\infty J_1(t\nu)\nu r(\nu)\,d\nu$ on $[a,b]$. This is obvious for $p > 5/2$, but we need to be more careful for $2 < p \le 5/2$, since $\int_0^\infty J_1(t\nu)\nu r(\nu)\,d\nu$ does not converge absolutely. However, for $n > m$, using (9.2.1) of Abramowitz and Stegun (1964),

$$\begin{aligned}
|f_n'(t) - f_m'(t)| &= \left|\int_m^n J_1(t\nu)\nu r(\nu)\,d\nu\right| \\
&= \left|\int_m^n \left(\frac{2}{\pi t\nu}\right)^{1/2}\left\{\cos\left(t\nu - \frac{3\pi}{4}\right) + O\left(\frac{1}{\nu}\right)\right\}\nu r(\nu)\,d\nu\right| \\
&= \left(\frac{2}{\pi t}\right)^{1/2}\left|\left\{\frac{1}{t}\sin\left(t\nu - \frac{3\pi}{4}\right)\nu^{1/2}r(\nu)\Big|_m^n\right\}\right. \\
&\qquad\left. - \frac{1}{t}\int_m^n \sin\left(t\nu - \frac{3\pi}{4}\right)\left\{\frac{1}{2}\nu^{-1/2}r(\nu) + r'(\nu)\right\}d\nu\right| \\
&\quad + O\left(\frac{1}{t^{1/2}}\int_m^n r(\nu)\,d\nu\right),
\end{aligned}$$

which tends to 0 uniformly for $t \in [a,b]$ as $m,n \to \infty$, since $|r'(\nu)| \in SR_{-p}$ by Proposition 1.8.1 of Bingham, Goldie and Teugels (1987).

Next, integrating by parts,

$$\begin{aligned}
C'(t) &= \int_0^\infty J_1(t\nu)\nu r(\nu)\,d\nu \\
&= \frac{1}{t}\int_0^\infty J_0(t\nu)\{r(\nu) + \nu r'(\nu)\}\,d\nu.
\end{aligned}$$

Since $r(\nu)$ and $\nu r'(\nu)$ are in SR_{-p+1} and hence regularly varying with exponent $-p+1$, (3.5) holds for $q = 1$ by applying (3.4) to $tC'(t)$. Now we can essentially repeat the argument to get the result for larger q. For example,

$$\begin{aligned}
C''(t) &= -\frac{1}{t^2}\int_0^\infty J_0(t\nu)\{r(\nu) + \nu r'(\nu)\}\,d\nu \\
&\quad + \frac{1}{t}\int_0^\infty J(t\nu)\nu\{r(\nu) + \nu r'(\nu)\}\,d\nu \\
&= -\frac{1}{t^2}\int_0^\infty J_0(t\nu)\{r(\nu) + \nu r'(\nu)\}\,d\nu \\
&\quad + \frac{1}{t^2}\int_0^\infty J_0(t\nu)\{r(\nu) + 3\nu r'(\nu) + \nu^2 r''(\nu)\}\,d\nu,
\end{aligned}$$

so that we can again apply (3.4) to $t^2 C''(t)$ since $r(\nu)$, $\nu r'(\nu)$ and $\nu^2 r''(\nu)$ are all in SR_{-p+1}. $\square$

APPENDIX B

PROOF OF (3.7). Define $D(x) = C(|x|)$ and use subscripts to indicate partial differentiation, so that, for example,

$$D_{j_1 j_2 j_3}(x) = \frac{\partial^3}{\partial x_{j_1}\,\partial x_{j_2}\,\partial x_{j_3}} D(x),$$

where $x = (x_1, x_2)$. It follows from (3.5) that

$$D_{j_1 \cdots j_\alpha}(x) \ll |x|^{-(2+\alpha)} g(|x|), \tag{B.1}$$

for all α and all $0 < |x| \le T$. For now, assume $\delta > \varepsilon$. Then

$$\begin{aligned}
&\operatorname{cov}\left(\int_A \{v(x)Z(x) - v(a)Z(a)\}\,dx, \int_B \{v(x)Z(x) - v(b)Z(b)\}\,dx\right)\\
&\quad = \int_{A\times B} \{v(x)v(y)D(x-y) - v(x)v(b)D(x-b)\\
&\qquad\qquad - v(a)v(y)D(a-y) + v(a)v(b)D(a-b)\}\,dx\,dy \qquad \text{(B.2)}\\
&\quad = D(a-b)\int_A \{v(x) - v(a)\}\,dx \int_B \{v(x) - v(b)\}\,dx\\
&\qquad + \sum_{\alpha=1}^{3} \sum_{j_1,\dots,j_\alpha=1}^{2} D_{j_1\cdots j_\alpha}(a-b) F_\alpha(j_1,\dots,j_\alpha) + O\big(\varepsilon^8 \delta^{-6} g(\delta^{-1})\big),
\end{aligned}$$

where

$$\begin{aligned}
F_\alpha(j_1,\dots,j_\alpha) = \int_{A\times B} \Bigg\{ & v(x)v(y) \prod_{l=1}^{\alpha} (x_{j_l} - y_{j_l} - a_{j_l} + b_{j_l})\\
& - v(x)v(b) \prod_{l=1}^{\alpha} (x_{j_l} - a_{j_l}) - v(a)v(y) \prod_{l=1}^{\alpha} (b_{j_l} - y_{j_l}) \Bigg\}\,dx\,dy.
\end{aligned}$$

Now, since a and b are centers of mass of A and B,

$$\int_A \{v(x) - v(a)\}\,dx \int_B \{v(x) - v(b)\}\,dx = O(\varepsilon^8). \tag{B.3}$$

Furthermore, for $\delta > \varepsilon$,

$$F_\alpha(j_1,\dots,j_\alpha) = O(\varepsilon^8), \tag{B.4}$$

for $1 \le \alpha \le 3$. For example, using the convention that any index appearing twice in a product is summed over,

$$\begin{aligned}
F_1(j) = \int_{A\times B} \Big[\{ & v(a)v(b) + v(a)v_k(b)(b_k - y_k) + v_k(a)v(b)(x_k - a_k)\\
& + v_k(a)v_l(b)(x_k - a_k)(b_k - y_k)\\
& + \tfrac{1}{2} v_{kl}(a)v(b)(x_k - a_k)(x_l - a_l)
\end{aligned}$$

$$
\begin{aligned}
&\quad + \tfrac{1}{2} v(a) v_{kl}(b)(b_k - y_k)(b_l - y_l)\big\}(x_j - y_j - a_j + b_j) \\
&\quad - v(a)\big\{v(b) + v_k(b)(b_k - y_k) \\
&\qquad\qquad + \tfrac{1}{2} v_{kl}(b)(b_k - y_k)(b_l - y_l)\big\}(b_j - y_j) \\
&\quad - v(b)\big\{v(a) + v_k(a)(x_k - a_k) \\
&\qquad\qquad + \tfrac{1}{2} v_{kl}(a)(x_k - a_k)(x_l - y_l)\big\}(x_j - a_j) + O(\varepsilon^4)\Big]\, dy\, dx \\
&= \int_{A\times B} \big\{v(a) v_k(b)(b_k - y_k)(x_j - a_j) \\
&\qquad + v_k(a) v(b)(x_k - a_k)(b_j - y_j) \\
&\qquad + v_k(a) v_l(b)(x_k - a_k)(b_l - y_l)(x_j - y_j - a_j + b_j) \\
&\qquad + \tfrac{1}{2} v_{kl}(a) v(b)(x_k - a_k)(x_l - a_l)(b_j - y_j) \\
&\qquad + \tfrac{1}{2} v(a) v_{kl}(b)(b_k - y_k)(b_l - y_l)(x_j - a_j)\big\}\, dy\, dx \\
&\quad + O(\varepsilon^8) \\
&= O(\varepsilon^8),
\end{aligned}
$$

again using that a and b are centers of mass. For $\delta > \varepsilon$, (3.7) follows from (B.1)–(B.4). For $\delta \le \varepsilon$, (3.7) follows directly from (3.4) using

$$
\begin{aligned}
&\bigg|\operatorname{cov}\bigg(\int_A \{v(x)Z(x) - v(a)Z(a)\}\, dx, \int_B \{v(x)Z(x) - v(b)Z(b)\}\, dx\bigg) \\
&\qquad\qquad - D(0)\int_A \{v(x) - v(a)\}\, dx \int_B \{v(x) - v(b)\}\, dx\bigg| \\
&\quad \ll \sup_{(x,y)\in A\times B} |D(x-y) - D(0)| \int_{A\times B} dy\, dx \\
&\quad \ll \varepsilon^{-2} g(\varepsilon^{-1}) \varepsilon^4,
\end{aligned}
$$

since the exponent of regular variation $-p$ satisfies $p > 2$. $\square$

APPENDIX C

PROOF OF (3.11). Consider $S_J \subset G$. Define $\alpha_J = w_n^{-1} m_J$, $\rho_J = (-\pi\alpha_J, \pi\alpha_J)^2$ and, for a function v,

$$
\begin{aligned}
&\mathscr{F}_J(\omega; v) = \int_{S_J} v(x) e^{i\omega' x}\, dx, \\
\text{(C.1)}\qquad &\mathscr{V}_J(\omega; v) = \bigg|\mathscr{F}_J(\omega; v) - \alpha_J^{-2} \sum_P v(c_{JP}) \exp(i\omega' c_{JP})\bigg|^2,
\end{aligned}
$$

so that $\mathrm{var}_J = \int f(\omega) \mathscr{V}_J(\omega; v)\, d\omega$. Define $u(x) = v(x) - v(w_n J)$, suppressing the dependence of u on J. Then

$$
\text{(C.2)}\qquad \mathscr{V}_J(\omega; v) \le 2 v(w_n J)^2 \mathscr{V}_J(\omega; 1) + 2\mathscr{V}_J(\omega; u),
$$

where $\mathscr{V}_J(\omega;1)$ means to take $v(x) \equiv 1$ in (C.1). By direct calculation,

$$\mathscr{V}_J(\omega;1) \ll \left|\frac{w_n\omega}{\alpha_J}\right|^4 \prod_{j=1}^{2}\left\{1+(w_n\omega_j)^2\right\}^{-1} \tag{C.3}$$

on ρ_J. A slight sharpening of Lemma 3.1 of Stein (1995a) provides a bound for the second term on the right-hand side of (C.2). Specifically, by inspecting the proof of this lemma we have the following: if a function v on $\mathbb{R}^d$ has partial derivatives through order $d+1$ on $[0,1]^d$ and D is a bound on the absolute value of v and these partial derivatives on $[0,1]^d$, then there exists a constant C independent of v and m such that

$$\begin{aligned}&\left|\int_{[0,1]^d} v(x)\exp(i\omega' x)\,dx - m^{-d}\sum_{\{1,\dots,m\}^d} v\left(\frac{J-h}{m}\right)\exp\left(\frac{i\omega'(J-h)}{m}\right)\right|\\&\quad\le \frac{CD}{m^2}\left\{1+\frac{(\log m)^{2d}(1+|\omega|^2)}{\prod_{j=1}^{d}(1+|\omega_j|)}\right\},\end{aligned}$$

for $\omega \in (-m\pi, m\pi)^d$, where $h = (\frac{1}{2},\dots,\frac{1}{2})'$. By a linear transformation of coordinates, we get

$$\mathscr{V}_J(\omega;u)^{1/2} \le \frac{2w_nCD}{\alpha_J^2}\left\{1+\frac{(\log m_J)^4(1+|w_n\omega|^2)}{\prod_{j=1}^{2}(1+|w_n\omega_j|)}\right\} \tag{C.4}$$

on ρ_J. Then (C.2)–(C.4) imply

$$\mathscr{V}_J(\omega;v) \ll \frac{w_n^2}{\alpha_J^4} + \frac{(\log m_J)^8(1+|w_n\omega|^4)}{\alpha_J^4\prod_{j=1}^{2}(1+|w_n\omega_j|^2)} \tag{C.5}$$

on ρ_J, from which it follows

$$\begin{aligned}\int_{\rho_J} f(\omega)\mathscr{V}_J(\omega;v)\,d\omega &\ll \frac{w_n^2}{\alpha_J^4} + \frac{(\log m_J)^8}{\alpha_J^4}\left\{g(w_n^{-1}) + m_J^3 g(\alpha_J)\right\}\\&= o\left(w_n^2 g(n^{1/2})\right),\end{aligned} \tag{C.6}$$

using g regularly varying at ∞ with exponent $-p$, $2<p<4$ and $\alpha_J n^{-1/2}$ uniformly bounded away from 0, for all n sufficiently large. Next, by Parseval's relation,

$$\begin{aligned}\int_{\rho_J^c} |\mathscr{F}(\omega;v)|^2\,d\omega &\ll \int_{\rho_J^c} |\mathscr{F}_J(\omega;1)|^2\,d\omega + \int_{\mathbb{R}^d} |\mathscr{F}_J(\omega;u)|^2\,d\omega\\&\ll \frac{w_n}{\alpha_J} + w_n^4 + \int_{S_J} u(x)^2\,dx\\&= o(w_n^2)\end{aligned} \tag{C.7}$$

and

$$\text{(C.8)}\qquad \begin{aligned}\int_{\rho_J} |\mathscr{F}_J(\omega; v)|^2 \, d\omega &= \int_{\mathbb{R}^d} |\mathscr{F}_J(\omega; v)|^2 \, d\omega - \int_{\rho_J^c} |\mathscr{F}_J(\omega; v)|^2 \, d\omega \\ &= (2\pi)^d \int_{S_J} v(x)^2 \, dx + o\big(w_n^2\big).\end{aligned}$$

Using (C.7),

$$\text{(C.9)}\qquad \begin{aligned}&\int_{\rho_J} \sum_K{}' f(\omega + 2\pi\alpha_J K) |\mathscr{F}_J(\omega + 2\pi\alpha_J K; v)|^2 \, d\omega \\ &\quad \ll \sum_K{}' g\big(\alpha_J |K|\big) \int_{\rho_J^c} |\mathscr{F}_J(\omega; v)|^2 \, d\omega \\ &\quad = o\big(w_n^2 g(n^{1/2})\big).\end{aligned}$$

Straightforward calculations and (C.5) yield

$$\text{(C.10)}\qquad \int_{\rho_J} \sum_K{}' f(\omega + 2\pi\alpha_J K) \mathscr{V}_J(\omega; v) \, d\omega = o\big(w_n^2 g(n^{1/2})\big).$$

Similar to the proofs of (2.5) and (3.1) in Stein (1995a), (C.9) and (C.10) imply

$$\text{(C.11)}\quad \mathrm{var}_J = \int_{\rho_J} \sum_K{}' f(\omega + 2\pi\alpha_J K) |\mathscr{F}_J(\omega; v)|^2 \, d\omega + o\big(w_n^2 g(n^{1/2})\big).$$

Finally, (3.11) follows from (C.6), (C.8), (C.11) and

$$\int_{\rho_J} \sum_K{}' \big| f(\omega + 2\pi\alpha_J K) - f(2\pi\alpha_J K) \big| \big|\mathscr{F}_J(\omega; v)\big|^2 \, d\omega = o\big(w_n^2 g(n)\big),$$

where this last bound is proven using a similar argument to the proof of (2.5) of Stein (1993). □

Acknowledgment. I would like to thank Peter McCullagh and Carlos Kenig for helpful discussions on conformal mapping.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. (1965). *Handbook of Mathematical Functions*. Dover, New York.

BABENKO, V. F. (1976). Asymptotically sharp bounds for the remainder for the best quadrature formulas for several classes of functions. *Math. Notes* **19** 187–193.

BABENKO, V. F. (1977). On the optimal error bound for cubature formulae on certain classes of continuous functions. *Anal. Math.* **3** 3–9.

BENHENNI, K. and CAMBANIS, S. (1992). Sampling designs for estimating integrals of stochastic processes. *Ann. Statist.* **20** 161–194.

BINGHAM, N. H. (1972). A Tauberian theorem for integral transforms of the Hankel type. *J. London Math. Soc.* **5** 493–503.

BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Regular Variation*. Cambridge Univ. Press.

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

CARRIER, G. F., KROOK, M. and PEARSON, C. E. (1966). *Functions of a Complex Variable: Theory and Technique*. McGraw-Hill, New York.

CARMAN, P. (1974). *n-Dimensional Quasiconformal* (*QCf*) *Mappings*. Abacus Press, Turnbridge Wells.

CASSELS, J. W. S. (1959). On a problem of Rankin about the Epstein zeta-function. *Proceedings of the Glasgow Mathematical Association* **6** 73–80.

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

DALENIUS, T., HÁJEK, J. and ZUBRZYCKI, S. (1961). On plane sampling and related geometrical problems. *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* **1** 125–150. Univ. California Press, Berkeley.

DAVIS, P. J. and RABINOWITZ, P. (1984). *Methods of Numerical Integration*, 2nd ed. Academic Press, New York.

DELVOS, F.-J. (1989). *R*-th order blending rectangle rules. In *Multivariate Approximation Theory* (C. K. Chui, W. Schempp and K. Zeller, eds.) **4** 107–114. Birkhäuser, Basel.

DIANANDA, P. H. (1964). Notes on two lemmas concerning the Epstein zeta-function. *Proceedings of the Glasgow Mathematical Association* **6** 202–204.

ENNOLA, V. (1964a). A lemma about the Epstein zeta-function. *Proceedings of the Glasgow Mathematical Association* **6** 198–201.

ENNOLA, V. (1964b). On a problem about the Epstein zeta-function. *Proc. Cambridge Philos. Soc.* **60** 855–875.

EUBANK, R. L., SMITH, P. L. and SMITH, P. W. (1981). Uniqueness and eventual uniqueness of optimal designs in some time series models. *Ann. Statist.* **9** 486–493.

HILLE, E. (1959). *Analytic Function Theory*. Blaisdell, New York.

LEVIN, M. and GIRSHOVICH, J. (1979). *Optimal Quadrature Formulas*. Teubner, Leipzig.

MATÉRN, B. (1960). *Spatial Variation. Meddelanden från Statens Skogsforskningsinstitut* **49**. Springer, Berlin. (2nd ed. 1986).

NIEDERREITER, H. (1992). *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM, Philadelphia.

PITT, L. D., ROBEVA, R. and WANG, D. Y. (1995). An error analysis for the numerical calculation of certain random integrals: Part 1. *Ann. Appl. Probab.* **5** 171–197.

POLOVINKIN, V. I. (1989). Asymptotically optimal cubature weight formulas. *Siberian Math. J.* **30** 289–297.

QUENOUILLE, M. H. (1949). Problems in plane sampling. *Ann. Math. Statist.* **20** 355–375.

RANKIN, R. A. (1953). A minimum problem for the Epstein zeta-function. *Proceedings of the Glasgow Mathematical Association* **1** 149–158.

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

SCHOENFELDER, C. (1982). Random designs for estimating integrals of stochastic processes: asymptotics. Report 6, Center for Stochastic Processes, Dept. Statistics, Univ. North Carolina.

SCHOENFELDER, C. and CAMBANIS, S. (1982). Random designs for estimating integrals of stochastic processes. *Ann. Statist.* **10** 526–538.

SHUSHBAEV, S. SH. (1989). Local minima of the Epstein zeta-function. *Math. Notes* **45** 83–88.

SOBOLEV, S. L. (1974). *Introduction to the Theory of Cubature Formulae*. Nauka, Moscow.

SOBOLEV, S. L. (1992). *Cubature Formulas and Modern Analysis: An Introduction*. Oxonian Press, New Delhi.

STEIN, M. L. (1993). Asymptotic properties of centered systematic sampling for predicting integrals of spatial processes. *Ann. Appl. Probab.* **3** 874–880.

STEIN, M. L. (1995a). Predicting integrals of random fields using observations on a lattice. *Ann. Statist.* **23** 1975–1990.

STEIN, M. L. (1995b). Predicting integrals of stochastic processes. *Ann. Appl. Probab.* **5** 158–170.

TITCHMARSH, E. C. (1939). *The Theory of Functions*, 2nd ed. Oxford Univ. Press.

TUBILLA, A. (1975). Error convergence rates for estimates of multidimensional integrals of random functions. Report 72, Dept. Statistics, Stanford Univ.

WAHBA, G. (1990). *Spline Models for Observational Data*. SIAM, Philadelphia.

YLVISAKER, D. (1975). Designs on random fields. In *A Survey of Statistical Design and Linear Models* (J. N. Srivastava, ed.) 593–607. North-Holland, Amsterdam.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CHICAGO
5734 UNIVERSITY AVENUE
CHICAGO, ILLINOIS 60637-1514