ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 26, Number 4, Fall 1996

INTEGRATION AND L_2-APPROXIMATION: AVERAGE CASE SETTING WITH ISOTROPIC WIENER MEASURE FOR SMOOTH FUNCTIONS

KLAUS RITTER AND GRZEGORZ W. WASILKOWSKI

ABSTRACT. We propose isotropic probability measures defined on classes of smooth multivariate functions. These provide a natural extension of the classical isotropic Wiener measure to multivariate functions from C^{2r}. We show that, in the corresponding average case setting, the minimal errors of algorithms that use n function values are $\Theta(n^{-(d+4r+1)/(2d)})$ and $\Theta(n^{-(4r+1)/(2d)})$ for the integration and L_2-approximation problems, respectively. Here d is the number of variables of the corresponding class of functions. This means that the minimal average errors depend essentially on the number d of variables. In particular, for d large relative to r, the L_2-approximation problem is intractable. The integration and L_2-approximation problems have been recently studied with measures whose covariance kernels are tensor products. The results for these measures and for isotropic measures differ significantly.

1. Introduction. We study the integration and L_2-approximation problems for multivariate functions f. For the integration problem, we want to approximate the integral of f, and for the function approximation problem, we want to recover f with respect to the L_2-norm. For both problems, we want to determine methods with minimal error among all methods that use n function values. Moreover, we want to know how these errors depend on the number n of evaluations, on the number d of variables of f, and on regularity of f.

Both problems have been extensively studied in the literature, see, e.g., [**15, 24, 25**] for hundreds of references. However, they are mainly addressed in the worst case setting with the algorithm cost and error measured by the worst performance with respect to a given class F of functions. Depending on the smoothness properties of functions from

Received by the editors on November 15, 1994.
Research of the first author supported in part by the Deutsche Forschungsgemeinschaft (DFG).
Research of the second author supported in part by the National Science Foundation under grant CCR-91-14042.

F, the minimal worst case errors strongly or only mildly depend on the dimensionality parameter d.

For instance, if F consists of functions with all partial derivatives of order up to r bounded by 1, both integration and approximation problems are intractable (prohibitively expensive) or even unsolvable since then the minimal worst case errors are proportional to $n^{-r/d}$. This is well known, see, e.g., [**15**, p. 36]. However, if F consists of functions with bounded mixed rth derivatives, the minimal worst case errors equal $\Theta(n^{-r}(\ln n)^{(d-1)/2})$, see [**3, 7**], and $O(n^{-r}(\ln n)^{(r+1)(d-1)})$, see [**22, 23**], for integration and L_2-approximation, respectively. This means that now the dependence on d is only through the exponent in $\ln n$ and a constant in the Θ- or O-notation. However, for $r = 0$, both problems are still unsolvable.

It is therefore important to see how difficult the problems are in an average case setting. In the average case setting, the class F is equipped with a probability measure μ, and the error of an algorithm is defined by its expectation with respect to μ. Like in the worst case setting where optimality of algorithms and minimal worst case errors depend on the properties of F, in the average case setting optimality of algorithms and minimal average errors depend on μ.

Recently, the average case setting for integration and L_2-approximation has been studied in [**17, 32, 33**] assuming that μ is the r-folded Wiener sheet measure. They proved that the minimal average errors equal $\Theta(n^{-r-1}(\ln n)^{(d-1)/2})$ and $\Theta(n^{-r-1/2}(\ln n)^{(d-1)(r+1)})$ for integration and L_2-approximation, respectively.

Wiener sheet measures are Gaussian measures with the covariance kernel being a tensor product of scalar covariance kernels. Hence, the mild dependence of minimal errors on d could be attributed to the tensor product properties of μ. Actually, similar bounds hold for a larger class of probability measures that are tensor product Gaussian, see [**18, 19**]. Therefore, it is important to see how the minimal average errors depend on d when μ does not have a tensor product form.

A classical example of a nontensor product measure is provided by the *isotropic* Wiener measure. Integration and L_2-approximation with such μ have been considered recently in [**31**]. It turns out that the minimal average errors for integration and L_2-approximation equal $\Theta(n^{-1/2-1/(2d)})$ and $\Theta(n^{-1/(2d)})$. Thus, they depend essentially on

d especially for L_2-approximation.

We think that the isotropicity is an important property, at least for a number of practical problems. However, the isotropic Wiener measure is concentrated on continuous functions for which, with probability one, the derivative does not exist at any point. Hence, it is not suitable for studying problems defined over classes of smooth functions. To remedy this, we propose a new measure that is a natural extension of the isotropic Wiener measure. This is an isotropic Gaussian measure and is concentrated on functions f with continuous partial derivatives of order up to $2r$.

We show that the minimal average errors for this measure are equal to $\Theta(n^{-(d+4r+1)/(2d)})$ for integration and $\Theta(n^{-(4r+1)/(2d)})$ for L_2-approximation. Clearly, these bounds depend essentially on d. The results concerning Wiener sheet measures and our results indicate the great difference between the average case settings with both measures. Of course, this difference occurs only for multivariate problems since for $d = 1$ both measures coincide.

Measures on spaces of smooth functions are often obtained from measures on spaces of irregular functions by some kind of smoothing. For instance, the r-folded Wiener sheet measure is obtained from the Wiener sheet measure by r-fold integration with respect to each variable. In this way the tensor product structure is preserved. In order to obtain an isotropic measure, we apply the rth power of the inverse Laplacian operator to the classical isotropic Wiener measure.

The paper is organized as follows. Section 2 contains the basic definitions and problem formulation. In Section 3 we give the construction of the isotropic measures on classes of smooth functions, and we analyze the corresponding reproducing kernel Hilbert space. The error bounds are obtained in Section 4, and the final section contains additional remarks, in particular on almost optimal methods.

2. Average errors: Basic definitions. We consider the following integration and function approximation problems for multivariate functions. Let $F \subset C^{2r}(D)$ be a space of functions with continuous derivatives of order up to $2r$. Since we are interested in isotropic measures, we take

$$D = \{x \in \mathbf{R}^d : |x| \leq 1\}$$

as the unit ball with respect to the Euclidean norm $|x| = (x_1^2 + \cdots + x_d^2)^{1/2}$. The space F is equipped with the norm $||f|| = \max_\alpha ||f^{(\alpha)}||_\infty$, where the maximum is with respect to all multi-indices $\alpha = [\alpha_1, \ldots, \alpha_d]$ with $\sum_{i=1}^d \alpha_i \leq 2r$.

For every $f \in F$ we want to approximate $S(f)$, where $S : F \to G$ with

$$S(f) = \mathrm{Int}\,(f) = \int_D f(x)\,dx \quad \text{and} \quad G = \mathbf{R}$$

for the *integration problem*, and

$$S(f) = \mathrm{App}_2(f) = f \quad \text{and} \quad G = L_2(D)$$

for the *approximation problem*.

An approximation $U_n(f)$ to $S(f)$ is computed based on *information* $N_n(f)$ that consists of n values of f taken at some points from D,

$$N_n(f) = [f(x_1), \ldots, f(x_n)].$$

Hence,

$$U_n(f) = \phi_n(N_n(f)),$$

where

$$\phi_n : N_n(F) \longrightarrow G$$

is an arbitrary (Borel measurable) mapping; ϕ_n is called an *algorithm* that uses N_n.

In the average case setting, we assume that the space F is endowed with a (Borel) probability measure μ. Then the *average error* of $U_n = \phi_n \circ N_n$ is defined by

$$e^{\mathrm{avg}}\,(U_n, S, \mu) = \left(\int_F ||S(f) - U_n(f)||_G^2 \mu(df) \right)^{1/2}.$$

The nth *minimal average error* is then the minimal error among all methods that use n function evaluations,

$$r_n^{\mathrm{avg}}\,(S, \mu) = \inf_{U_n} e^{\mathrm{avg}}\,(U_n, S, \mu),$$

i.e., minimization is with respect to the mapping ϕ_n as well as to the knots x_i. For a more detailed discussion, see, e.g., [**25**].

We study the asymptotic order of the nth minimal average errors $r_n^{\mathrm{avg}}(S,\mu)$. Furthermore, we determine methods using n function evaluations such that their errors differ from $r_n^{\mathrm{avg}}(S,\mu)$ at most by a multiplicative constant.

3. Isotropic Wiener measure for smooth functions. In this section we provide the definition and basic properties of the measure $\mu = w_r$ studied in this paper.

We begin by recalling the *classical* isotropic Wiener measure w_0, see, e.g., [**1, 5, 12, 14**]. This is the zero mean Gaussian measure on $F_0 = C(D)$ with covariance kernel

$$K_0(x,y) := \int_{F_0} f(x)f(y)w_0(df) = \frac{|x|+|y|-|x-y|}{2}.$$

Since $K_0(Qx,Qy) = K_0(x,y)$ for any orthogonal transform Q on $\mathbf{R}^d$, the measure w_0 is isotropic, i.e., it is invariant with respect to any orthogonal transform of D. Moreover, with probability one, any f from F_0 does not have any derivative.

To introduce an isotropic measure w_r on a class of regular functions, we proceed as follows. Let $\Delta = \sum_{i=1}^d \partial^2/\partial x_i^2$ denote the Laplace operator. For a nonnegative integer r, let

$$F_r = \{f \in C^{2r}(D) : f|_{\partial D} = \cdots = (\Delta^{r-1}f)|_{\partial D} = 0\}.$$

The space F_r equipped with $||f|| = \max_\alpha ||f^{(\alpha)}||_\infty$ is a separable Banach space, and we consider the Borel σ-algebra on F_r.

The operator Δ^r defines a bounded linear injection $F_r \to F_0$. Define

$$T_r : F_0 \to F_r$$

by

$$T_r f = \begin{cases} \Delta^{-r} f & \text{if } f \in \Delta^r(F_r), \\ 0 & \text{otherwise.} \end{cases}$$

Clearly, $T_r(\Delta^r f) = f$ for any $f \in F_r$ and $\Delta^r(T_r f) = f$ for any $f \in \Delta^r(F_r)$.

We need the following result for the Poisson equation in Hölder spaces, see, e.g., [**6**, p. 99]. Let $C^{k,\lambda}(D)$ denote the Banach space of functions on D whose kth derivatives satisfy a Hölder condition with exponent λ.

Proposition 1. *The Laplace operator defines an isomorphism* $\{f \in C^{k+2,\lambda}(D) : f|_{\partial D} = 0\} \to C^{k,\lambda}(D)$, *if* $k \in \mathbf{N}_0$ *and* $0 < \lambda < 1$.

A measurable mapping $F_0 \to F_r$ is called weakly measurable linear operator, if it is linear on a measurable linear subspace $V \subset F_0$ with $w_0(V) = 1$, see [**8**].

Lemma 1. T_r *is a weakly measurable linear operator and* $w_0(\Delta^r(F_r)) = 1$.

Proof. Due to a theorem by Kuratowski, see [**26**, p. 5], $\Delta^r(A) \subset F_0$ is measurable for any measurable $A \subset F_r$. This implies the measurability of T_r. Clearly, T_r is linear on $\Delta^r(F_r)$.

Observe that $C^{0,\lambda}(D) \subset F_0$ is measurable. We have $w_0(C^{0,\lambda}(D)) = 1$ if and only if $0 \le \lambda < 1/2$, see [**1**, p. 202] and $C^{0,\lambda}(D) \subset \Delta^r(F_r)$ for $0 < \lambda \le 1$ due to Proposition 1. Therefore, $w_0(\Delta^r(F_r)) = 1$. □

We study measures w_r which are obtained from the classical isotropic Wiener measure by smoothing with some power of Δ^{-1}.

Theorem 1. *Let*

$$w_r = T_r w_0,$$

i.e., $w_r(A) = w_0(T_r^{-1}(A)) = w_0(\Delta^r(A))$ *for any measurable* $A \subset F_r$. *Then* w_r *is an isotropic zero mean Gaussian measure on* F_r.

Proof. Lemma 1 implies that w_r is Gaussian with zero mean, see [**8, 10, 27**]. The measure w_0 is isotropic, and Δ^r commutes with orthogonal transforms of D. Therefore, w_r is isotropic, too. □

It is well known that the reproducing kernel Hilbert space H_μ generated by the covariance kernel of a measure μ plays an important role in analysis of average errors. For instance, for the integration problem,

nth minimal average errors equal the nth minimal worst case errors:

$$
\begin{aligned}
r_n^{\text{avg}}(\text{Int},\mu) &= r_n^{\text{wor}}(\text{Int},H_\mu) \\
&:= \inf_{N_n}\inf_{\phi_n}\sup_{||h||_\mu\le 1}|\text{Int}(h)-\phi_n(N_n(h))| \\
&= \inf_{N_n}\sup_{||h||_\mu\le 1, N_n,(h)=0}\text{Int}(h).
\end{aligned} \tag{1}
$$

This property has been used in many papers, see e.g., [**9, 11, 13, 16, 17, 18, 19, 20, 25, 28, 29, 31, 33, 35**]. For the L_2 approximation problem we only have the inequality

$$
\begin{aligned}
a_d^{-1/2}\cdot r_n^{\text{avg}}(\text{App}_2,\mu) &\le r_n^{\text{wor}}(\text{App}_\infty,H_\mu) \\
&:= \inf_{N_n}\inf_{\phi_n}\sup_{||h||_\mu\le 1}||h-\phi_n(N_n(h))||_\infty,
\end{aligned} \tag{2}
$$

where a_d denotes the volume of the unit ball $D\subset \mathbf{R}^d$, see [**33**]. Here $||\cdot||_\mu$ denotes the norm in H_μ. Therefore, in the following subsection we provide some characterization of the Hilbert spaces which are generated by the covariance kernels of the measures w_r.

3.1. *The reproducing kernel Hilbert space.* We begin by recalling some basic properties of reproducing kernel Hilbert spaces generated by Gaussian measures, see, e.g., [**2, 26, 29**].

Let μ be a zero mean Gaussian measure defined on a space $F\subset C^{2r}(D)$. The covariance kernel of μ is denoted by K_μ. Then the corresponding reproducing kernel Hilbert space $H_\mu\subset F$ is the space generated by finite linear combinations of $K_\mu(\cdot,y)$ for $y\in D$ with the inner product $\langle\cdot,\cdot\rangle_\mu$ defined by

$$\langle K_\mu(\cdot,x),K_\mu(\cdot,y)\rangle_\mu = K_\mu(x,y).$$

Hence, K_μ is the reproducing kernel of H_μ. Moreover, for a complete orthonormal system $\{h_i\}_i$ in H_μ which is orthogonal in $L_2(D)$, we have

$$K_\mu(x,y)=\sum_{i=1}^{\infty}h_i(x)\cdot h_i(y)$$

and

$$f(x) = \sum_{i=1}^{\infty} \xi_i(f) \cdot h_i(x). \tag{3}$$

Here

$$\xi_i(f) = ||h_i||_{L^2(D)}^{-2} \int_D f(y) h_i(y)\, dy,$$

and convergence in (3) is understood in mean square sense with respect to μ. Observe that $\{\xi_i\}_i$ forms a sequence of independent random variables with standard normal distribution.

Now let $K_r = K_{w_r}$ be the covariance kernel of w_r, and let $H_r = H_{w_r}$ be the corresponding reproducing kernel Hilbert space. The norm in H_r is denoted by $|| \cdot ||_r$.

Lemma 2. *$H_0 \subset \Delta^r(F_r)$ and $H_r = \{\Delta^{-r} h : h \in H_0\}$. Moreover,*

$$\|h\|_r = \|\Delta^r h\|_0$$

for any $h \in H_r$.

Proof. Since $w_0(\Delta^r(F_r)) = 1$, we obtain $H_0 \subset \Delta^r(F_r)$, see [**8**]. The representation theorem for weakly measurable linear operators [**8**] and (3), with $\mu = w_0$, imply

$$T_r f(x) = \sum_{i=1}^{\infty} \xi_i(f) \cdot T_r h_i(x) = \sum_{i=1}^{\infty} \xi_i(f) \cdot \Delta^{-r} h_i(x)$$

with convergence in the mean square sense. Therefore,

$$\begin{aligned} K_r(x,y) &= \int_{F_0} T_r f(x) T_r f(y) w_0(df) \\ &= \sum_{i=1}^{\infty} \Delta^{-r} h_i(x) \cdot \Delta^{-r} h_i(y). \end{aligned}$$

Clearly, $H := \{\Delta^{-r} h : h \in H_0\}$, equipped with the scalar product $\langle f, g \rangle_H = \langle \Delta^r f, \Delta^r g \rangle_0$, is a Hilbert space. Moreover, $K_r(\cdot, y) \in H$.

The measurability of $h \mapsto h(y) = T_r(\Delta^r h)(y)$ implies continuity, see [**4, 8**], and therefore H has a reproducing kernel. From

$$h(y) = \sum_{i=1}^{\infty} \langle h, \Delta^{-r} h_i \rangle_H \cdot \Delta^{-r} h_i(y) = \langle h, K_r(\cdot, y) \rangle_H$$

we get $H_r = H$ and $||h||_r^2 = \langle h, h \rangle_H$. □

In Lemmas 3 and 4, we determine spaces X and Y with $X \subset H_r \subset Y$, which are suitable for a worst case analysis of the integration and approximation problems.

Consider the kernel K_0 as a function on $\mathbf{R}^d \times \mathbf{R}^d$, and let Φ denote the Hilbert space of functions on $\mathbf{R}^d$ having this reproducing kernel. The following property of Φ is due to [**14**] for odd d and [**5**] for arbitrary d. Let $\| \cdot \|_\Phi$ denote the norm on Φ. In what follows, we write c to denote positive (perhaps different) constants which may only depend on d. Then

$$\{\varphi \in C_0^\infty(\mathbf{R}^d) : \varphi(0) = 0\} \subset \Phi \tag{4}$$

and

$$||\varphi||_\Phi = c \cdot \left(\int_{\mathbf{R}^d} |\Delta^{(d+1)/4} \varphi(y)|^2 \, dy \right)^{1/2}$$

on this subspace. For $d+1$ not divisible by 4, $\Delta^{(d+1)/4}$ is understood in the generalized sense, see, e.g., [**21**].

Due to [**2**] we have

$$H_0 = \{\varphi|_D : \varphi \in \Phi\} \tag{5}$$

and

$$\|h\|_0 = \min\{\|\varphi\|_\Phi : \varphi \in \Phi, \varphi|_D = h\} \tag{6}$$

for any $h \in H_0$.

Lemma 3. *$H_r \subset C^{2r,1/2}(D)$ and the embedding is continuous.*

Proof. Let $\langle \cdot, \cdot \rangle$ denote the Euclidean scalar product in $\mathbf{R}^d$, and let

$$q(x,y) = \cos(2\pi\langle x,y\rangle) + \sin(2\pi\langle x,y\rangle) - 1.$$

Then $q(\cdot,y)/|y|^{(d+1)/2} \in L_2(\mathbf{R}^d)$ and $Q : L_2(\mathbf{R}^d) \to \Phi$, given by

$$(Qg)(x) = c \cdot \int_{\mathbf{R}_d} q(x,y)/|y|^{(d+1)/2} \cdot g(y)\, dy$$

is an isometric isomorphism, see [**5**]. Observe that

$$\begin{aligned} |q(x+u,y) - q(x,y)| &= |\cos(2\pi\langle x+u,y\rangle) - \cos(2\pi\langle x,y\rangle) \\ &\qquad + \sin(2\pi\langle x+u,y\rangle) - \sin(2\pi\langle x,y\rangle)| \\ &\le c \cdot \min(|\langle u,y\rangle|, 1). \end{aligned}$$

This yields

$$\begin{aligned} &\int_{\mathbf{R}^d} (q(x+u,y) - q(x,y))^2/|y|^{d+1}\, dy \\ &\qquad \le c \cdot \left(|u|^2 \cdot \int_{\{|y| \le 1/|u|\}} |y|^{-d+1}\, dy + \int_{\{|y| \ge 1/|u|\}} |y|^{-d-1}\, dy \right) \\ &\qquad = c \cdot |u| \end{aligned}$$

and

$$\begin{aligned} &|(Qg)(x+u) - (Qg)(x)| \\ &\qquad \le c \cdot \int_{\mathbf{R}^d} |q(x+u,y) - q(x,y)|/|y|^{(d+1)/2} \cdot |g(y)|\, dy \\ &\qquad \le c \cdot \|g\|_{L_2(\mathbf{R}^d)} \cdot |u|^{1/2}. \end{aligned}$$

Using (5) we see that any function in H_0 is Hölder continuous with exponent $1/2$. Moreover, the closed graph theorem implies the continuity of the respective embedding. This proves the statement for $r = 0$ and, together with Proposition 1 and Lemma 2, for $r > 0$. □

Lemma 4. $\{\varphi|_D : \varphi \in C^\infty(\mathbf{R}^d),\ \operatorname{supp} \varphi \subset \operatorname{int} D\backslash\{0\}\} \subset H_r$, *and the norm on this subspace satisfies*

$$\|\varphi|_D\|_r \le \|\Delta^r \varphi\|_\Phi.$$

Proof. If $\varphi \in C^\infty(\mathbf{R}^d)$ satisfies $\operatorname{supp}\varphi \subset \operatorname{int}(D)\backslash\{0\}$, then $\Delta^r\varphi \in \Phi$ and therefore $\Delta^r\varphi|_D \in H_0$, see (4) and (5). We obtain $\varphi|_D \in H_r$ because of Lemma 2. Using (6), the estimate on the norm follows. □

4. Average error bounds.

Theorem 2.

$$r_n^{\text{avg}}(\text{App}_2, w_r) = \Theta(n^{-(4r+1)/(2d)})$$

and

$$r_n^{\text{avg}}(\text{Int}, w_r) = \Theta(n^{-(d+4r+1)/(2d)}).$$

Proof. It is known, see [**31**], that

$$r_n^{\text{avg}}(\text{App}_2, w_r) \geq n^{1/2} \cdot r_{2n}^{\text{avg}}(\text{Int}, w_r).$$

Hence, to prove the theorem it is enough to show that $r_n^{\text{avg}}(\text{App}_2, w_r)$ is bounded from above by $O(n^{-(4r+1)/(2d)})$ and that $r_n^{\text{avg}}(\text{Int}, w_r)$ is bounded from below by $\Omega(n^{-(d+4r+1)/(2d)})$.

The upper bound follows from

$$\begin{aligned} r_n^{\text{avg}}(\text{App}_2, w_r) &\leq r_n^{\text{wor}}(\text{App}_\infty, H_r) \\ &= O(r_n^{\text{wor}}(\text{App}_\infty, C^{2r,1/2}(D))) \\ &= \Theta(n^{-(4r+1)/(2d)}). \end{aligned}$$

Here the inequality is due to [**33**], see (2). The first equality follows from Lemma 3, and the second equality is due to [**15**, p. 34].

To prove the lower bound for the integration problem we proceed as in [**31**]. Consider arbitrary information $N_n(f) = [f(x_1), \ldots, f(x_n)]$ with $x_i \in D$. There are n equal-size cubes $Q_1, \ldots, Q_n \subset D$ with centers $v_1, \ldots, v_n$ that satisfy the following properties. For every j,

$$0 \notin Q_j \quad \text{and} \quad x_i \notin Q_j \quad \text{for all } i, \tag{7}$$

$$Q_i \cap Q_j = \varnothing \quad \text{if } i \neq j, \tag{8}$$

$$\text{the sidelength } \lambda \text{ of } Q \text{ equals } c \cdot n^{-1/d}. \tag{9}$$

As in the previous section, c denotes a positive constant independent of n. Let $\psi \in C^\infty(\mathbf{R})$ be a nonnegative function with $\psi(t) = 0$ if and only if $t \geq 1/16$, and define

$$\varphi_i(x) = \psi(\lambda^{-2} \cdot |x - v_i|^2).$$

Due to (8) and (9), these functions have pairwise disjoint supports, each contained in $Q_j \subset D$. Furthermore, $\varphi_i|_D \in H_r$ holds because of Lemma 4. Consider

$$f(x) = \frac{f_0(x)}{||f_0||_r} \quad \text{with} \quad f_0 = \sum_{i=1}^n \varphi_i|_D.$$

Obviously, $||f||_r = 1$ and, because of (7), $N_n(f) = 0$. Hence, due to (1), we only need to show that $\text{Int}\,(f) \geq c \cdot n^{-(d+4r+1)/(2d)}$. For this end, note that $\text{Int}\,(\varphi_i) = c/n$ and thus

$$\text{Int}\,(f_0) = c. \tag{10}$$

Moreover,

$$\Delta^r \varphi_i(x) = g(\lambda^{-1} \cdot (x - v_i)) \cdot c \cdot n^{2r/d},$$

where $g = \Delta^r \psi(|\cdot|^2)$. Using Lemma 4 and the estimate (3) in [**31**], we obtain

$$\begin{aligned} ||f_0||_r &\leq \left\| \sum_{i=1}^n \Delta^r \varphi_i \right\|_\Phi \\ &\leq c \cdot n^{2r/d} \left\| \sum_{i=1}^n g(\lambda^{-1}(\cdot - v_i)) \right\|_\Phi \\ &\leq c \cdot n^{2r/d} \cdot n^{1/2+1/(2d)}. \end{aligned}$$

This and (10) imply that $\text{Int}\,(f) \geq c \cdot n^{-(d+4r+1)/(2d)}$ as needed. □

5. Remarks.

5.1. *Almost optimal methods.* We discuss linear methods $U_n = \phi_n \circ N_n$ whose average errors are proportional to the minimal errors $r_n^{\text{avg}}(S, w_r)$. Such U_n are said to be almost optimal.

We begin with the L_2-approximation problem. The order of the nth minimal average errors for this problem coincides with the order of the nth minimal worst case errors for L_∞-approximation on the unit ball of the space $C^{2r,1/2}(D)$. This follows from Theorem 2 and [**15**, p. 34]. Moreover, inequality (2), stated in terms of nth minimal errors, also holds for errors of linear methods U_n, see [**33**]. That is,

$$e^{\text{avg}}(U_n, \text{App}_2, w_r) \le e^{\text{wor}}(U_n, \text{App}_\infty, H_r) := \sup_{||f||_r \le 1} ||f - U_n(f)||_\infty$$

for $U_n(f) = \sum_{i=1}^n f(x_i, n) \cdot g_{i,n}$ with knots $x_{i,n} \in D$ and functions $g_{i,n} \in L_2(D)$. Hence, due to Lemma 3, it is sufficient to find linear methods with

$$e^{\text{wor}}(U_n, \text{App}_\infty, C^{2r,1/2}) := \sup_{\|f\|_{C^{2r,1/2}} \le 1} \|f - U_n(f)\|_\infty = \Theta(n^{-(4r+1)/(2d)}).$$

Such methods are known, see, for instance, [**15**, p. 34]. One can take knots $x_{i,n}$ which form a uniform grid in D. The corresponding algorithm ϕ_n computes piecewise polynomials of degree at most $2r$ in each variable which interpolate f on subgrids containing $(2r+1)^d$ points. Hence, we see that a fairly simple method is almost optimal.

For the integration problem, the situation is much more complicated since (almost) optimal selection of sample points $x_{i,n}$ is unknown. We only know (almost) optimal *randomized* methods even though randomization does not help in the average case setting. Indeed, let $U_n = \phi_n \circ N_n$ be a family of (almost) optimal methods for the L_2-approximation problem, e.g., piecewise polynomial interpolation on a uniform grid. Consider

$$\psi_n^{\vec{t}}(M_n^{\vec{t}}(f)) = \text{Int}\,(U_n(f)) + \frac{\lambda}{n} \sum_{i=1}^n (f - U_n(f))(t_i). \tag{11}$$

Here λ is the volume of D and $M_n^{\vec{t}}(f)$ is the following randomized information. It contains all values of $N_n(f)$ as well as the values $f(t_i)$ for n random points $t_1, \dots, t_n \in D$, each generated independently with uniform distribution on D. Then the average error of such a randomized method does not exceed $(\lambda/n)^{1/2} e^{\mathrm{avg}}(U_n, \mathrm{App}_2, w_r)$, see, e.g., [**31**], which proves (almost) optimality of $\psi_n^{\vec{t}} \circ M_n^{\vec{t}}$.

5.2. *Complexity.* Our main results are stated in terms of minimal average errors among all algorithms that use n function values. Obviously, the same results hold if the permissible information operators include evaluation of derivatives. Moreover, we could allow adaptive selection of sample points, both deterministic and randomized.

Theorem 2 can also be used to determine the average complexity of both integration and L_2-approximation problems. For a detailed discussion of complexity, see, e.g., [**25**]. We only recall that the average ε-complexity of a problem, $\mathrm{comp}^{\mathrm{avg}}(\varepsilon, S, \mu)$, is the minimal expected cost among all methods $\phi \circ N$ whose average error does not exceed ε. Here we allow any information N consisting of a number of function (and/or derivatives) values; both the number of evaluations and the sample points can be chosen randomly and/or adaptively. Also, ϕ can be a random mapping. We assume that the cost of one function evaluation is c, whereas basic arithmetic operations have unit cost $1 \le c$.

From [**30**] we know that for Gaussian measures μ, $\mathrm{comp}^{\mathrm{avg}}(\varepsilon, S, \mu)$ is proportional to $n(\varepsilon) = \min\{l : r_l^{\mathrm{avg}}(S, \mu) \le \varepsilon\}$ if the nth minimal average error $r_n^{\mathrm{avg}}(S, \mu)$ is a semiconvex function of n. From Theorem 2 we see that, for $\mu = w_r$ and $S \in \{\mathrm{Int}, \mathrm{App}_2\}$ the corresponding nth minimal average errors are semi-convex. Hence,

$$\mathrm{comp}^{\mathrm{avg}}(\varepsilon, \mathrm{App}_2, w_r) = \Theta(\varepsilon^{-2d/(4r+1)})$$

and

$$\mathrm{comp}^{\mathrm{avg}}(\varepsilon, \mathrm{Int}, w_r) = \Theta(\varepsilon^{-2d/(d+4r+1)}).$$

Moreover, nonadaptive linear methods (as discussed in the previous subsection) are almost optimal also from the complexity point of view.

5.3. *Weighted integration.* Suppose we want to approximate a weighted integral

$$S(f) = \mathrm{Int}_\rho(f) = \int_D f(x) \cdot \rho(x)\, dx$$

for a given nonzero weight function $\rho \in C(D)$. The corresponding nth minimal average errors are of order $n^{-(d+4r+1)/(2d)}$ as in the case $\rho \equiv 1$. The lower bound follows as in the proof of Theorem 2 with the only difference being that the cubes Q_i are from a ball on which $|\rho(x)|$ is bounded from below by a positive constant. The upper bound is achieved analogously to (11): Apply the classical Monte Carlo method for Int_ρ to the function $f - U_n(f)$ and add $\mathrm{Int}_\rho(U_n(f))$.

5.4. *Boundary conditions.* Due to the construction of w_r, we study functions which vanish, together with some derivatives, at the boundary of D. This restriction can be removed by introducing random boundary conditions. We sketch this modification for the case $r = 1$.

The operator Δ and the restriction on the boundary ∂D define a bounded linear injection $C^2(D) \to C(D) \times C^2(\partial D)$. Let T denote the inverse of this operator, extended by zero to be defined on the whole space $C(D) \times C^2(\partial D)$. Moreover, let ν denote a zero mean isotropic Gaussian measure on $C^2(\partial D)$, such that the corresponding reproducing kernel Hilbert space is a subset of $C^{2,1/2}(\partial D)$. See [**29, 34**] for a construction of such measures.

Let $C(D) \times C^2(\partial D)$ be equipped with the product measure $w_0 \otimes \nu$. Proceeding as in Section 3, one can show that the image measure $\mu = T(w_0 \otimes \nu)$ is a zero mean isotropic Gaussian measure on $C^2(D)$. The reproducing kernel Hilbert space H_μ is contained in $C^{2,1/2}(D)$. Therefore, the upper bounds from Theorem 2 (and of course the lower bounds) are also valid with the measure μ.

Acknowledgment. We thank Erich Novak for valuable remarks and comments.

REFERENCES

1. R.J. Adler, *The geometry of random fields*, Wiley, New York, 1981.

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

3. V.A. Bykovskij, *On exact order of optimal quadrature formulas for spaces of functions with bounded mixed derivatives*, Dalnevostochnoi Center of Academy of Sciences, Vladivostok, USSR, 1985 (in Russian).

4. J.P.R. Christensen, *Borel structures in groups and semi-groups*, Math. Scand. **28** (1971), 124–128.

5. Z. Ciesielski, *On Lévy's Brownian motion with several-dimensional time*, in *Probability-Winter School*, Lecture Notes in Math. **472** (1975), 29–56.

6. Y.V. Egorov, M.A. Shubin, eds., *Partial differential equations* I, Encyclopedia of Mathematical Sciences, Vol. 30, Springer-Verlag, Berlin, 1992.

7. K.K. Frolov, *An upper bound of the error of quadratures of some classes of functions*, Dokl. Akad. Nauk SSSR **231** (1976), 818–821 (in Russian).

8. M. Gattinger, *Representation theorems for operators and measures on abstract Wiener spaces*, in *Measure theory applications to stochastic analysis*, Lecture Notes in Math. **695** (1978), 239–245.

9. G.S. Kimeldorf and G. Wahba, *Spline functions and stochastic processes*, Sankhya, Sec. A, **32** (1970), 173–180.

10. M.A. Kon, K. Ritter and A.G. Werschulz, *On the average case solvability of ill-posed problems*, J. Complexity **7** (1991), 220–224.

11. F.M. Larkin, *Gaussian measure in Hilbert space and application in numerical analysis*, Rocky Mountain J. Math. **2** (1972), 379–421.

12. P. Lévy, *Processus stochastique et mouvement Brownien*, Gauthier-Villars, Paris, 1948.

13. C.A. Micchelli and G. Wahba, *Design problems for optimal surface interpolation*, in *Approximation theory and applications* (Z. Ziegler, ed.), Academic Press, New York, 1981.

14. G.M. Molchan, *On some problems concerning Brownian motion in Lévy's sense*, Theory of Prob. Appl. **12** (1967), 747–755.

15. E. Novak, *Deterministic and stochastic error bounds in numerical analysis*, Lecture Notes in Math. **1349**, Springer-Verlag, Berlin, 1988.

16. E. Parzen, *An approach to time series analysis*, Ann. Math. Stat. **32** (1962), 951–989.

17. S.H. Paskov, *Average case complexity of multivariate integration for smooth functions*, J. Complexity **9**, 291–312.

18. K. Ritter, G.W. Wasilkowski and H. Woźniakowski, *On multivariate integration for stochastic processes*, in *Numerical integration* (H. Braßand and G. Hämmerlin, eds.), Birkhäuser Verlag, Basel, 1993, 331–347.

19. ———, *Multivariate integration and approximation for random fields satisfying Sacks-Ylvisaker conditions*, Ann. Appl. Prob. **5** (1995), 518–540.

20. J. Sacks and D. Ylvisaker, *Statistical design and integral approximation*, in Proc. 12th Bienn. Semin. Can. Math. Congr. (R. Pyke, ed.), Canad. Math. Soc., Montreal, 1970, 115–136.

21. E.M. Stein, *Singular integrals and differentiability properties of functions*, Princeton University Press, Princeton, New Jersey, 1970.

22. V.N. Temlyakov, *Approximate recovery of periodic functions of several variables*, Math. USSR-Sb. **56** (1987), 249–261.

23. ———, *On approximate recovery of functions with bounded mixed derivative*, J. Complexity **9** (1993), 41–59.

24. ———, *Approximation of periodic functions*, Nova Science, New York, 1993.

25. J.F. Traub, G.W. Wasilkowski and H. Woźniakowski, *Information-based complexity*, Academic Press, New York, 1988.

26. N.N. Vakhania, V.I. Tarieladze and S.A. Chobanyan, *Probability distributions on Banach spaces*, Reidel, Dordrecht, 1987.

27. N.N. Vakhania, *Gaussian mean boundedness of densely defined linear operators*, J. Complexity **7** (1991), 225–231.

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

29. ———, *Spline models for observational data*, SIAM, Philadelphia, 1990.

30. G.W. Wasilkowski, *Information of varying cardinality*, J. Complexity **2** (1986), 204–228.

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

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

33. ———, *Average case complexity of linear multivariate problems*, Part 1: *Theory*, Part 2: *Applications*, J. Complexity **8** (1992), 337–372, 373–392.

34. M.I. Yadrenko, *Spectral theory of random fields*, Optimization Software, New York, 1983.

35. D. Ylvisaker, *Designs on random fields*, in *A survey of statistical design and linear models* (J. Srivastava, ed.), North-Holland, Amsterdam, 1975.

Mathematisches Institut, Universität Erlangen-Nürnberg, Bismarckstr. 1 1/2, 91054 Erlangen, Germany
E-mail: ritter@mi.uni-erlangen.de

Department of Computer Science, University of Kentucky, Lexington, KY 40506-0046
E-mail: greg@cs.engr.uky.edu