The Annals of Statistics
1996, Vol. 24, No. 5, 2081–2096

ASYMPTOTIC OPTIMALITY OF REGULAR SEQUENCE DESIGNS

BY KLAUS RITTER

Universität Erlangen-Nürnberg

We study linear estimators for the weighted integral of a stochastic process. The process may only be observed on a finite sampling design. The error is defined in a mean square sense, and the process is assumed to satisfy Sacks–Ylvisaker regularity conditions of order $r \in \mathbb{N}_0$. We show that sampling at the quantiles of a particular density already yields asymptotically optimal estimators. Hereby we extend the results of Sacks and Ylvisaker for regularity $r = 0$ or 1, and we confirm a conjecture by Eubank, Smith and Smith.

1. Introduction. Let $X(t)$, $t \in [0,1]$, be a centered stochastic process which is at least continuous in quadratic mean. For a known function $\rho \in L_2([0,1])$ we want to estimate the weighted integral

$$\mathrm{Int}_\rho(X) = \int_0^1 X(t)\rho(t)\,dt.$$

We consider linear estimators I_n which are based on n observations of X. Hence

$$I_n(X) = \sum_{i=1}^n a_i X(t_i),$$

with sampling points $0 \le t_1 < \cdots < t_n \le 1$ and coefficients $a_i \in \mathbb{R}$. The error of I_n is defined in a mean square sense by

$$e(I_n, \rho, K) = \Big(E\big(\mathrm{Int}_\rho(X) - I_n(X)\big)^2\Big)^{1/2}.$$

Here E denotes the expectation, K denotes the covariance kernel of X and the error depends on X only through K.

It is well known how to choose the coefficients a_i optimally if the sampling design $T_n = \{t_1, \dots, t_n\}$ is fixed and if K is known. Let $K_{T_n} = (K(t_i, t_j))_{i,j}$ denote the covariance matrix of the observations, and let $b_{T_n} = (\int_0^1 K(s, t_i) \times \rho(s)\,ds)_i$. Then I_n^* has minimal error in the class of all linear estimators that

Received February 1995; revised October 1995.
AMS 1991 *subject classifications*. Primary 62K05, 41A55; secondary 60G12, 62M99, 65D30.
Key words and phrases. Integral estimation, asymptotically optimal designs, regular sequence designs, Sacks–Ylvisaker conditions.

use the sampling design T_n iff $a = (a_i)_i$ solves $K_{T_n} a = b_{T_n}$. The corresponding error is

$$e(T_n, \rho, K) = e(I_n^*, \rho, K)$$
$$= \left(\int_0^1 \int_0^1 K(s,t)\rho(s)\rho(t)\, ds\, dt - b_{T_n}^t K_{T_n}^{-1} b_{T_n} \right)^{1/2}$$

if K_{T_n} is nonsingular, which may be assumed without loss of generality.

In this paper we address the design problem of how to minimize $e(T_n, \rho, K)$ over all n-point designs T_n. The infimum

$$e_n(\rho, K) = \inf_{T_n} e(T_n, \rho, K)$$

is called the nth minimal error, and any design T_n with $e(T_n, \rho, K) = e_n(\rho, K)$ is called optimal.

The design problem for integral estimation was introduced by Suldin (1959, 1960), who considered the Brownian motion X where $K(s,t) = \min(s,t)$. It turned out that $T_n = \{2i/(2n+1): i = 1, \dots, n\}$ is the optimal n-point design for the constant weight function $\rho = 1$. The optimal coefficients given T_n are $a_i = 2/(2n+1)$, and this gives the minimal error $e_n(\rho, K) = (\sqrt{3}\,(2n+1))^{-1}$.

We briefly present two of several equivalent formulations for the design problem. See Sacks and Ylvisaker (1970b), Cambanis (1985) and Traub, Wasilkowski and Woźniakowski (1988) for details and further equivalences.

Let $H(K)$ denote the Hilbert space with reproducing kernel K. This space consists of real-valued functions h: $[0,1] \to \mathbb{R}$ such that $K(\cdot, t) \in H(K)$ and $h(t) = \langle h, K(\cdot, t)\rangle_K$ for any $t \in [0,1]$ and any $h \in H(K)$. See, for example, Aronszajn (1950), Parzen (1959) and Wahba (1990). For any estimator I_n the (average) error $e(I_n, \rho, K)$ coincides with the maximal error of I_n on the unit ball in $H(K)$, that is,

$$e(I_n, \rho, K) = \sup\left\{ \left| \mathrm{Int}_\rho(h) - I_n(h) \right| : h \in H(K), \|h\|_K \le 1 \right\},$$

where $\|\cdot\|_K$ denotes the norm in $H(K)$. Hence the design problem for integral estimation is equivalent to finding the worst case optimal quadrature formula on a unit ball in a reproducing kernel Hilbert space. The general problem of optimal quadrature formulas goes back to Nikolskij (1950).

Consider the linear regression model

$$Y(t) = \beta f(t) + X(t),$$

with an unknown constant $\beta \in \mathbb{R}$, where

$$f(t) = \int_0^1 K(s,t)\rho(s)\, ds$$

is known. Then $(b_{T_n}^t K_{T_n}^{-1} b_{T_n})^{-1}$ is the variance of the best linear unbiased estimator for β which is based on the design T_n. Hence the design problems for integral estimation and for linear regression are equivalent if f and ρ are related as above. In both cases $b_{T_n}^t K_{T_n}^{-1} b_{T_n}$ has to be maximized over all

n-point designs T_n. In the context of a regression model with correlated errors, the design problem was formulated by Sacks and Ylvisaker (1966).

The design problem is solved exactly only in a few cases. Therefore, Sacks and Ylvisaker (1966) have introduced the notion of asymptotic optimality. A sequence $(T_n)_n$ of n-point designs is called asymptotically optimal if

$$e(T_n, \rho, K) \approx e_n(\rho, K), \quad \text{that is,} \lim_{n\to\infty} \frac{e(T_n, \rho, K)}{e_n(\rho, K)} = 1.$$

Regular sequences of designs are of particular interest. The corresponding sampling points are defined as quantiles of a fixed continuous density ψ on $[0,1]$; see Sacks and Ylvisaker (1970a, 1970b). We have $T_n = \{t_{1,n}, \dots, t_{n,n}\}$ with

$$\int_0^{t_{i,n}} \psi(t)\, dt = \frac{i-1}{n-1} \int_0^1 \psi(t)\, dt,$$

and we use the notation $(T_n)_n = \mathrm{RS}(\psi)$. Regular sequences enjoy several advantages. They are easy to describe and implement, and they are also well suited for an error analysis. These sequences are quasi-uniform if $\psi > 0$; moreover, $t_{2i-1,2n-1} = t_{i,n}$. For nonparametric regression, regular sequences are studied in Speckman (1985) and Golubev and Nussbaum (1990).

The quality of a regular sequence for integral estimation clearly depends on ψ, ρ and K, and the following questions arise: do regular sequences already lead to asymptotically optimal designs, that is, does there exist a density ψ such that

$$(*) \qquad e(T_n, \rho, K) \approx e_n(\rho, K) \quad \text{for } (T_n)_n = \mathrm{RS}(\psi)?$$

In case of a positive answer: is it possible to construct such a density without knowing the covariance kernel K of the process X precisely?

Positive answers are given in some situations. Sacks and Ylvisaker (1966) consider processes X which are nowhere differentiable in quadratic mean, and they obtain $(*)$ with $\psi = |\rho|^{2/3}$. The result is proven for a class of kernels K which is defined by certain regularity conditions. In particular, the Brownian motion kernel is covered. If X' exists in quadratic mean and if the covariance kernel $K^{(1,1)}$ of X' satisfies the regularity conditions, then $(*)$ holds with $\psi = |\rho|^{2/5}$; see Sacks and Ylvisaker (1970a, 1970b). The order of the minimal errors is

$$(**) \qquad e_n(\rho, K) \approx b_r \cdot \left(\int_0^1 |\rho(t)|^{2/(2r+3)}\, dt \right)^{(2r+3)/2} n^{-(r+1)}.$$

Here r is the regularity in quadratic mean, $r = 0$ or 1, and

$$b_r^2 = \frac{|B_{2r+2}|}{(2r+2)!},$$

where B_k denotes the kth Bernoulli number.

Eubank, Smith and Smith (1981, 1982) consider the r-fold integrated Brownian motion, possibly pinned with some derivatives at $t = 1$. They obtain $(*)$ with $\psi = |\rho|^{2/(2r+3)}$ and $(**)$ for these specific processes of higher regularity. Eubank, Smith and Smith (1982), Remark 1, conjecture that their results hold for more general kernels as those which satisfy Sacks–Ylvisaker conditions of order r.

Benhenni and Cambanis (1992a) consider slightly different regularity conditions and give further support to the conjecture. They determine the asymptotic order of $e(T_n, \rho, K)$ for any regular sequence $(T_n)_n = \mathrm{RS}(\psi)$. The smallest asymptotic constant is again obtained for $\psi = |\rho|^{2/(2r+3)}$, and the right-hand side of $(**)$ is equivalent to $e(T_n, \rho, K)$ in this case. Since lower bounds for $e_n(\rho, K)$ were not obtained, the optimality $(*)$ and the equivalence $(**)$ remained open.

In this paper we prove the conjecture from Eubank, Smith and Smith (1982). The proof is based on results for integral estimation by Barrow and Smith (1979) and Eubank, Smith and Smith (1981). Furthermore, we use results on the set $H(K)$ under Sacks–Ylvisaker conditions, which are due to Ritter, Wasilkowski and Woźniakowski (1995). Finally, we provide some asymptotic equivalences of norms.

Related results for integral estimation based on Hermite data $X(t_i), \dots, X^{(r)}(t_i)$ are known. See Sacks and Ylvisaker (1970a, 1970b), Wahba (1971, 1974), Hájek and Kimeldorf (1974) and Benhenni and Cambanis (1992b), where analogs to $(*)$ and $(**)$ are obtained. The problem of reconstructing X in the L_2-norm is studied by Speckman (1979), Su and Cambanis (1993) and Müller-Gronbach (1996). Integration, reconstruction and differentiation of X based on noisy observations $X(t_i) + \varepsilon_i$ are studied by Plaskota (1992) and Ritter (1996).

Much less is known in the multivariate case for random fields $X(t)$, $t \in D$, with $D \subset \mathbb{R}^d$ and $d > 1$. Woźniakowski (1991, 1992) and Paskov (1993) study integral estimation and L_2-reconstruction for the r-fold integrated Brownian sheet. Wasilkowski (1993, 1994) and Ritter and Wasilkowski (1996) study both problems for (smooth) isotropic Brownian motion. Sharp bounds are obtained in these papers. Bounds which depend on the smoothness of K are given in Ylvisaker (1975), Wittwer (1978), Micchelli and Wahba (1981), Ritter, Wasilkowski and Woźniakowski (1993) and Weba (1995). Sharp bounds for tensor products of kernels which satisfy the Sacks–Ylvisaker conditions are obtained in Ritter, Wasilkowski and Woźniakowski (1995). However, in the multivariate case only weakly asymptotically optimal designs are constructed and asymptotic constants are unknown.

2. Results and remarks. Regularity in quadratic mean of the stochastic process X is specified by the regularity of its covariance kernel at the diagonal in $[0,1]^2$. We denote one-sided limits at this diagonal in the following way. Let

$$\Omega_+ = \left\{(s,t) \in (0,1)^2 \colon s > t\right\}, \qquad \Omega_- = \left\{(s,t) \in (0,1)^2 \colon s < t\right\},$$

and let cl A denote the closure of a set A. Suppose that L is a continuous function on $\Omega_+ \cup \Omega_-$ such that $L|_{\Omega_j}$ is continuously extendable to cl Ω_j for $j \in \{+, -\}$. By L_j we denote the extension of L to $[0,1]^2$ which is continuous on cl Ω_j and on $[0,1]^2 \setminus \text{cl}\, \Omega_j$.

A covariance kernel K on $[0,1]^2$ satisfies the Sacks–Ylvisaker conditions of order $r \in \mathbb{N}_0$ if the following three conditions hold:

(A) $K \in C^{r,r}([0,1]^2)$, the partial derivatives of $L = K^{(r,r)}$ up to order 2 are continuous on $\Omega_+ \cup \Omega_-$ and continuously extendable to cl Ω_+ as well as to cl Ω_-.

(B) There is a constant $\alpha > 0$ with

$$L_-^{(1,0)}(s,s) - L_+^{(1,0)}(s,s) = \alpha, \qquad 0 \le s \le 1.$$

(C) $L_+^{(2,0)}(s, \cdot) \in H(L)$ for all $0 \le s \le 1$ and

$$\sup_{0 \le s \le 1} \left\| L_+^{(2,0)}(s, \cdot) \right\|_L < \infty.$$

In a series of papers Sacks and Ylvisaker (1966, 1968, 1970a, 1970b) have introduced these conditions to study the design problem. A detailed discussion and various examples are given there. If K fulfills (A) and corresponds to a wide sense stationary process, then the difference in (B) is always a nonnegative constant. For arbitrary kernels which satisfy (A) and (C), this difference is constant, too.

For $r = 0$, the conditions are satisfied in particular if $K(s,t)$ is given by $\min(s,t)$, $1 - |s-t|$ or $\exp(-|s-t|)$. Kernels of higher regularity may be obtained by r-fold integration of a corresponding process with deterministic or stochastic boundary conditions. Modifications of (A), (B) and (C) are used by several authors. A partial list of references includes Benhenni and Cambanis (1992a, 1992b), Su and Cambanis (1993) and Müller-Gronbach (1996).

Henceforth we assume

$$\alpha = 1$$

to simplify the notation. We define

$$J_{\rho,r}(\psi) = \left(\int_0^1 \rho(t)^2 \psi(t)^{-(2r+2)}\, dt \right)^{1/2} \left(\int_0^1 \psi(t)\, dt \right)^{r+1}.$$

THEOREM 1. *Let K satisfy the Sacks–Ylvisaker conditions of order $r \in \mathbb{N}_0$ together with the boundary conditions*

$$K^{(r,k)}(\cdot, 0) = 0, \qquad k = 0, \dots, r-1.$$

Then

$$e(T_n, \rho, K) \approx b_r J_{\rho,r}(\psi) n^{-(r+1)}$$

for any regular sequence $(T_n)_n = \mathrm{RS}(\psi)$ which is generated by a positive, continuous density ψ.

The upper bound in the next theorem follows immediately from

$$J_{\rho,r} = \inf\{J_{\rho,r}(\psi) : \psi \text{ positive, continuous density}\},$$

where

$$J_{\rho,r} = \left(\int_0^1 |\rho(t)|^{2/(2r+3)}\, dt\right)^{(2r+3)/2};$$

see Sacks and Ylvisaker (1970a, 1970b).

THEOREM 2. *Under the assumptions of Theorem* 1, *the minimal errors satisfy* $(**)$, *that is*,

$$e_n(\rho, K) \approx b_r J_{\rho,r} n^{-(r+1)}.$$

Note that $J_{\rho,r}(\psi) = J_{\rho,r}$ for $\psi = |\rho|^{2/(2r+3)}$, and this density depends on K only through its regularity r in the Sacks–Ylvisaker conditions. Therefore, Theorems 1 and 2 imply positive answers to our questions, at least if ρ is continuous and never 0.

COROLLARY. *Under the assumptions of Theorem* 1, *the asymptotic optimality* $(*)$ *holds with* $\psi = |\rho|^{2/(2r+3)}$ *if* ρ *is continuous and never* 0. *For arbitrary* $\rho \in L_2([0,1])$, *suitable regular sequences yield errors which differ from* $e_n(\rho, K)$ *by arbitrary small constants if n is sufficiently large.*

REMARK 1. Theorems 1 and 2 and the corollary are due to Sacks and Ylvisaker (1966, 1970a, 1970b) for the case $r = 0$ and 1. These authors have even obtained the asymptotic optimality of $|\rho|^{2/(2r+3)}$ for any continuous ρ if $r = 0$ and, under mild assumptions on the 0's of ρ, if $r = 1$. Benhenni and Cambanis (1992a) obtain Theorem 1 for arbitrary $r \in \mathbb{N}_0$ and $\rho, \psi \in C^{r+2}([0,1])$ under slightly different regularity conditions for K.

Eubank, Smith and Smith (1981, 1982) show Theorems 1 and 2 and the corollary for some specific kernels. In particular, they consider the reproducing kernel P_r of the Hilbert space

$$H(P_r) = \{h \in W_2^{r+1}([0,1]) : h^{(k)}(0) = h^{(k)}(1) = 0 \text{ for } k = 0, \dots, r\}$$

equipped with the norm

$$\|h\|_{P_r} = \|h^{(r+1)}\|_2.$$

Note that P_r is the covariance kernel of an r-fold integrated Brownian motion, pinned with all derivatives at $t = 1$. For $r = 0$ we have the Brownian bridge kernel $P_0(s,t) = \min(s,t) - st$. Furthermore, Barrow and Smith (1979) prove Theorems 1 and 2 and the corollary for the maximal error

$$\sup\{|\mathrm{Int}_\rho(h) - I_n(h)| : h \in W_2^{r+1}([0,1]),\ \|h^{(r+1)}\|_2 \le 1\}$$

on the unit ball with respect to the Sobolev seminorm $\|h^{(r+1)}\|_2$.

REMARK 2. The upper bound $e_n(\rho, K) = O(n^{-(r+1)})$ already follows from condition (A); see Sacks and Ylvisaker (1970b). Condition (B) restricts the smoothness of the process X, but it is not sufficient to show that the above bound is sharp. While (A), (B) and (C) lead to sharp bounds for regularity $r = 0$, Example 1 from Ritter, Wasilkowski and Woźniakowski (1995) shows that this is not true for $r > 0$. The conditions only guarantee the asymptotic upper bounds in Theorems 1 and 2. Thus we use the boundary conditions $K^{(r,k)}(\cdot, 0) = 0$ for $k = 0, \dots, r-1$, which state that $X^{(r)}$ and $X^{(k)}(0)$ are uncorrelated.

Suppose that a stationary process $Y(t)$, $t \in [0,1]$, is given, whose covariance kernel L satisfies the Sacks–Ylvisaker conditions of order $r = 0$. Mitchell, Morris and Ylvisaker (1990) and Lasinger (1993) show how to preserve stationarity by r-fold integration of Y. Clearly, the covariance kernel K of the integrated process X satisfies the Sacks–Ylvisaker conditions of order r. However, the boundary conditions cannot be satisfied if X is stationary. Nevertheless, one can show that the conclusions of Theorems 1 and 2 are true, as long as the variances $K^{(k,k)}(0,0)$ of the processes $X^{(k)}$ are sufficiently large.

Weba (1991) considers L_p-processes X with continuous derivatives up to order $r = 2m + 2$ in an L_p-sense. He obtains upper bounds for the errors, defined in an L_p-sense, of specific quadrature formulas. For instance, the Romberg method yields the error bound $O(n^{-r})$, as in the classical case.

REMARK 3. So far we have studied integral estimation based on samples $X(t_i)$ only. Now we discuss estimators

$$I_n(X) = \sum_{i=1}^{n} a_i X^{(k_i)}(t_i)$$

which use derivatives of order $k_i \in \{0, \dots, r\}$. We compare such estimators by their errors and by the total number n of observations.

Due to results by Zhensykbaev (1983) derivatives do not help for nonnegative ρ and certain kernels K, as those considered by Eubank, Smith and Smith (1981, 1982). The minimal error in the class of all linear estimators I_n which may use derivatives is attained by an estimator with $k_i = 0$.

It follows that derivatives do not help at least asymptotically under the assumptions of Theorem 1 if $\rho \geq 0$. The minimal errors in the two classes of estimators with and without derivatives are equivalent. The proof applies Zhensykbaev's result to $K = P_r$ and is similar to the proof of the lower bounds in Theorem 2.

Sacks and Ylvisaker (1970a, b) study estimators which use Hermite data, that is,

$$I_n(X) = \sum_{i=1}^{l} \sum_{k=0}^{r} a_{i,k} X^{(k)}(t_i) \quad \text{with } n = l(r+1).$$

They show that the errors of optimal coefficient estimators which are based on a regular sequence $(T_n)_n = \mathrm{RS}(\psi)$ are equivalent to

$$\tilde{b}_r \cdot J_{\rho,r}(\psi) l^{-(r+1)}, \tag{1}$$

where

$$\tilde{b}_r = \frac{(r+1)!}{((2r+2)!(2r+3)!)^{1/2}}.$$

Moreover, the minimal errors in the class of all linear estimators which use Hermite data at l sampling points are equivalent to

$$\tilde{b}_r J_{\rho,r} l^{-(r+1)}. \tag{2}$$

Clearly, $b_0 = \tilde{b}_0$, but note that $b_1 = \tilde{b}_1$, too. Since

$$\lim_{r\to\infty} \frac{\tilde{b}_r l^{-(r+1)}}{b_r n^{-(r+1)}} = \lim_{r\to\infty} (r+1)^{r+1} \frac{\tilde{b}_r}{b_r} = \infty,$$

we see that Hermite data are disadvantageous, from this point of view, if n and r are large. Moreover, it may be impracticable to use high-order derivatives for the integration of smooth processes.

Wahba (1971, 1974) and Hájek and Kimeldorf (1974) construct covariance kernels from differential operators in the following way. Let $\mathscr{D}^i f = f^{(i)}$ and let

$$\mathscr{L} = \sum_{i=0}^{r+1} \beta_i \mathscr{D}^i$$

denote the linear differential operator of order $r+1$ with coefficients $\beta_i \in C^i([0,1])$ and β_{r+1} never being 0. Furthermore, let G denote the Green's function for the initial value problem $\mathscr{L} f = u$ and $f^{(i)}(0) = 0$ for $i = 0, \dots, r$. Clearly,

$$K(s,t) = \int_0^1 G(s,u) G(t,u)\, du \tag{3}$$

defines a positive definite function, and any process with covariance kernel K has exactly r derivatives in quadratic mean.

Wahba (1971), in a particular case, and Hájek and Kimeldorf (1974), in the general case, obtain (1) and (2), with ρ replaced by ρ/β_{r+1}, when Hermite data are available. Vector processes are studied in Wittwer (1976). Another generalization is due to Wahba (1974). She considers kernels K such that the corresponding centered Gaussian process is equivalent to the centered Gaussian process with a kernel given by (3). For two kernels of the form (3), this equivalence holds iff the leading coefficients β_{r+1} of the differential operators $\mathscr{L}$ coincide.

Our proof for integral estimation based on samples $X(t_i)$ also applies to covariance kernels K of the form (3). The conclusions of Theorems 1 and 2 and the corollary are true if ρ is replaced by ρ/β_{r+1} again.

Comparing the two cases, sampling data $X(t_i)$ or sampling Hermite data, we see the same dependence on ψ of the quality of regular sequence designs. Furthermore, regular sequence designs lead to asymptotically optimal estimators in both cases. We add that equidistant knots always yield the error bound $O(n^{-(r+1)})$. However, they are arbitrarily bad with respect to the asymptotic constant for some weight functions ρ.

REMARK 4. We already know how to construct asymptotically optimal regular sequences if only the regularity r of the covariance kernel K in the Sacks–Ylvisaker conditions is known. However, we have studied optimal coefficient estimators once a design is fixed, and optimal coefficients depend on the covariances $K(t_i, t_j)$ and $\int_0^1 K(s, t_i)\rho(s)\,ds$.

On the other hand, one can also use suitable coefficients which depend on K only through r to get still asymptotically optimal estimators. For $r = 0$ and $r = 1$, a construction is given in Sacks and Ylvisaker (1970b). For arbitrary r, Benhenni and Cambanis (1992a) show that the asymptotical behavior does not change if we switch from optimal coefficient estimators to weighted Gregory formulas, given a regular sequence RS(ψ). The Gregory formulas only depend on ρ, ψ and r. Combining this result from Benhenni and Cambanis (1992a) with Theorem 2, the asymptotic optimality of the Gregory formulas with $\psi = |\rho|^{2/(2r+3)}$ follows. For integral estimation based on Hermite data, a weighted Rodriguez formula enjoys the same properties; see Benhenni and Cambanis (1992b). Practical experiments are presented in Benhenni and Cambanis (1992a, b). Additional results are presented in Cambanis (1985) and Istas and Laredo (1994).

Further estimators with simple coefficients are defined as weighted integrals of natural polynomial splines of degree $2r + 1$ which interpolate the data $X(t_i)$. It would be interesting to know whether these estimators, based on RS($|\rho|^{2/(2r+3)}$), are asymptotically optimal, too.

3. Proofs. Henceforth let K be a covariance kernel which satisfies the Sacks–Ylvisaker conditions of order $r \in \mathbb{N}_0$ with $\alpha = 1$ in (B). Assume that the boundary conditions $K^{(r,k)}(\cdot, 0)$ for $k = 0, \dots, r-1$ hold. For any n-point design T_n let

$$B(K, T_n) = \{h \in H(K) : \|h\|_K \le 1,\ h(t) = 0 \text{ for } t \in T_n\}.$$

From the worst case formulation of the design problem, we get

$$e(I_n, \rho, K) \ge \sup\{|\mathrm{Int}_\rho(h)| : h \in B(K, T_n)\}$$

for any linear estimator which is based on T_n. We have equality for optimal coefficient estimators; see Traub, Wasilkowski and Woźniakowski (1988), page 76. Hence

$$e(T_n, \rho, K) = \sup\{|\mathrm{Int}_\rho(h)| : h \in B(K, T_n)\}.$$

Therefore, we study the Hilbert space $H(K)$ and, in particular, the unit ball $B(K, T_n)$ in the subspace of all functions from $H(K)$ which vanish at the sampling design.

A proof of the following result is given in Ritter, Wasilkowski and Woźniakowski (1995) in the case of boundary conditions $K^{(0,k)}(\cdot, 0)$ for $k = 0, \dots, r-1$. The proof is easily adapted to the more general boundary conditions from Theorem 1.

LEMMA 1.

$$H(P_r) \subset H(K) \subset W_2^{r+1}([0,1]),$$

where $H(P_r)$ *is defined in Remark* 1.

Embeddings between reproducing kernel Hilbert spaces are continuous, and hence we have

$$\begin{aligned} c_1 e(T_n, \rho, P_r) &\le e(T_n, \rho, K) \\ &\le c_2 \sup\Big\{ \big|\mathrm{Int}_\rho(h)\big| : h \in W_2^{r+1}([0,1]),\ \|h^{(r+1)}\|_2 \le 1, \\ &\qquad\qquad h(t) = 0 \text{ for } t \in T_n \Big\}, \end{aligned}$$

with constants $c_i > 0$ which only depend on K. To get tight bounds, we compare the (semi)norms $\|h\|_K$ and $\|h^{(r+1)}\|_2$ for functions h which vanish on T_n.

Let T_n consist of sampling points $0 \le t_1 < \cdots < t_n \le 1$. Put $t_0 = 0$ and $t_{n+1} = 1$ to define

$$\delta = \delta_{T_n} = \max_{i=1,\dots,n+1} (t_i - t_{i-1}).$$

The following estimate can be verified by induction.

LEMMA 2. *If* $n \ge r+1$ *and* $h \in W_2^{r+1}([0,1])$ *with* $h(t) = 0$ *for* $t \in T_n$, *then*

$$\|h^{(k)}\|_2 \le \frac{(r+1)!}{k!} \delta^{r-k+1} \|h^{(r+1)}\|_2$$

for $k = 0, \dots, r$.

LEMMA 3. *There exists a constant* $c > 0$, *depending only on* K, *with the following property. If* $n > 2r+1$ *and* $h \in B(K, T_n)$ *then*

$$\left(\int_{t_{r+1}}^{t_{n-r}} h^{(r+1)}(s)^2 \, ds \right)^{1/2} \le 1 + c\delta.$$

If $n \ge r$ *and* $h \in B(P_r, T_n)$, *then*

$$\|h\|_K \le 1 + c\delta.$$

PROOF. First, we consider the case $r = 0$. Let $\hat{K}$ denote the restriction of K to $[t_1, t_n]^2$, and let

$$\hat{h} = \sum_{j=1}^{m} b_j \hat{K}(\cdot, u_j), \tag{4}$$

with distinct $t_1 \le u_j \le t_n$ such that $\hat{h}(t) = 0$ for $t \in T_n$. Integration by parts yields

$$\int_{t_1}^{t_n} \hat{h}'(s)^2\, ds = \sum_{j,k=1}^{m} b_j b_k \Bigg(K(t_n, u_j) K_+^{(1,0)}(t_n, u_k) - K(t_1, u_j) K_-^{(1,0)}(t_1, u_k) + K(u_j, u_k)\big(K_-^{(1,0)}(u_k, u_k) - K_+^{(1,0)}(u_k, u_k)\big) - \int_{t_1}^{t_n} K(s, u_j) K_+^{(2,0)}(s, u_k)\, ds \Bigg).$$

Observe that $0 = \hat{h}(t_1) = \hat{h}(t_n) = \sum_{j=1}^m b_j \cdot K(t_1, u_j) = \sum_{j=1}^m b_j \cdot K(t_n, u_j)$, and due to (B) we have

$$\sum_{j,k=1}^{m} b_j b_k K(u_j, u_k)\big(K_-^{(1,0)}(u_k, u_k) - K_+^{(1,0)}(u_k, u_k)\big) = \sum_{j,k=1}^{m} b_j b_k K(u_j, u_k) = \|\hat{h}\|_{\hat{K}}^2.$$

Thus

$$\|\hat{h}\|_{\hat{K}}^2 - \int_{t_1}^{t_n} \hat{h}'(s)^2\, ds = c_1,$$

where

$$c_1 = \sum_{j,k=1}^{m} b_j b_k \int_{t_1}^{t_n} K(s, u_j) K_+^{(2,0)}(s, u_k)\, ds = \sum_{k=1}^{m} b_k \int_{t_1}^{t_n} \hat{h}(s) K_+^{(2,0)}(s, u_k)\, ds.$$

For $t_1 \le s \le t_n$ let $\hat{g}_s$ denote the restriction of $K_+^{(2,0)}(s, \cdot)$ to $[t_1, t_n]$. From (C) we get $\hat{g}_s \in H(\hat{K})$. Hence

$$\sum_{k=1}^{m} b_k K_+^{(2,0)}(s, u_k) = \sum_{k=1}^{m} b_k \big\langle \hat{g}_s, \hat{K}(\cdot, u_k) \big\rangle_{\hat{K}} = \langle \hat{g}_s, \hat{h} \rangle_{\hat{K}}$$

and

$$|c_1| = \left| \int_{t_1}^{t_n} \hat{h}(s) \langle \hat{g}_s, \hat{h} \rangle_{\hat{K}}\, ds \right| \le c \|\hat{h}\|_{\hat{K}} \int_{t_1}^{t_n} \big|\hat{h}(s)\big|\, ds \le c\delta \|\hat{h}\|_{\hat{K}} \left(\int_{t_1}^{t_n} \hat{h}'(s)^2\, ds \right)^{1/2},$$

with

$$c = \sup_{t_1 \le s \le t_n} \|\hat{g}_s\|_{\hat{K}} \le \sup_{0 \le s \le 1} \big\| K_+^{(2,0)}(s, \cdot) \big\|_K < \infty;$$

see (C) and Lemma 2. Therefore, we obtain

$$\left| \|\hat{h}\|_{\hat{K}}^2 - \int_{t_1}^{t_n} \hat{h}'(s)^2\, ds \right| \le c\delta \|\hat{h}\|_{\hat{K}} \left(\int_{t_1}^{t_n} \hat{h}'(s)^2\, ds \right)^{1/2}.$$

This implies

$$(5)\qquad (1+c\delta)^{-1}\|\hat{h}\|_{\hat{K}} \le \left(\int_{t_1}^{t_n} \hat{h}'(s)^2\,ds\right)^{1/2} \le (1+c\delta)\|\hat{h}\|_{\hat{K}}.$$

Since functions $\hat{h}$ of the form (4) which vanish on T_n are dense in the space of all functions from $H(\hat{K})$ which vanish on T_n, the estimate (5) even holds on the latter space. By $\|h|_{[t_1,t_n]}\|_{\hat{K}} \le \|h\|_K$ the first estimate from Lemma 3 follows. By Lemma 1, $h/\|h\|_K \in B(K, T_n \cup \{0,1\})$ if $h \in B(P_0, T_n)$, and the second estimates follows.

Now we consider the case $r > 0$. Due to (A) we have $K \in C^{r,r}([0,1]^2)$, and this yields the following facts; see Ritter, Wasilkowski and Woźniakowski (1995). By $Uh = h^{(r)}$, a bounded linear operator from $H(K)$ to $H(L)$ is defined, where $L = K^{(r,r)}$. Moreover, $U^*L(\cdot,t) = K^{(0,r)}(\cdot,t)$, which implies $UU^*g = g$ for any $g \in H(L)$.

Because of the boundary conditions

$$(U^*L(\cdot,t))^{(k)}(0) = K^{(k,r)}(0,t) = 0,$$

and therefore

$$(U^*g)(t) = \int_0^t g(u)\frac{(t-u)^{r-1}}{(r-1)!}\,du,$$

that is, the adjoint of U is given by r-fold integration.

Suppose that $h \in W_2^{r+1}([0,1])$ vanishes on T_n. Then $h^{(r)}$ has 0's, $0 < z_1 < \cdots < z_{n-r} < 1$, with $z_1 \le t_{r+1}$ and $z_{n-r} \ge t_{n-r}$ as well as $\max_i(z_i - z_{i-1}) \le (r+1)\delta$. Here $z_0 = 0$ and $z_{n-r+1} = 1$. Clearly, L satisfies the Sacks–Ylvisaker conditions of order $r = 0$. Using $\|h^{(r)}\|_L \le \|h\|_K$ and Lemma 3 with $r = 0$, we get

$$\begin{aligned}\left(\int_{t_{r+1}}^{t_{n-r}} h^{(r+1)}(s)^2\,ds\right)^{1/2} &\le \left(\int_{z_1}^{z_{n-r}} h^{(r+1)}(s)^2\,ds\right)^{1/2} \le (1+c(r+1)\delta)\|h^{(r)}\|_L\\ &\le 1 + c(r+1)\delta\end{aligned}$$

for $h \in B(K, T_n)$.

If $h \in B(P_r, T_n)$, then $h \in H(K)$ (see Lemma 1) and $h^{(r)} \in H(P_0)$ with norm bounded by 1. Moreover, $h = U^*Uh$. Hence

$$\|h\|_K = \|h^{(r)}\|_L \le 1 + c(r+1)\delta$$

follows from Lemma 3 with $r = 0$. $\square$

By means of Lemma 3 we reduce Theorems 1 and 2, as stated with Sacks–Ylvisaker conditions, to the particular versions which are due to Barrow and Smith (1979) and Eubank, Smith and Smith (1981); see Remark 1.

PROOF OF THEOREM 1. Let $(T_n)_n = \mathrm{RS}(\psi)$, where ψ is a positive, continuous density on $[0,1]$. By c we denote positive constants, which only depend on K, ρ and ψ and which may have different values. Clearly,

$$\max_{i=2,\dots,n} (t_{i,n} - t_{i-1,n}) \le \frac{c}{n}.$$

Fix $0 < a < 1/2$ and let $h \in B(K, T_n)$. Note that $\|h^{(r+1)}\|_2 \le c$ by Lemma 1, and $\|h\|_2 \le cn^{-(r+1)}$ by Lemma 2. Therefore,

$$\begin{aligned} \left|\int_0^a h(t)\rho(t)\,dt\right| &\le c\|\rho 1_{[0,a]}\|_2 n^{-(r+1)}, \\ \left|\int_{1-a}^1 h(t)\rho(t)\,dt\right| &\le c\|\rho 1_{[1-a,1]}\|_2 n^{-(r+1)}. \end{aligned} \tag{6}$$

If n is sufficiently large, then

$$\left(\int_{a/2}^{1-a/2} h^{(r+1)}(s)^2\,ds\right)^{1/2} \le 1 + \frac{c}{n}$$

according to Lemma 3. Take $g \in C^{r+1}(\mathbb{R})$ with $g = 0$ on $[0, a/2] \cup [1 - a/2, 1]$, $g = 1$ on $[a, 1-a]$ and $0 < g < 1$ otherwise. Then

$$\begin{aligned} \left\|(gh)^{(r+1)}\right\|_2 \le {} & \left(\int_{a/2}^{1-a/2} h^{(r+1)}(s)^2\,ds\right)^{1/2} \\ & + c \sup_{k=1,\dots,r+1} \left(\|g^{(k)}\|_\infty \|h^{(r+1-k)}\|_2\right) \le A_n, \end{aligned}$$

with

$$A_n = 1 + \frac{c}{n}\left(1 + \sup_{k=1,\dots,r+1} \|g^{(k)}\|_\infty\right)$$

due to Lemma 2. Clearly, $\lim_n A_n = 1$. Define $\zeta = \rho \cdot 1_{[a,1-a]}$ and note that $\mathrm{Int}_\zeta(h) = \mathrm{Int}_\zeta(gh)$. We get

$$\begin{aligned} e(T_n, \zeta, K) &= \sup\left\{\left|\mathrm{Int}_\zeta(h)\right| : h \in B(K, T_n)\right\} \\ &\le A_n \sup\Big\{\left|\mathrm{Int}_\zeta(h)\right| : h \in W_2^{r+1}([0,1]), \\ &\qquad\qquad \|h^{(r+1)}\|_2 \le 1,\ h(t) = 0 \text{ for } t \in T_n\Big\}. \end{aligned}$$

From Barrow and Smith (1979) we know that the right-hand side is equivalent to $b_r J_{\zeta,r}(\psi) n^{-(r+1)}$, and therefore

$$\limsup_{n\to\infty} \left(n^{r+1} e(T_n, \zeta, K)\right) \le b_r J_{\zeta,r}(\psi) \le b_r J_{\rho,r}(\psi).$$

Together with (6) this implies

$$\limsup_{n\to\infty} \left(n^{r+1} e(T_n, \rho, K)\right) \le b_r J_{\rho,r}(\psi) + c\left(\|\rho_{1[0,a]}\|_2 + \|\rho 1_{[1-a,1]}\|_2\right).$$

Letting $a \to 0$, the asymptotic upper bound for $e(T_n, \rho, K)$ follows.

Similarly, Lemma 3 gives

$$
\begin{aligned}
e(T_n, \rho, K) &\geq \sup\{|\mathrm{Int}_\rho(h)|: h \in B(K, T_n) \cap H(P_r)\} \\
&\geq (1 + c/n)^{-1} \sup\{|\mathrm{Int}_\rho(h)|: h \in B(P_r, T_n)\} \qquad (7)\\
&= (1 + c/n)^{-1} e(T_n, \rho, P_r)
\end{aligned}
$$

if $n \geq r$. From Eubank, Smith and Smith (1981) we know that the right-hand side is equivalent to $b_r J_{\rho, r}(\psi) n^{-(r+1)}$, and therefore

$$\liminf_{n \to \infty} \left(n^{r+1} e(T_n, \rho, K)\right) \geq b_r J_{\rho, r}(\psi),$$

which completes the proof. □

Proof of Theorem 2. Take any asymptotically optimal sequence of designs T_n. Without loss of generality, we may assume

$$\lim_{n \to \infty} \delta_{T_n} = 0.$$

Analogously to (7) we get

$$e(T_n, \rho, K) \geq \left(1 + c\delta_{T_n}\right)^{-1} e(T_n, \rho, P_r)$$

and

$$
\begin{aligned}
\liminf_{n \to \infty} \left(n^{r+1} e_n(\rho, K)\right) &\geq \liminf_{n \to \infty} \left(n^{r+1} e(T_n, \rho, P_r)\right) \\
&\geq \lim_{n \to \infty} \left(n^{r+1} e_n(\rho, P_r)\right) = b_r \cdot J_{\rho, r}.
\end{aligned}
$$

Here the equality is due to Eubank, Smith and Smith (1981).

The asymptotic upper bound for $e_n(\rho, K)$ is already a consequence of Theorem 1. □

REFERENCES

Aronszajn, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** 337–404.

Barrow, D. L. and Smith, P. W. (1979). Asymptotic properties of optimal quadrature formulas. In *Numerische Integration* (G. Hämmerlin, ed.) *Internat. Ser. Numer. Math.* **45** 54–66. Birkhäuser, Basel.

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

Benhenni, K. and Cambanis, S. (1992b). Sampling designs for estimating integrals of stochastic processes using quadratic mean derivatives. In *Approximation Theory* (G. A. Anastassiou, ed.) 93–123. Dekker, New York.

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

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.

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

Golubev, G. K. and Nussbaum, M. (1990). A risk bound in Sobolev class regression. *Ann. Statist.* **18** 758–778.

HÁJEK, J. and KIMELDORF, G. S. (1974). Regression designs in autoregressive stochastic processes. *Ann. Statist.* **2** 520–527.

ISTAS, J. and LAREDO, C. (1994). Estimation d'intégrales de processus aléatoires à partir d'observations discrétisées. *C. R. Acad. Sci. Paris Sér. I Math.* **319** 85–88.

LASINGER, R. (1993). Integration of covariance kernels and stationarity. *Stochastic Process. Appl.* **45** 309–318.

MICCHELLI, C. A. and WAHBA, G. (1981). Design problems for optimal surface interpolation. In *Approximation Theory and Applications* (Z. Ziegler, ed.) 329–347. Academic Press, New York.

MITCHELL, T., MORRIS, M. and YLVISAKER, D. (1990). Existence of smoothed stationary processes on an interval. *Stochastic Process. Appl.* **35** 109–119.

MÜLLER-GRONBACH, T. (1996). Optimal designs for approximating the path of a stochastic process. *J. Statist. Plann. Inference* **49** 371–385.

NIKOLSKIJ, S. M. (1950). On the problem of approximation estimate by quadrature formulas (in Russian). *Uspekhi Mat. Nauk* **5** 165–177.

PARZEN, E. (1967). Statistical inference on time series by Hilbert space methods. I. In *Time Series Analysis Papers* (E. Parzen, ed.) 251–382. Holden-Day, San Francisco.

PASKOV, S. H. (1993). Average case complexity of multivariate integration for smooth functions. *J. Complexity* **9** 291–312.

PLASKOTA, L. (1992). Function approximation and integration on the Wiener space with noisy data. *J. Complexity* **8** 301–323.

RITTER, K. (1996). Almost optimal differentiation using noisy data. *J. Approx. Theory*. **86** 293–309.

RITTER, K. and WASILKOWSKI, G. W. (1996). Integration and L_2-approximation: average case setting with isotropic Wiener measure for smooth functions. *Rocky Mountain J. Math.* **26**. To appear.

RITTER, K., WASILKOWSKI, G. W. and WOŹNIAKOWSKI, H. (1993). On multivariate integration for stochastic processes. In *Numerical Integration* (H. Brass and G. Hämmerlin, eds.) *Internat. Ser. Numer. Math.* **112** 331–347. Birkhäuser, Basel.

RITTER, K., WASILKOWSKI, G. W. and WOŹNIAKOWSKI, H. (1995). Multivariate integration and approximation for random fields satisfying Sacks–Ylvisaker conditions. *Ann. Appl. Probab.* **5** 518–540.

SACKS, J. and YLVISAKER, D. (1966). Designs for regression with correlated errors. *Ann. Math. Statist.* **37** 68–89.

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

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

SACKS, J. and YLVISAKER, D. (1970b). Statistical design and integral approximation. In *Proc. 12th Biennial Seminar Canad. Math. Congress.* (R. Pyke, ed.) 115–136. Canad. Math. Soc., Montreal.

SPECKMAN, P. (1979). L_p approximation of autoregressive Gaussian processes. Technical report, Dept. Statistics, Univ. Oregon, Eugene.

SPECKMAN, P. (1985). Spline smoothing and optimal rates of convergence in nonparametric regression models. *Ann. Statist.* **13** 970–983.

SU, Y. and CAMBANIS, S. (1993). Sampling designs for estimation of a random process. *Stochastic Process. Appl.* **46** 47–89.

SULDIN, A. V. (1959). Wiener measure and its applications to approximation methods. I (in Russian). *Izv. Vyssh. Uchebn. Zaved. Mat.* **13** 145–158.

SULDIN, A. V. (1960). Wiener measure and its applications to approximation methods. II (in Russian). *Izv. Vyssh. Uchebn. Zaved. Mat.* **18** 165–179.

TRAUB, J. F., WASILKOWSKI, G. W. and WOŹNIAKOWSKI, H. (1988). *Information-Based Complexity*. Academic Press, New York.

WAHBA, G. (1971). On the regression design problem of Sacks and Ylvisaker. *Ann. Math. Statist.* **42** 1035–1043.

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

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

WASILKOWSKI, G. W. (1993). Integration and approximation of multivariate functions: average case complexity with isotropic Wiener measure. *Bull. Amer. Math. Soc.* (*N.S.*) **28** 308–314. Full version (1994) *J. Approx. Theory* **77** 212–227.

WEBA, M. (1991). Quadrature of smooth stochastic processes. *Probab. Theory Related Fields* **87** 333–347.

WEBA, M. (1995). Cubature of random fields by product-type integration rules. *Comput. Math. Appl.* **30** 229–234.

WITTWER, G. (1976). Versuchsplanung im Sinne von Sacks–Ylvisaker für Vektorprozesse. *Math. Operationsfersch Ser. Statist.* **7** 95–105.

WITTWER, G. (1978). Über asymptotisch optimale Versuchsplanung im Sinne von Sacks–Ylvisaker. *Math. Operationsforsch Ser. Statist.* **9** 61–71.

WOŹNIAKOWSKI, H. (1991). Average case complexity of multivariate integration. *Bull. Amer. Math. Soc.* (*N.S.*) **24** 185–194.

WOŹNIAKOWSKI, H. (1992). Average case complexity of linear multivariate problems. I: theory; II: applications. *J. Complexity* **8** 337–372, 373–392.

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

ZHENSYKBAEV, A. A. (1983). Extremality of monosplines of minimal deficiency. *Math. USSR-Izv.* **21** 461–482.

MATHEMATISCHES INSTITUT
UNIVERSITÄT ERLANGEN-NÜRNBERG
BISMARCKSTRASSE 1 1/2
91054 ERLANGEN
GERMANY
E-MAIL: ritter@mi.uni-erlangen.de