Complexity of Deterministic and Randomized methods on Multivariate Integration Problem for the Classes $H_p^\Lambda(I^d)$

Fang Gensun

Department of Mathematics, Beijing Normal University,
Beijing 100875, China
E-mail: fanggs@bnu.edu.cn

Ye Peixin

School of Mathematical Sciences and LPMC
Nankai University
Tianjin 300071, China
E-mail: yepx@nankai.edu.cn

*Project supported by the Natural Science Foundation of China (Grant No. 10071006) and the Program of “One Hundred Distinguished Chinese Scientists” of the Chinese Academy of Sciences."
We introduce a new classes of multivariate functions $H^\Lambda_p(I^d)$, which include some classical classes of functions such as classical, anisotropic Sobolev and Hölder-Nikolskii class as its special cases or its subclasses in the sense of continuously imbedding, and study the computational complexity of the integration problem with respect to these classes. We prove that when $p > 1$ the $\varepsilon$-complexity in the randomized setting and in the average case setting are significantly smaller than that of in the worst case deterministic setting.

*Keywords:* Computational complexity; worst case; average case; randomized setting; generalized Hölder-Nikolskii class

1. Introduction

It is known that the most problems of scientific computation must be solved with uncertainty. We often have only partial and contaminated information which is priced about such problems. The computational complexity of a problem is the minimal cost among all algorithms of computing an approximating solution with error at most $\varepsilon$, it is the intrinsic difficulty as measured by the minimal computational resources required for its approximation solution, and is a fundamental invariant of computer science.

In this paper, we study the computational complexity of the high-dimensional integration problem which has attracted much attention due to its applications in computational mathematics [4, 16, 18, 26, 28], in finance [2, 12, 17], in physics and engineering [9, 10] and in statistics [6].

Let $I^d := [0, 1]^d$ and $C(I^d)$ be the space of continuous functions on it, endowed with the supremum norm. Let $F_d$ be a bounded subset of $C(I^d)$. Set

$$S(f) = \int_{I^d} f(x) \, dx, \quad (1)$$

we assume that the function $f$ is unknown, instead, we can compute their *standard* information $N(f)$, consisting only of function values at $n$ points $x_1, \ldots, x_n \in I^d$, that is

$$N(f) = [f(x_1), \ldots, f(x_n)], \quad (2)$$

the points and the number $n$ of them (called the cardinality of $N$) can be selected adaptively or randomly, i.e., for adaptive information $N(f)$, $x_j$'s depending on previously computed values $f(x_1), \ldots, f(x_{j-1})$, and the cardinality $n = n(f)$ varies with $f$ based on the computed values. For the randomized information $N(f)$, the points $x_j$ and the cardinality $n(f)$ may also depend on an outcome of a random process. We want to approximate the integral $S$ by a numerical method $\tilde{S}$ of the form

$$\tilde{S}(f) = \phi(N(f)) = \phi(f(x_1), \ldots, f(x_n)),$$

where $\phi: \mathbb{R}^n \to \mathbb{R}$ is an arbitrary mapping.

Denoted the set of all information operators $N$ by $I_n$ and the set of all numerical method $\tilde{S}$ by $A_n$.

$$I_n = \{ N : F_d \to \mathbb{R}^n, N(f) = (f(a_1), \ldots, f(a_n)), a_i \in I^d, i = 1, \ldots, n \},$$
$A_n = \{ \tilde{S} : F_d \rightarrow \mathbb{R}, \tilde{S} = \phi \circ N \text{ with any } \phi : \mathbb{R}^n \rightarrow \mathbb{R}, \text{ and } N \in I_n \}$. Then we can define the notions of error and cost with respect to various settings.

The worst case error of a deterministic method $\tilde{S}$ on $F_d$ is given by

$$\Delta_{\text{max}}(\tilde{S}, F_d) := \sup_{f \in F_d} |S(f) - \tilde{S}_n(f)|.$$  (3)

The $n$-th minimal worst case deterministic error is defined by

$$e_{\text{wor}}^n(F_d) := \inf_{\tilde{S} \in A_n} \Delta_{\text{max}}(\tilde{S}, F_d).$$ (3)

Assume that $F_d$ is equipped with a probability measure $\mu$. Then the average case error of $\tilde{S}$ on $F_d$ is defined by:

$$\Delta_{\mu}(\tilde{S}, F_d) = \int_{F_d} |S(f) - \tilde{S}(f)| \, d\mu(f),$$

where $\mu$ is a Borel probability measure on $(F_d, \| \cdot \|_{\infty})$.

The $n$-th minimal average case deterministic error on $F_d$ with respect to $\mu$ is

$$e_{\text{avg}}^n(F_d, \mu) = \inf_{\tilde{S} \in A_n} \Delta_{\mu}(\tilde{S}, F_d).$$ (4)

Let $(\Omega, \Sigma, \rho)$ be a probability space, i.e. $\Omega$ is a nonempty set, $\Sigma$ a $\sigma$-algebra of subset of $\Omega$, and $\rho$ a probability measure on $\Omega$. A Monte Carlo method in $A_n$ is a couple

$$U := \left( (\Omega, \Sigma, \rho), (\tilde{S}_\omega)_{\omega \in \Omega} \right)$$
where $\tilde{S}_\omega \in A_n$ for any $\omega \in \Omega$.

The worst case error of randomized algorithm $U$ on $F_d$ is

$$\Delta_{\text{ran}}(U, F_d) = \sup_{f \in F_d} \int_{\Omega} |S(f) - \tilde{S}_\omega(f)| \, d\rho(\omega),$$

where $\underline{\Delta}_{\text{ran}}$ denote the set of all Monte Carlo methods $U$ by $M_n$.

The $n$-th minimal worst case randomized error on $F_d$ is

$$e_{\text{ran}}^n(F_d) = \inf_{U \in M_n} \Delta_{\text{ran}}(U, F_d).$$ (5)

Then we define the $\varepsilon$-complexity of various setting which is the minimal cost of computing the integration within a prescribed accuracy $\varepsilon$, see [15]. Without loss of generality, we measure the cost of a method by the number $n$ of function values which enter into it. Therefore the worst case complexity $\text{comp}_{\text{wor}}^n(F_d)$ can be defined as

$$\text{comp}_{\text{wor}}^n(F_d) = \inf \left\{ n : e_{\text{wor}}^n(F_d) \leq \varepsilon \right\}.$$
Analogously, we can define the quantities $\text{comp}^{\text{ran}}(\varepsilon, F_d)$ and $\text{comp}^{\text{avg}}(\varepsilon, F_d)$ respectively. We are ready to define tractability, see [16, 27], we say that the multivariate integration problem for a sequence $F_d$ of spaces in the worst case setting is tractable, if there exist non-negative constants $a$, $b$, and $c$ such that

$$\text{comp}^{\text{wor}}(\varepsilon, F_d) \leq c d^b \varepsilon^{-a}, \quad d \in \mathbb{N}, \quad \varepsilon \in (0, 1),$$

where the constant $c$ does not depend on $\varepsilon$ and $d$, the infimum of the numbers $a$ and $b$ are called the exponents of tractability. The tractability means that there exists an algorithm whose error is at most $\varepsilon$ and whose standard information cost is bounded by a polynomial in the dimension $d$ and in $\varepsilon^{-1}$. We can define tractability in the average and randomized setting.

Detailed information about the computational complexity may be found in the literatures [24, 23] and the recent paper [28].

In the following, we use asymptotic notation $a_n \ll b_n$ for two positive sequences $a_n$ and $b_n$, which means that there exist some constant $c > 0$ and some $n_0 \in \mathbb{N}$ such that $a_n \leq c b_n$ for all $n \geq n_0$. If $a_n \ll b_n$ and $b_n \ll a_n$, then we write $a_n \asymp b_n$.

Let $L^p(I^d)$, $1 \leq p \leq \infty$ be $p$-th Lebesgue spaces with the usual norms. For $\alpha \in \mathbb{Z}^d_+$, define $|\alpha| := \sum_{j=1}^d \alpha_j$, For the isotropic (classical) Sobolev classes (with the imbedding conditions $kp > d$) which are defined by

$$W^k_p(I^d) := \left\{ f \in L^p(I^d) : \sum_{|\alpha|=k} \| D^\alpha f \|_p \leq 1 \right\}, \quad 1 \leq p \leq \infty,$$

Bakhvalov, see [1], obtained

$$e^\text{wor}_n(W^k_\infty(I^d)) \asymp n^{-\frac{k}{2}},$$

Novak [13, 14] extended this result and proved that

$$e^\text{wor}_n(W^k_p(I^d)) \asymp n^{-\frac{k}{2}}, \quad 1 \leq p < \infty.$$

It is clear that the worst case error in (7) and (8) is unsatisfactory when $d$ is large relative to $k$, so Novak [13, 14] considered its average and stochastic error, and proved

$$e^\text{ran}_n(W^k_p(I^d)) \asymp n^{-\frac{k}{2}-l},$$

$$\sup_\mu e^\text{avg}_n(W^k_p(I^d), \mu) \asymp n^{-\frac{k}{2}-l},$$

where $l = 1/2$ for $p = 2$ and $l = 1 - 1/p$ for $1 \leq p < 2$. We know from these results that when $p > 1$, then the average and stochastic error of suitable methods are essential smaller than the worst case error of arbitrary methods, it is important for the applications, especially if $d$ is much bigger than the smoothness $k$, in this sense we conclude that the isotropic Sobolev classes are suitable for the average case and randomized setting. A problem arises naturally: are there more general classes of functions which is suitable for the average case or randomized...
setting that is, we want to find some classes of functions on which the error bounds of the integration problem in the average case or randomized setting are significantly better than that in the worst case setting.

To state our main ideas and results, we recall the definitions of some fundamental classes of functions.

1) Let $T^d$ be $d$-dimensional torus, the periodic anisotropic Sobolev class $SW_{p,\beta}^k(T^d)$, $k \in \mathbb{R}^d$, $\beta \in \mathbb{R}^d$ consists of all functions $f(x)$ which have the following integral representation for each $j \in \{1, \ldots, d\}$,

$$f(x) = \frac{1}{2\pi} \int_0^{2\pi} \phi_j(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_d)F_k(x_j - y, \beta_j)dy,$$

where

$$F_k(x, \beta) = 1 + 2\sum_{j=1}^\infty j^{-k} \cos(jx - \frac{\beta\pi}{2}), \quad \beta \in \mathbb{R}$$

is the Bernoulli kernel and $\phi_j$ is the fractional partial derivative of order $k_j$ of $f$ in the sense of Weil, see [22].

2) Let $F_d$ be the H"older-Nikolskii class $H_p^k(T^d)$, which is the set of functions $f \in L_p(T^d)$ such that

$$\|f\|_p \leq 1, \quad \|\Delta_h^j(f)\|_p \leq |h|^{\eta(k_j)},$$

where $l_i = [k_j] + 1, j = 1, \ldots, d, [k_j]$ denotes the integer part of $k_j$. It follows from Temlyakov [21, 22] that if $1 \leq p \leq \infty$, and $\eta(k)p > d$, then

$$\inf_{\tilde{S} \in \tilde{A}_n} \Delta_{\max}(\tilde{S}, F_d) \asymp n^{-\frac{\eta(k)}{d}}, \quad (10)$$

where $F_d$ denotes $SW_p^k(T^d)$ or $H_p^k(T^d)$, and $\tilde{A}_n$ is the collection of the simplest linear method, namely

$$\tilde{A}_n = \{ \tilde{S} \in A_n : \tilde{S}(f) = \sum_{i=1}^n c_if(x_i) \}, \quad \text{(11)}$$

and

$$\eta(k) = \frac{d}{\sum_{j=1}^d \frac{1}{k_j}} \quad \text{(12)}$$

is the harmonic average of the numbers $k_j, j = 1, \ldots, d$. It is clear that $\tilde{A}_n$ is a subset of $A_n$. By means of periodization technique, Temlyakov [21] extended the above results to non-periodic class $H_p^k(I^d)$, that is, (10) is still true for $H_p^k(I^d)$, therefore, we conclude that

$$\inf_{\tilde{S} \in \tilde{A}_n} \Delta_{\max}(\tilde{S}, H_p^k(I^d)) \asymp n^{-\frac{\eta(k)}{d}}.$$
We notice that the distributions of smoothness index of the isotropic and anisotropic Sobolev, and Hölder-Nikolskii classes can be viewed as the boundary of a complete set. To see this, we recall some definitions of [3], a set $\Lambda \subset \mathbb{Z}_+^d$ is said to be complete, if it is bounded and satisfies: $\alpha \in \Lambda$ and $\alpha' \leq \alpha$, implies that $\alpha' \in \Lambda$, and the boundary of $\Lambda$ is defined by

$$\partial \Lambda := \{ \alpha : \alpha \notin \Lambda, \text{ and if } \alpha' < \alpha, \text{ then } \alpha' \in \Lambda \},$$

(13)

where $\alpha' < \alpha$ means $\alpha' \leq \alpha$ and $\alpha'_j < \alpha_j$ at least for a $j \in \{1, \ldots, d\}$. It is easy to see that the smoothness index $\{1 \in \mathbb{Z}_+^d : |l| = k\}$ of the classical Sobolev class $W_p^k(I^d)$ is the boundary of complete set $\Lambda_1 := \{1 \in \mathbb{Z}_+^d : |l| < k\}$. The smoothness index $\{k_j e_j : j = 1, \ldots, d\}$ of the anisotropic Hölder-Nikolskii class $H_p^k(I^d)$ is the boundary of complete set $\Lambda_2 := \{1 \in \mathbb{Z}_+^d : l_j < k_j, j = 1, \ldots, d\}$ where $e_j$ denotes the unit vector in the $j$-th coordinate direction.

Based on above observation, we want to define a new class which includes above known classes as its special cases or its subclasses in the sense of continuously imbedding. For an $\alpha \in \mathbb{Z}_+^d$, $t \in \mathbb{R}_+^d$, define

$$\Delta^\alpha_t f, x) := \Delta^\alpha_{t_1} \cdots \Delta^\alpha_{t_d} (f, x),$$

where $\Delta^\alpha_{t_j}$ is the usual $\alpha_j$-th forward difference of step length $t_j$ with respect to $x_j$, $j = 1, \ldots, d$. For

$$I^d(\alpha, h) := \{x : (x_j + s_j \alpha_j)^d \in I^d \text{ for all } s \leq h, s \in \mathbb{R}_+^d\},$$

define

$$\omega_{\alpha}(f, h, I^d_p) := \sup_{0 \leq t \leq h} \|\Delta^\alpha_t(f, \cdot)\|_p(I^d(\alpha, h)),$$

Now we define a new class of functions $H_p^\Lambda(I^d)$ which is called the generalized Hölder-Nikolskii classes of functions

$$H_p^\Lambda(I^d) = \{ f \in L_p(I^d) : \omega_{\alpha}(f, h, I^d_p) \leq h^\alpha, \alpha \in \partial \Lambda \},$$

(14)

where $h^\alpha := \prod_{j=1}^d h_j^{\alpha_j}$. In the next section, we will see that the generalized Sobolev class $W_p^\Lambda(I^d)$ which is defined by

$$W_p^\Lambda(I^d) := \left\{ f \in L_p(I^d) : \sum_{\alpha \in \partial \Lambda} \|\Delta^\alpha f\|_p + \|f\|_p \leq 1 \right\},$$

can be continuously imbedded into the generalized Hölder-Nikolskii class $H_p^\Lambda(I^d)$.

It is known that anisotropic classes are the appropriate setting for describing and analyzing functions with different smoothness properties with respect to different directions and allow one to study systematically linear anisotropic but coercive partial differential problems and also to investigate nonlinear anisotropic boundary problems, see [8]; and generalized Sobolev classes play an important role in finite element analysis and the establishment of general setting for multi-spline approximation, see [3]. According to our definition, the class $H_p^\Lambda(I^d)$ includes some these important known classes as its special cases or its subclasses in the sense of continuously imbedding. Thus the investigation of the integration problem on the classes of functions $H_p^\Lambda(I^d)$
may lead to a more general results than before. Particularly, we will prove that when \( p > 1 \), this class is suitable for average case analysis and randomized setting.

2. Statement of main results

The smoothness of generalized Hölder-Nikolskii class \( H^\Lambda(I^d) \) is slight weaker than that of the Sobolev class, meanwhile unlike the classical and anisotropic sobolev classes, we don’t know the precise distribution of the smoothness index of our new classes of functions, both make this class of functions more complicate to study. In fact there is a trade off between the universality and accuracy. However due to the following imbedding theorem, we can reduce the lower estimate of error bound of the integration problem on our new classes to generalized Sobolev classes and the upper estimates to anisotropic Hölder-Nikolskii classes.

**Theorem 1.** Let \( 1 \leq p \leq \infty \). Then the following continuous imbedding relations hold

\[
W^\Lambda_p(I^d) \hookrightarrow H^\Lambda_p(I^d) \hookrightarrow H^{k_\Lambda}_p(I^d),
\]

where \( k_\Lambda = (k_1, \ldots, k_d) \) satisfies \( k_j e_j \in \partial \Lambda \) for \( j = 1, \ldots, d \).

It is known from Theorem 1 and the Nikolskii imbedding theorem [11] that when \( \eta(k_\Lambda)p > d \) the class \( H^\Lambda_p(I^d) \) can be continuously imbedded into \( C(I^d) \).

In the following, we set \( \bar{k} = \max_{k \in \partial \Lambda} |k| \).

**Theorem 2.** Let \( F_d \) be \( H^\Lambda_p(I^d) \) and \( \eta(k_\Lambda)p > d \). Then

\[
n^{-\frac{d}{\bar{k}}} \ll e^\text{wor}_n(F_d) \ll n^{-\frac{\eta(k_\Lambda)}{\bar{k}}}, \quad \varepsilon^{-\frac{d}{2}} \ll \text{comp}^\text{wor}(\varepsilon, F_d) \ll \varepsilon^{-\frac{d}{\eta(k_\Lambda)}}.
\]

By virtue of Theorem 2, we have immediately

**Corollary 1.** Let \( F_d \) be one of the isotropic classes \( W^k_p(I^d) \) or \( H^k_p(I^d) \) and \( kp > d \). Then

\[
e^\text{wor}_n(F_d) \asymp n^{-\frac{k}{\bar{k}}}, \quad \text{comp}^\text{wor}(\varepsilon, F_d) \asymp \varepsilon^{-\frac{d}{\bar{k}}}.
\]

Corollary 1 was proved by Bakhvalov [1], and Novak [13], see (7) and (8).

**Corollary 2.** Let \( F_d \) be one of the anisotropic classes \( W^k_p(I^d) \) or \( H^k_p(I^d) \), and \( \eta(k)p > d \). Then

\[
e^\text{wor}_n(F_d) \asymp n^{-\frac{\eta(k)}{\bar{k}}}, \quad \text{comp}^\text{wor}(\varepsilon, F_d) \asymp \varepsilon^{-\frac{d}{\eta(k)}}.
\]

**Remark 1.**

a) The upper estimate of Theorem 2 follows from the second imbedding relationship in Theorem 1, and the results of Temlyakov [21, 22], hence, we only have to prove its lower estimate.

b) The upper estimate of Corollary 2 was obtained by Temlyakov in [21, 22], but the lower estimate is new, since the deterministic methods \( A_n \), see (??), which includes linear, nonlinear methods, adaptive methods, and methods with varying cardinality, is wider than that of Temlyakov (the simplest linear methods \( A_n \)), see (11).
Theorem 3. Let $F_d$ be one of the classes $W_p^\Lambda(I^d)$ or $H_p^\Lambda(I^d)$ with $\eta(k_\Lambda)p > d$. Then
\[
\frac{\pi}{2} - l < e_{n}^{\text{ran}}(F_d) < n^{-\frac{\eta(k_\Lambda)}{d}-l}, \quad \frac{\pi}{2} - l < e_{n}^{\text{avg}}(F_d, \mu) < n^{-\frac{\eta(k_\Lambda)}{d}-l}.
\] (17)

It follows from Theorem 3, we get

Corollary 3. Let $F_d$ be one of the isotropic classes $W_p^k(I^d)$ or $H_p^k(I^d)$ with $kp > d$. Then
\[
e_{n}^{\text{ran}}(F_d) \asymp n^{-\frac{k}{d}-l}, \quad \sup_{\mu} e_{n}^{\text{avg}}(F_d, \mu) \asymp n^{-\frac{k}{d}-l}.
\]

Corollary 4. Let $F_d$ be one of anisotropic the classes $W_p^k(I^d)$ or $H_p^k(I^d)$ with $\eta(k)p > d$. Then
\[
e_{n}^{\text{ran}}(F_d) \asymp n^{-\frac{n(k)}{d}-l}, \quad \sup_{\mu} e_{n}^{\text{avg}}(F_d, \mu) \asymp n^{-\frac{n(k)}{d}-l}.
\]

Corollary 4 was proved in [5].

Remark 2. By above Theorems, we conclude that in the randomized setting and average case setting, we have an essential better error bound of the integration problem on the classes $H_p^\Lambda(I^d)$, $p > 1$, than that of in the worst case setting. The most important case is that of moderate smoothness ($\eta(k_\Lambda)$ or $\overline{k}$) and large dimension $d$, as a consequence of imbedding condition $\eta(k_\Lambda)p > d$, $p$ should be appropriately large. Meanwhile, when $1 \leq p < \infty$, this imbedding conditions enable us to obtain at least the convergence rate $n^{-1/p}$ for $e_{n}^{\text{wor}}(H_p^\Lambda(I^d))$, this rate does not depend on the dimension $d$. Thus we can choose an exponent $a = p$ in (6) which satisfies the requirement of tractability for $\varepsilon^{-1}$, however we could not claim that the problem is tractable in this case, since we don’t know enough about the constant $b$ in (6), for the case of $p = \infty$, there is no dependence between $\eta(k_\Lambda)$ and $d$, then we can claim that the integration problem on $H_p^\Lambda(I^d)$ is intractable for a fixed smoothness $k$. More details for this problem see [].

3. The proof of results

Proof of Theorem 1. Let $M_k$ denote the $B$-spline of one variable of order $k$ with knots $0, 1, \ldots, k$ normalized to have integral one on $\mathbb{R}$. For $t > 0$ and $\alpha \in \partial \Lambda$, let
\[
M_\alpha(t, x) := \prod_{j=1}^{d} (t_j)^{-1} M_{\alpha_j}(t_j^{-1} x_j).
\] If $f \in W_p^\Lambda(I^d)$, then it follows from the known formula [19] on $B$-spline that
\[
\Delta_{t}^{\alpha}(f, x) = t^{\alpha} \int_{\mathbb{R}^d} D^{\alpha} f(x + u) M_{\alpha}(t, u) du,
\]
hence
\[ \omega_\alpha(f, \mathbf{h}, I^d)_p \leq \mathbf{h}^\alpha \| D^\alpha f \|_p \leq \mathbf{h}^\alpha, \]
which proves the first imbedding relationship. It follows from the definition of complete set \( \Lambda \) that for any \( 1 = (l_1, \ldots, l_d) \in \Lambda \) implies \( l_j e_j \in \Lambda, \ j = 1, \ldots, d, \) and then it follows from the definition of the boundary \( \partial \Lambda, \) there is a unique vector \( k_\Lambda = (k_1, \ldots, k_d) \) satisfying \( k_j e_j \in \partial \Lambda. \) This combines the definition of the Hölder-Nikolskii class implies the second imbedding relationship. The proof of Theorem 1 is complete.

**Lemma 1.** [14] Let \( F_d \) be a class of bounded functions and assume that \( f_1, \ldots, f_{2n} \) have mutually disjoint supports such that \( S(f_i) \geq \epsilon, \) and \( \sum_{i=1}^{2n} \delta_i f_i \in F_d \) for all \( \delta_i \in \{-1, 1\}, i = 1, \ldots, 2n. \) Then
\[ e_n(\omega_\alpha(F_d)) \geq n \epsilon. \] (18)

**Proof of Theorem 2.** The upper estimate is known from Theorem 1 and the works of Temlyakov [21, 22] (see also Remark 1). Thus to prove Theorem 2, it suffices for us to prove the lower bound. By virtue of Lemma 1, we only need to show that there exist \( 2n \) functions \( f_1, \ldots, f_{2n} \) which have mutually disjoint supports such that \( \sum_{i=1}^{2n} \delta_i f_i \in W_p^\Lambda(I^d), S(f_i) \geq \epsilon \) and \( \epsilon \asymp n^{-p/d-1}. \) Let
\[ \psi(x) = \begin{cases} a \prod_{j=1}^d \sin^{2\pi} 2\pi x_j, & x \in I^d \\ 0, & \text{elsewhere.} \end{cases} \] (19)
Choose \( a > 0 \) such that \( \psi \in W_p^\Lambda(I^d). \) For given \( n \) we set \( m = \lfloor (2n)^{1/d} \rfloor. \) We divide \( I^d \) into \( m^d \) equal-size cubes with edge length \( 1/m \) and centers \( x_i, i = 1, \ldots, m^d. \) Consider the functions
\[ \psi_i(x) = m^{-\bar{K}} \psi(m(x - x_i)), \quad i = 1, \ldots, m^d. \] (20)
then it is easy to see that \( \psi_i, i = 1, \ldots, m^d, \) satisfy our requirements. The proof of Theorem 2 is complete.

**Remark 3.** In the upper estimate of Theorem 3, as illustration in he beginning of section 2, we reduce the to \( H^k(I^d) \) we will use the stochastic interpolation method which reduces the integration problem to approximation problem. This method is already used by Novak to treat classical Sobolev class. However, we have some difficulties in the study of the class, the main difficulties of include its slight weak smoothness, and different smoothness in different directions. To overcome these difficulties, we split the cube \( I^d \) into some sub-rectangles according to the possible different smoothness in different coordinate directions. Then using the general results of Dahmen, (see [3]) about the coordinate degree local-polynomial approximation, we obtain the estimate of local error relative to each sub-rectangle, and then by a summing property of \( \omega_\Lambda \) which can be derived from the equivalence of the \( K \)-functional \( K_\Lambda \) with the modulus of smoothness \( \omega_\Lambda, \) we get the global error estimate in the randomized setting.

**Lemma 2.** [14] Let \( V \subset L_2([a, b]^d) \) be a linear space with \( \dim V = n \) and \( 1 \in V. \) Then there is a stochastic quadrature formula \( U \in M_n \) for \( S(f) = \int_{[a, b]^d} f(x) dx \) of the form \( S_n(f) = \)
\[ \sum_{i=1}^{n} c_i(\omega) f(a_i(\omega)) \] with the following properties:

\begin{enumerate}
  \item \( E(U(f)) = S(f) \), for all \( f \in L_1([a, b]^d) \);
  \item \( U(f) = S(f) \), for all \( f \in V \);
  \item \( E((U(f) - S(f))^2) \leq (b - a)^d \inf_{g \in V} S((f - g)^2) \), for all \( f \in L_2([a, b]^d) \);
\end{enumerate}

(21)

where \( E(\cdot) \) means expectation of a random variable.

**Lemma 3.**

\( i) \) For any rectangle \( Q \) there exist constants \( C_1, C_2 > 0 \) such that for each vector \( h \in \mathbb{R}^d_+ \), we have

\[ C_1 \omega(\Lambda(f, h, Q)) \leq K_1(f, h, Q) \leq C_2 \omega(\Lambda(f, h, Q)), \]

\[ \text{for all } f \in L_1(Q) ; \]

\[ \text{for } h \in \mathbb{R}^d_+ , \]

(22)

where

\[ \omega(\Lambda(f, h, Q)) = \sum_{\alpha \in \partial \Lambda} \omega(\alpha, f, h, Q) , \]

\[ K_1(f, h, Q) = \inf_{g \in L_1(Q)} \left\{ \| f - g \|_{L_1(Q)} + \sum_{\alpha \in \partial \Lambda} h^\alpha \| D^\alpha g \|_{L_1(Q)} \right\} , \]

(23)

\( \Lambda_0(Q) \) is the set of functions \( g \in L_1(Q) \) for which \( D^\alpha g \in L_1(Q) \) for all \( \alpha \in \partial \Lambda \).

\( ii) \) If \( Q \) is a rectangle in \( \mathbb{R}^d \) with side length vector \( h \in \mathbb{R}^d_+ \), then for each \( f \in L_p(Q) \), there exists a polynomial \( p \in P_\Lambda(d) \) with

\[ \| f - p \|_{L_p(Q)} \leq C_3 \omega(\Lambda(f, h, Q)) , \]

(24)

where \( P_\Lambda(d) := \text{span} \{ t^\alpha : \alpha \in \Lambda \} \).

Note that the constants \( C_1, C_2, \text{ and } C_3 \) depend at most on \( \Lambda \) and \( p \).

To state our next lemma we recall the \( L_p^N \) norm on \( \mathbb{R}^N \) which is defined by \( \| x \|_{L_p^N} = (\sum_{i=1}^{N} |x_i|^p)^{1/p} \) for \( x = (x_1, \ldots, x_N) \in \mathbb{R}^N \).

**Lemma 4.** Let \( \{ Q_i \} \) be a collection of interior disjoint rectangles satisfying \( \bigcup_{i=1}^{N} Q_i = I^d \). Then

\[ \| \{ \omega(\Lambda(f, h, Q_i)) \}_{i=1}^{N} \|_{L_p^N} \leq \omega(\Lambda(f, h, I^d)) \]

(1 \leq p < \infty).

**Proof.** Let \( \| \cdot \|_p^{(i)} \) denote the \( L_p^\star \)-norm on the \( i \)-th rectangle. It follows from part (i) of lemma
3 that
\[
\sum_{i=1}^{N} \omega_{\Lambda}^{p}(f,h,Q_{i}) \ll \sum_{i=1}^{N} \inf_{g \in L_{p}^{\Lambda}(Q_{i})} \left( \|f - g\|_{p}^{(i)} + \sum_{\alpha \in \partial \Lambda} h_{\alpha} \|D_{k} g\|_{p}^{(i)} \right)
\]
\[
\ll \sum_{i=1}^{N} \inf_{g \in L_{p}^{\Lambda}(I^{d})} \left( \|f - g\|_{p}^{(i)} + \sum_{\alpha \in \partial \Lambda} h_{\alpha} \|D_{k} g\|_{p}^{(i)} \right)
\]
\[
= \inf_{g \in L_{p}^{\Lambda}(I^{d})} \left( \|f - g\|_{p}^{(i)} + \sum_{\alpha \in \partial \Lambda} h_{\alpha} \|D_{k} g\|_{p}^{(i)} \right)
\]
\[
\ll \omega_{\Lambda}^{p}(f,h,I^{d})_{p}.
\]

The inequality
\[
\omega_{\Lambda}(f,h,I^{d})_{p} \leq \|\{\omega_{\Lambda}(f,h,Q_{i})\}_{i=1}^{N}\|_{L_{p}^{\infty}}
\]
follows directly from the definition of \(\omega_{\Lambda}\). The proof of the lemma is complete.

**Lemma 5.**

Let \(F_{d}\) be a set of integrable functions and \(f_{i}, i = 1, \ldots, 4n\) with the following conditions:

(a) i) the \(f_{i}\) have disjoint supports and satisfy \(S(f_{i}) = \int f_{i} \geq \delta\) for \(i = 1, \ldots, 4n\),

ii) for all \(\delta_{i} \in \{-1, +1\}\), the function \(\sum_{i=1}^{4n} \delta_{i} f_{i}\) is an element of \(F_{d}\). Then

\[e_{n}^{\text{ran}}(F_{d}) \geq \frac{1}{2} \delta n^{1/2}.\] (25)

(b) We assume that instead of ii) in statement (a) the property

ii)’ for all \(i = 1, \ldots, 4n\) the functions \(\pm f_{i}\) are an element of \(F_{d}\) holds. Then

\[e_{n}^{\text{ran}}(F_{d}) \geq \delta.\] (26)

**Proof of Theorem 3.** It follows from Theorem 1, we prove the upper bound for the class \(H_{p}^{k}(I^{d})\).

For convenience, we write \(k\) for \(k_{\Lambda}\) and \(\omega_{k}\) for \(\omega_{\Lambda_{2}}\). Denote \(P_{k}(d) := \{x^{\alpha} : \alpha_{j} < k_{j}, j = 1, \ldots, d\}\) and \(d_{k} = \dim P_{k}(d)\). We have by the monotonicity of \(L_{p}\) norms,

\[H_{p}^{k}(I^{d}) \hookrightarrow H_{2}^{k}(I^{d}), \quad 2 < p \leq \infty,
\]

therefore, we only have to prove the case of \(p = 2\). Let an integer \(n > d_{k}\) be given and let

\[n_{j} := \left\lceil \left( \frac{n}{d_{k}} \right)^{\frac{n_{j}}{2}} \right\rceil, \quad N := \prod_{j=1}^{d} n_{j}.
\]
Then we split $I^d$ into $N$ congruent rectangles with side length vector $h = \left(\frac{1}{n_1}, \ldots, \frac{1}{n_d}\right)$. In each rectangle $Q_i$ there is a randomized method $U_i \in M_{d_k}$ such that

$$E(U_i(f)) = \int_{Q_i} f(x) dx, \quad (f \in L_1(Q_i));$$
$$U_i(f) = \int_{Q_i} f(x) dx, \quad (f \in P_k(I_d));$$
$$E \left( (U_i(f) - \int_{Q_i} f(x) dx)^2 \right) \ll \frac{1}{N} \inf_{p \in P_k(I_d)} \|f - p\|_{L_2(Q_i)}^2, \quad (f \in L_2(Q_i)).$$

Applying this method independently to each of the $N$ rectangles yields a method

$$U(f) = \sum_{i=1}^{N} U_i(f),$$

it is easy to see that $U \in M_n$, using part (ii) of Lemma 3 and lemma 4, the error of $U$ can be estimated as follows:

$$E \left( (U(f) - \int_{I_d} f(x) dx)^2 \right) = \sum_{i=1}^{N} E \left( (U_i(f) - \int_{Q_i} f(x) dx)^2 \right) \ll N^{-1} \sum_{i=1}^{N} \omega_k^2(f, h, Q_i) \tag{27}$$

$$\ll N^{-1} \omega_k^2(f, h, I_d),$$

with our choice of $N$ and $h$, we yield

$$\Delta_{\text{ran}}(U, H_k^p(I_d)) \ll n^{-\frac{1}{2}} \sum_{j=1}^{d} n_j^{-k_j} \ll n^{-\frac{n(k) - \frac{1}{2}}{d}}$$

and hence

$$e_{\text{ran}}(H_k^p(I_d)) \ll n^{-\frac{n(k) - \frac{1}{2}}{d}}.$$ 

If $1 \leq p < 2$, the corresponding estimate can be derived from the imbedding relation (see [11]),

$$H_k^p(I_d) \hookrightarrow C(p, k) H_2^{(1-\frac{d}{n(k) - \frac{1}{2}})}(I_d),$$

where $\beta := 1/p - 1/2$, we complete the upper estimate of Theorem 3.

Now we deal with the lower bounds. Set $m = \left[(4n)^{1/d}\right]$, we split $I^d$ into $m^d$ equal-size cubes as in the proof of Theorem 2 and define

$$\psi_i = m^{-\frac{d}{p}} \psi(m(x - x_i)), \quad \tilde{\psi}_i = m^{-\frac{d}{p} + \beta} \psi(m(x - x_i)), \quad i = 1, \ldots, m^d.$$
Then $\psi_i$ and $\tilde{\psi}_i$, $i = 1, \ldots, m^d$ satisfy the assumptions of Lemma 5 with $\delta \asymp n^{-\frac{p}{2} - \frac{1}{2}}$ in the case $p \geq 2$ and $\delta \asymp n^{-\frac{p}{d} - 1 + 1/p}$ in the case $1 \leq p < 2$, respectively. The proof of the estimate on randomized error of Theorem 3 is complete.

The upper bound of average error follows from the upper estimates of randomized error, and the proof of lower bound of average error is similar to randomized error, we omit it, see [13, 14] for more details. The proof of Theorem 3 is complete.

Acknowledgement

We would like to express our deep gratitude to the referees and Professor G. Alistair Watson, their critique and excellent remarks, suggestions undoubtedly improved the presentation of this paper.

References


