The Exact Exponent of Sparse Grid Quadratures in the Weighted Case

Leszek Plaskota

Department of Mathematics, Informatics, and Mechanics, Warsaw University, ul. Banacha 2, 02-097 Warsaw, Poland
E-mail: leszekp@mimuw.edu.pl

and

Grzegorz W. Wasilkowski

Department of Computer Science, University of Kentucky, 777 Anderson Hall, Lexington, Kentucky 40506
E-mail: greg@cs.uky.edu

Received February 15, 2000; revised March 24, 2000; accepted April 6, 2000; published online September 20, 2001

This paper provides a lower bound on the exponent of tractability for Sparse Grid Quadratures for multivariate integration of functions from a certain class of weighted tensor product spaces. This lower bound is sharp since it matches a corresponding upper bound of G. W. Wasilkowski and H. Woźniakowski (1999, J. Complexity 15, 402–447). It also shows that, for slowly decreasing weights, the exponent of Sparse Grid Quadratures is far from being optimal. © 2001 Elsevier Science

1. INTRODUCTION

We consider strong tractability of the following class of multivariate integration problems. Given a positive integer $d$, let $\mathbb{H}_d$ be a tensor product

\[ \mathbb{H}_d = \mathbb{H}_1 \times \cdots \times \mathbb{H}_d. \]
space, $\mathcal{H}_d = \bigotimes_{k=1}^d H_k$, where $H_k$ are Hilbert spaces of scalar functions with the inner-products $\langle f, g \rangle_{H_k} = f(0)g(0) + \gamma_k^{-1} \int_0^1 f'(x)g'(x)\,dx$. Here $\{\gamma_k\}_{k=1}^\infty$ is a non-increasing sequence of positive numbers. These numbers are referred to as weights and $\mathcal{H}_d$ are referred to as weighted tensor product spaces. The spaces $\mathcal{H}_d$ provide a model for problems with decreasing importance of consecutive variables. For $f \in \mathcal{H}_d$ we want to approximate the integral

$$S_d(f) = \int_{[0,1]^d} f(x)\,dx$$

by quadratures (algorithms) that only use function evaluations as information operations and, in particular, we would like to know for which weights $\{\gamma_k\}_{k=1}^\infty$ such integration problem is strongly tractable.

We say that a problem is strongly tractable if the cost (i.e., number of function evaluations) needed to reduce the initial error by a factor $\varepsilon$ is bounded by $O(\varepsilon^{-a})$ independently of the dimension $d$. The infimum with respect to such $a$ is then called the exponent of strong tractability.

The tractability of such multivariate integration problems was studied, in particular, by Sloan and Woźniakowski [3] and Novak and Woźniakowski [1]. They showed that strong tractability is equivalent to $\sum_{k=1}^\infty \gamma_k < \infty$. Moreover, the exponent of strong tractability is then between 1 and 2. However, their proof is non-constructive, i.e., it does not identify the corresponding efficient quadratures.

A constructive approach was undertaken in [5] by proposing and studying the class of Weighted Tensor Product (WTP) algorithms. Although more general spaces $\mathcal{H}_d$ and operators $S_d$ are considered, the integration problem defined as above constitutes an important application there. More specifically, suppose $\gamma_k = \Theta(k^{-p})$ for $p > 1$. Then explicit WTP algorithms are provided with cost of reducing the error by $\varepsilon$ bounded from above by, roughly, $O(\varepsilon^{-a})$ with $a \leq \max\{1, 2/(p-1)\}$. For $p \geq 3$, this gives $a \leq 1$ and, hence, implies optimality of such WTP algorithms with $a = 1$. However, for $p < 3$, $2/(p-1)$ is only an upper bound on the exponent and its sharpness has been unknown. Note that if it were sharp, this and the result of Sloan and Woźniakowski [3] would imply that WTP algorithms are not optimal for $p \in (1, 2)$. Actually, they would be very far from being optimal for $p \approx 1$ since $\lim_{p \to 1^+} 2/(p-1) = +\infty$, whereas the exponent of tractability (among all algorithms) is not greater than 2.

The main result of this paper (Theorem 1) implies that the upper bound $\max\{1, 2/(p-1)\}$ is sharp. Actually, we study even a larger class of algorithms, the class of Sparse Grid Quadratures (SGQ) introduced in [2], that includes WTP algorithms. Let $\mathcal{A}_p$ be the set of positive numbers $\alpha$ for
which there exist SGQ with the cost of the error reduction bounded by $O(\varepsilon^{-a})$. Then Theorem 1 and the upper bounds of Wasilkowski and Woźniakowski [5] immediately give the following corollary.

**Corollary 1.** If the weights are given as $\gamma_k = \Theta(k^{-p})$, where $p > 1$, then the minimal exponent among all Sparse Grid Quadratures is

$$\inf_{x \in \mathcal{A}_p} x = \max \left\{ 1, \frac{2}{p-1} \right\}.$$  

2. THE LOWER BOUND

Let

$$\gamma_1 \geq \gamma_2 \geq \gamma_3 \geq \cdots > 0$$

be a given infinite sequence of weights, and let $H_k$ be the reproducing kernel Hilbert space (RKHS) of functions $f : [0, 1] \to \mathbb{R}$ with the reproducing kernel (RK)

$$R_k(x, y) = 1 + \gamma_k \cdot \min\{x, y\}.$$  

We define the space $\mathcal{H}_d$ of functions $f : [0, 1]^d \to \mathbb{R}$ as the tensor product

$$\mathcal{H}_d = \bigotimes_{k=1}^d H_k.$$  

Obviously, $\mathcal{H}_d$ is also an RKHS and its kernel is

$$\mathcal{K}_d(x, y) = \prod_{k=1}^d R_k(x_k, y_k).$$  

We wish to approximate the integral (1) using quadratures $Q(f) = \sum_{j=1}^n \alpha_j f(t_j)$. The quality of a quadrature is measured by its error

$$e(Q) = \|S_d - Q\|_{\mathcal{H}_d} = \sup_{1/1 \leq f \in 1} |S_d(f) - Q(f)|,$$

where $\| \cdot \|_{\mathcal{H}_d}$ denotes both the vector- and operator-norm in $\mathcal{H}_d$, correspondingly.
We consider a special class of quadratures which are given as follows. For each integer \( k \), we first select a sequence of points \( \{t_{kj}\}_{j=1}^{\infty} \in [0, 1] \), and let \( Q_k^i \) be the one-dimensional quadratures that use the points \( t_{kj} \), \( 1 \leq j \leq i \), and minimize the error in the space \( H_k \). Thus \( Q_k^i \) is the \( H_k \)-orthogonal projection of \( S_i \) onto the subspace spanned by \( R_k(t_{kj}, \cdot) \), \( 1 \leq j \leq k \). For a multi-index \( j = (j_1, \ldots, j_d) \), we let

\[
U_j = \bigotimes_{k=1}^{d} (Q_k^i - Q_k^{i-1}).
\]

\( Q_0^i = 0 \) by convention. The quadratures that we consider are given as

\[
\mathcal{P}_n^d = \sum_{j \in \mathcal{P}_n} U_j,
\]

where \( \mathcal{P}_n \subset \mathbb{N}^d \) is a set of multi-indices of cardinality \( n \). Here \( n \) is an arbitrary positive integer. We assume that \( \mathcal{P}_n \) is consistent, which means that if \( j \in \mathcal{P}_n \) and \( i \leq j \) (coordinate-wise) then also \( i \in \mathcal{P}_n \). Note that \( \mathcal{P}_n^d \) is uniquely determined by \( \mathcal{P}_n \).

This kind of quadratures is a generalization of Smolyak’s construction and was considered by Wasilkowski and Woźniakowski [5] for general tensor product linear problems, and by Plaskota [2] for “unweighted” integration. For the purpose of this paper, we call them \( \text{SGQ} \).

Any SGQ corresponding to a consistent \( \mathcal{P}_n \) has the following remarkable property; namely, it uses exactly \( n \) points which are \( t_j = (t_{j_1}, \ldots, t_{j_d}) \) for \( j \in \mathcal{P}_n \), and it is optimal among all quadratures that use the same set of points. Moreover, due to orthogonality of \( U_j \)’s in \( \mathcal{H}_d \), we also have

\[
e(\mathcal{P}_n^d) = \sqrt{\sum_{j \in \mathcal{P}_n} \|U_j\|_{\mathcal{H}_d}^2} = \sqrt{\sum_{j \in \mathcal{P}_n} \prod_{k=1}^{d} \|Q_k^j - Q_k^{j-1}\|_{H_k}^2}.
\]

See [2, 4, 5] for these and other properties of SGQ.

We now define the (minimum) exponent of SGQ as the infimum over all \( \alpha \geq 0 \) that have the following property. For any \( d \geq 1 \) and for any \( \varepsilon > 0 \) there exists a consistent set \( \mathcal{P}_n \) such that for the corresponding quadrature \( \mathcal{P}_n^d \) we have

\[
e(\mathcal{P}_n^d) \leq \varepsilon \cdot \|S_d\|_{\mathcal{H}_d}
\]

and

\[
n \leq K \cdot \varepsilon^{-\alpha}.
\]

Here \( K \) is a constant independent of \( d \) and \( \varepsilon \), but it may depend on \( \alpha \).
Our main result reads as follows.

**THEOREM 1.** Suppose that $\sum_{k=1}^{\infty} \gamma_k < \infty$. Then the exponent of Sparse Grid Quadratures is bounded from below by

$$\alpha^* = \frac{2}{\beta^* - 1},$$

where

$$\beta^* = \sup \left\{ \beta \leq 3 : \sum_{k=1}^{\infty} \gamma_k^{1/\beta} < +\infty \right\}$$

$(2/0 = +\infty$, by convention$)$.

Theorem 1 provides a lower bound for the exponent of SGQ. From [5] it follows that this bound is sharp. Hence $\alpha^*$ is the exact exponent of tractability of SGQ, and Corollary 1 follows.

**Remark 1.** Note that we can equivalently define $\beta^*$ in Theorem 1 as

$$\beta^* = \sup \{ \beta \leq 3 : \gamma_k = O(k^{-\beta}) \}.$$

Indeed, if $\sum_{k=1}^{\infty} \gamma_k^{1/\beta} < \infty$ then $\gamma_k^{1/\beta} = O(1/k)$, i.e., $\gamma_k = O(k^{-\beta})$. On the other hand, if $\gamma_k = O(k^{-\beta})$ then for any $\omega > 1$ we have $\sum_{k=1}^{d} \gamma_k^{\omega/\beta} = \sum_{k=1}^{d} O(k^{-\omega}) = O(1)$.

**Remark 2.** In the definition of the exponent, we formally demand that $e(\mathcal{H}^d) \leq \varepsilon \|S_d\|_{\mathcal{H}^d}$, that is, we are interested in reducing the initial error by $\varepsilon$. Since for $\sum_{k=1}^{\infty} \gamma_k < \infty$ the norms $\|S_d\|_{\mathcal{H}^d} = \prod_{k=1}^{d} (1 + \gamma_k/3)^{1/2}$ are uniformly bounded in $d$, this definition is equivalent to the one in which we replace (2) by

$$e(\mathcal{H}^d) \leq \varepsilon,$$

i.e., when we want the absolute error to be at most $\varepsilon$. In particular, Theorem 1 holds for both (2) or (3).

**Remark 3.** Consider the integration problems in the spaces $\tilde{\mathcal{H}}_d = \bigotimes_{k=1}^{d} H_d$. That is, for each particular $d$, the weights $\gamma_k$, $1 \leq k \leq d$, are all replaced by $\gamma_d$, so that $\tilde{\mathcal{H}}_d$ is the $d$-tensor product of $H_d$. Since $\gamma_k$ decreases, integration in $\tilde{\mathcal{H}}_d$ is not harder than in $\mathcal{H}_d$. However, it turns out that the exponents in both cases are the same. This can be seen by applying the proof of Theorem 1.
3. PROOF OF THEOREM 1

In view of Remark 2, it suffices to prove the theorem with (2) replaced by (3). We do this in three steps.

1. Observe that the exponent of given SGQ depends only on the values

\[ u_j^k = \| Q_{j+1}^k - Q_j^k \|_{\mathcal{H}_0}, \]

where \( k \geq 1 \) and \( j \geq 0 \). Hence, we can speak of exponents of quadratures as well as exponents of (doubly indexed) sequences \( \{ w_j^k \} \). The only technical difference is that, in the latter case, the error \( e(\tilde{P}_n^a) \) should be replaced by

\[ \tilde{e}(\{ w_j^k \}, P_n) = \sqrt{\sum_{j \in \mathcal{P}_n} \prod_{k=1}^d w_{j,k}^k}. \]

We first present a lower bound for the exponent of a sequence of the form

\[
\begin{align*}
  w_0^k &= 1 + a \gamma_k, \\
  w_j^k &= b \gamma_k j^{-q}, \quad \text{for } j \geq 1,
\end{align*}
\]

where \( a \) and \( b \) are some positive reals, \( a \geq b \), and \( q > 1 \).

For \( \delta > 0 \) (which will be chosen later), we select the set of indices as

\[
P_n^a = \left\{ j \geq 0 : \prod_{k=1}^d w_{j,k}^k \geq \delta \right\},
\]

where, as before, \( n = n(\delta) \) is the cardinality of \( P_n^a \). Obviously, \( P_n^a \) is consistent, due to monotonicity of \( \{ w_j^k \}_{j \geq 1} \) for any \( k \geq 1 \). Furthermore, due to the same argument, that selection minimizes the "error" (5) with respect to all \( P_n \) of \( n \) indices. Indeed, to minimize the error, we have to select the \( n \) largest \( \prod_{k=1}^d w_{j,k}^k \), and this is what we do by selecting all the products at least equal to \( \delta \).

We now bound the cardinality \( n \) of \( P_n^a \) and the "error" \( e_d = \tilde{e}(\{ w_j^k \}, P_n) \) from below. For the cardinality, we let

\[
P^{(1)} = \{ j \in P_n^a : \text{exactly one component of } j \text{ is different than } 0 \}.\]
Then
\[ n \geq \#P(1) \geq \sum_{k=1}^{d} \#\{ j \geq 1 : b \gamma_k j^{-q} \geq \delta \} \geq \sum_{k=1}^{d} \lfloor (b \gamma_k / \delta)^{1/q} \rfloor. \]

For the error, we let
\[ P^{(2)} = \{ j \neq P^*_n : \text{exactly one component of } j \text{ is different than } 0 \}. \]

Denoting
\[ A = \prod_{k=1}^{\infty} w_k^d = \prod_{k=1}^{\infty} (1 + a \gamma_k) < \infty, \]
where the finiteness of \( A \) follows from \( \sum_{k=1}^{\infty} \gamma_k < \infty \), we have
\[ e_{2,d}^2 \geq \sum_{j \neq P^*_n} \prod_{k=1}^{d} w_k^j = \sum_{k=1}^{d} \left( b \gamma_k \right) \sum_{j=1}^{\infty} j^{-q} \]
\[ = C_d \sum_{k=1}^{d} \gamma_k^{1/q}. \]

We now set a special value of \( \delta = b \gamma_d \). Since for \( x \geq 1 \) we have \( |x| > x/2 \),
\[ n \geq \sum_{k=1}^{d} \lfloor (\gamma_k / \gamma_d)^{1/q} \rfloor \geq 2^{-1} \gamma_d^{1/q} \sum_{k=1}^{d} \gamma_k^{1/q}. \]  
(7)

To estimate the error for such \( \delta \), we use the fact that \( \sum_{j=1}^{\infty} j^{-q} \geq \int_{\gamma} \gamma^{-q} \, d\gamma = s^{1-q} (q-1)^{-1} \), and that for \( x > 1 \) we have \( |x|+1 < 2x \). We obtain
\[ e_{2,d}^2 \geq \frac{b}{2^{1-q}(q-1)} \sum_{k=1}^{d} \gamma_k \left( \frac{\gamma_d}{A \gamma_k} \right)^{1-1/q} = C_1 \cdot \gamma_d^{1-1/q} \sum_{k=1}^{d} \gamma_k^{1/q}, \]  
(8)

where \( C_1 = 2^{1-q} (q-1)^{-1} A^{1/q-1} \).

Let \( \alpha > 0 \). Then, by (7) and (8), we have
\[ n \geq C_K(d)^{1+\alpha/2} \varepsilon_2^\alpha, \]
(9)

where \( C = C_1^{1/2} / 2, \) and
\[ K_\alpha(d) = \gamma_d^{\alpha/(\alpha+2) - 1/q} \sum_{k=1}^{d} \gamma_k^{1/q}. \]
Since $P^*$ optimally selects $n$ indices, for any other selection $P$ with $\varepsilon(\{u^k\}, P) \leq \varepsilon_d$, the cardinality $\#P$ is bounded from below by the right hand side of (9). Thus if
\[
\limsup_{d \to \infty} K_\alpha(d) = +\infty \tag{10}
\]
then $\#P$ cannot be uniformly in $d$ bounded by $K_{\varepsilon_d}^\alpha$, and the exponent for the given values (6) of $w^k_j$ is at least $\alpha$.

Observe that (10) holds for all $\alpha < 2/(q-1)$. Letting $\beta = 1 + 2/\alpha$ we obtain that the exponent of $\{w^k_j\}$ is at least
\[
\alpha_1 = 2/(\beta^*_1 - 1), \tag{11}
\]
where
\[
\beta^*_1 = \sup \left\{ \beta \in [1, q] : \limsup_{d \to \infty} \gamma^1_{d^{1/\beta}} < \infty \right\}. \tag{12}
\]

2. We now show that
\[
\beta^*_1 = \beta^*_2, \tag{13}
\]
where $\beta^*_1$ is given by (12) and
\[
\beta^*_2 = \sup \left\{ \beta \in [1, q] : \sum_{j=1}^{\infty} \gamma^1_j < \infty \right\}.
\]

Suppose first that for some $\beta < q$, $\sum_{j=1}^{\infty} \gamma^1_j < \infty$. Then $\gamma^1_j = O(1/j)$, i.e., $\gamma_j = O(j^{-\beta})$. Hence,
\[
\gamma^1_{d^{1/\beta}} \sum_{j=1}^{d} \gamma^1_j = O(d^{-\beta + 1/\beta - 1/\alpha} d^{1-\beta/\alpha}) = O(1),
\]
as $d \to \infty$, which proves $\beta^*_1 \geq \beta^*_2$.

Suppose now that for $\beta < q$, $\sum_{j=1}^{\infty} \gamma^1_j = \infty$. Let $0 < \omega < 1$. We choose $p > 1$ such that $p\omega < 1$ and $p\beta < q$. Then there are infinitely many $k$’s for which $\gamma_k^{1/\beta} \geq k^{-p}$. By monotonicity of $\{\gamma_j\}$, for such $k$’s we have
\[
\gamma_k^{\omega/\beta} \sum_{j=1}^{k} \gamma_j^{1/\beta} \geq k \cdot \gamma_k^{\omega/\beta} \geq k \cdot \left( \frac{1}{k} \right)^{\sup} \to \infty,
\]
as \( k \) increases to infinity. Since \( \omega \) can be arbitrarily close to 1, we have \( \beta_n^k \geq \beta_0^k \), and (13) follows.

3. Let \( w_k^j \) and \( u_k^j \) be as in (6) and (4), respectively. We now show that it is possible to choose \( q = 3 \) and parameters \( a \) and \( b \) in (6) such that for any SGQ and corresponding \( \{u_k^j\} \) we have

\[
\sum_{n=1}^{\infty} w_k^j \leq \sum_{n=1}^{\infty} u_k^j, \tag{14}
\]

for all \( n \geq 0 \) and \( k \geq 1 \). Then we can use Lemma 4 of [2] to show that for any SGQ the exponent of the corresponding sequence \( \{u_k^j\} \) is not smaller than the exponent of \( \{w_k^j\} \) given by (11) and (12). Equality (13) will complete the proof. (The proof in [2] was formally for the case of \( w_k^j \) independent of \( k \), but it can be obviously modified for sequences that vary with \( k \).)

To show (14) we fix \( k \) and, for convenience, we drop the superscript \( k \). That is, we want to approximate the integral \( \int_0^1 f(x) \, dx \) for \( f \) in the unit ball of the RKHS \( H \) with RK \( R(x, y) = 1 + \gamma \min \{x, y\} \), where \( \gamma > 0 \). Let \( e^2_n \) be the squared minimum error of quadratures that use \( n \) samples. Obviously, \( e^2_0 = 1 + \gamma / 3 \). Using standard calculations we find that for \( n \geq 1 \)

\[
e^2_n = 3 \min_{0 \leq h < 1} \left\{ h^3 \frac{1 + \gamma h / 4}{1 + \gamma h} + \frac{(1-h)^3}{(2n-1)^2} \right\}. \tag{15}
\]

(The best choice of points is \( t^*_n = h^* + (j-1)(1-h^*)/(n-1/2) \), where \( h^* \) is the optimal \( h \) in (15)). The quantity that we minimize in (15) is bounded from below by \( h^3 / 4 + (1-h)^3 / (2n-1)^2 \). Minimizing this we obtain

\[
e^2_n \geq \frac{\gamma}{3(2n+1)^2} =: s_n.
\]

We now let

\[
a = 8/27, \quad b = 16/675, \quad \text{and} \quad q = 3
\]

in (6). Observe that then \( w_0 = s_0 - s_1 \) (where \( s_0 = e^2_0 \)) and for \( j \geq 1 \)

\[
w_j \leq \frac{\gamma}{3} \left( \frac{1}{(2j+1)^2} - \frac{1}{(2j+3)^2} \right) = s_j - s_{j+1}.
\]
Hence, for all $n \geq 0$,
\[
\sum_{j=n}^{\infty} w_j \leq \sum_{j=n}^{\infty} (s_j - s_{j+1}) = s_n \leq \epsilon_n^2,
\]
and (14) follows from the fact that for any choice of the quadratures $\{Q_j\}$ we have
\[
\epsilon_n^2 \leq \epsilon(Q_n)^2 = \sum_{j=n+1}^{\infty} \|Q_j - Q_{j-1}\|_u^2 = \sum_{j=n}^{\infty} u_j.
\]

The proof of Theorem 1 is complete.

REFERENCES