Institut für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Random approximation of finite sums

P. Mathé

submitted: 22nd September 1992

Institut für Angewandte Analysis und Stochastik
Hausvogteiplatz 5-7
D – O 1086 Berlin
Germany

Preprint No. 11
Berlin 1992

1991 Mathematics Subject Classification. Primary 41 A 44; Secondary 65 C 05, 65 D 30.
Key words and phrases. Monte Carlo integration, optimal error, least favorable distribution.

Herausgegeben vom
Institut für Angewandte Analysis und Stochastik
Hausvogteiplatz 5-7
D – O 1086 Berlin

Fax: + 49 30 2004975
e-Mail (X.400): c=de;a=dbp;p=iaas-berlin;s=preprint
e-Mail (Internet): preprint@iaas-berlin.dbp.de

RANDOM APPROXIMATION OF FINITE SUMS

PETER MATHÉ

September 2, 1992

ABSTRACT. This paper is devoted to a detailed study of the randomized approximation of finite sums, i.e., sums $\sum_{j=1}^{m} x_j, \quad x \in \mathbb{R}^m$, where m is supposed to be large, shall be approximated with information on n coordinates, only. The error is measured on balls in $l_p^m, \quad 1 \leq p \leq \infty$. Main emphasis is laid on the exact solution of the problems stated below. In most cases we obtain both, an optimal method for the Monte Carlo setting and the description of least favorable distributions for the average case setting, exhibiting results obtained in a previous paper by the author, [Mat92]. Moreover, the solution of the finite-dimensional problem is applied to the Monte Carlo integration of continuous functions. Finally, this knowledge is used to study some of the properties, the optimal methods possess.

1. INTRODUCTION

This paper is devoted to a detailed study of the randomized approximation of finite sums. Although this is no problem of numerical relevance, it may serve as a model study of Monte Carlo integration on $C(0,1)$, the continuous functions on the unit interval, equipped with the supremum norm $\|\cdot\|_\infty$. This application is studied in the last section.

Main emphasis is laid on the exact solution of the problems stated below. This work shall exhibit results obtained in a previous paper by the author, see [Mat92].

Since we are able to compute the optimal methods in many cases explicitely we can study some of their properties. The usual (crude) Monte Carlo integration is motivated by the law of large numbers, cf. [Erm71, §2.4] : If $f \in C(0,1)$ is any function, then $\int_0^1 f(x)dx = \mathbf{E}f = \lim_{n\to\infty} 1/n \sum_{j=1}^{n} f(\xi_j), a.s.$, where the ξ_j are independent random variables, distributed uniformly on $[0,1]$. A simple computation proves

$$\left(\mathbf{E}\left|\int_0^1 f(x)dx - \frac{1}{n}\sum_{j=1}^{n} f(\xi_j)\right|^2\right)^{1/2} \leq \frac{\|f\|_\infty}{\sqrt{n}}.$$

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 41 A 44; Secondary 65 C 05, 65 D 30.

Key words and phrases. Monte Carlo integration, optimal error, least favorable distribution.

Another standard argument proves that this typical Monte Carlo rate $n^{-1/2}$ cannot be improved on the class of all functions with $\|f\|_\infty \leq 1$. Thus at the level of the rate of convergence we cannot expect to gain new insight.

We will however compute the exact optimal error and provide a strictly optimal random method.

2. Optimal approximation of finite sums

We shall study stochastic methods to approximate sums with a large number of summands. Since we are dealing with the discrete case in this section the underlying probabilities, describing the stochastic methods can be treated as combinatorial.

The analysis carried out below is in the framework and notation of Information-based Complexity, see [TWW88]. Additional background information can be obtained there. Summation shall be considered as a linear functional S^m on $\mathbb{R}^m$, where m represents the length of the sum. To be precise, let $S^m : \mathbb{R}^m \to \mathbb{R}$ be defined as

$$S^m(x) := \sum_{j=1}^m x_j, \quad x = (x_j)_{j=1}^m \in \mathbb{R}^m.$$

Weighted sums can be represented similarly as $S_\sigma^m : \mathbb{R}^m \to \mathbb{R}$ via

$$S_\sigma^m(x) := \sum_{j=1}^m \sigma_j x_j, \quad x = (x_j)_{j=1}^m \in \mathbb{R}^m.$$

(We always shall assume, that the weights $\sigma = (\sigma_j)_{j=1}^m$ are positive and arranged in decreasing order, $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$.) We shall consider S^m or S_σ^m as functionals acting on the closed unit balls $B_p^m := \left\{x \in \mathbb{R}^m, \ \sum_{j=1}^m |x_j|^p \leq 1\right\} \subset \mathbb{R}^m, 1 \leq p \leq \infty$, the usual modification for $p = \infty$, i.e., B_p^m is the class of problem elements.

Given $n \in \mathbb{N}$ we shall allow any deterministic method u of the form

$$u(x) := \sum_{j \in I} c_j x_j, \quad x \in \mathbb{R}^m,$$

where $I \subset \{1, \ldots, m\}$ is any subset of at most n elements and $(c_j)_{j=1}^m \in \mathbb{R}^m$ are arbitrary weights (Observe, that the values of $c_j, j \notin I$ are not needed in the computation). Thus the class of admissible methods is

$$\mathfrak{M}_n(\mathbb{R}^m) := \left\{ u(x) = \sum_{j \in I} c_j x_j, \quad \operatorname{card}(I) \leq n, c_j \in \mathbb{R} \right\}.$$

All these methods u are linear (in x) functionals on $\mathbb{R}^m$, which can be written as $u(x) = \sum_{j=1}^m \chi_I(j) c_j x_j$, where χ_I is the indicator function of the subset $I \subset \{1, \ldots, m\}$.

Let us, for completeness, review the worst-case problem of approximating S_σ^m by methods $u \in \mathfrak{M}_n(\mathbb{R}^m)$ here, i.e., the error of any method u on B_p^m is defined as

$$e(S_\sigma^m, u, B_p^m) := \sup\left\{|S_\sigma^m(x) - u(x)|, \quad x \in B_p^m\right\},$$

and

$$e(S_\sigma^m, \mathfrak{M}_n(\mathbb{R}^m), B_p^m) := \inf\left\{e(S_\sigma^m, u, B_p^m), \quad u \in \mathfrak{M}_n(\mathbb{R}^m)\right\}.$$

Proposition 1. *For any $p, 1 \leq p \leq \infty$ and $1 \leq n \leq m$ we have*

(1) $e(S_\sigma^m, \mathfrak{M}_n(\mathbb{R}^m), B_p^m) = \left(\sum_{j=n+1}^m |\sigma_j|^{p'}\right)^{1/p'}$

(2) $e(S^m, \mathfrak{M}_n(\mathbb{R}^m), B_p^m) = (m-n)^{1/p'}$.

(where p' is conjugate to p, i.e., $1/p + 1/p' = 1$)

Proof: It is easy to see that the method $u(x) = \sum_{j=1}^n \sigma_j x_j$ yields the error stated in the theorem. To prove that this is best possible, let u with I and $(c_j)_{j=1}^m$ be chosen. We have

$$\sup_{x \in B_p^m} |S_\sigma^m(x) - u(x)|^{p'} = \|S_\sigma^m - u\|_{p'}^{p'} = \sum_{j \in I} |\sigma_j - c_j|^{p'} + \sum_{j \notin I} |\sigma_j|^{p'} \geq \sum_{j \notin I} |\sigma_j|^{p'}.$$

Since the weights are arranged in decreasing order we obtain the lower estimate. □

The above result may serve as a reference to be compared with the estimate obtained for random methods.

To proceed we shall introduce random (stochastic) methods. Stochastic methods are considered as probabilities on the class of (admissible) deterministic methods. Therefore the class $\mathfrak{M}_n(\mathbb{R}^m)$ can be identified with a special Borel subset of $\mathbb{R}^m$ and it shall be equipped with the respecting Borel σ-algebra, denoted by $\mathcal{F}$ (we suppress the parameters n, m). Any probability P on $[\mathfrak{M}_n(\mathbb{R}^m), \mathcal{F}]$ will be called a Monte Carlo method on $\mathfrak{M}_n(\mathbb{R}^m)$. This approach is equivalent to the following (more intuitive) procedure. One chooses any n-subset I of $\{1, \ldots, m\}$ at random, further weights $c = (c_1, \ldots, c_m)$ are randomly chosen and the composition $u_{I,c}(x) = \sum_{j \in I} c_j x_j$ gives our random method in $\mathfrak{M}_n(\mathbb{R}^m)$. We shall however provide an optimal method, which only chooses the set I of indices at random.

It is well-known that the mapping $(u, x) \to u(x)$ is product measurable (where B_p^m is equipped with the Borel-σ-algebra, too), cf. [Mat92]. The error of any random method P on $[\mathfrak{M}_n(\mathbb{R}^m), \mathcal{F}]$ is defined as usual by

$$e(S_\sigma^m, P, x) := \left(\int |S_\sigma^m(x) - u(x)|^2 dP(u)\right)^{1/2}$$

at any specific element $x \in B_p^m$, whereas

$$e(S_\sigma^m, P, B_p^m) := \sup_{x \in B_p^m} \left(\int |S_\sigma^m(x) - u(x)|^2 dP(u)\right)^{1/2}$$

characterizes the overall quality of P. We shall look for an optimal method P minimizing the above quantity, i.e.,

$$e^{mc}(S_\sigma^m, \mathfrak{M}_n(\mathbb{R}^m), B_p^m) := \inf\left\{e(S_\sigma^m, P, B_p^m), \quad P \text{ a probability on } \mathfrak{M}_n(\mathbb{R}^m)\right\}.$$

It has been proven by the author [Mat92] that under the above assumptions the Monte Carlo error $e^{mc}(S_\sigma^m, \mathfrak{M}_n(\mathbb{R}^m), B_p^m)$ is equal to the average-case error

$$(1) \qquad e^{avg}(S_\sigma^m, \mathfrak{M}_n(\mathbb{R}^m), B_p^m) := \sup\left\{e(S_\sigma^m, \mathfrak{M}_n(\mathbb{R}^m), \mu), \quad \mu(B_p^m) = 1\right\}.$$

Here μ is any Borel probability on B_p^m and

$$e(S_\sigma^m, \mathfrak{M}_n(\mathbb{R}^m), \mu) := \inf\left\{\left(\int |S_\sigma^m(x) - u(x)|^2 d\mu(x)\right)^{1/2}, \quad u \in \mathfrak{M}_n(\mathbb{R}^m)\right\}$$

denotes the μ-average-case error. Any probability μ maximizing the right-hand side in (1) is said to be least favorable.

The aim of this section is the description of both the optimal random methods P and least favorable distributions μ for the problems S^m on $B_p^m, 1 \le p \le \infty$.

Let us start with a result for S_σ^m in case $p = 1$, similar to the example treated in [Mat92].

Theorem 1. *For any* $1 \le n \le m$ *we have*

(1)

$$e^{mc}(S_\sigma^m, \mathfrak{M}_n(\mathbb{R}^m), B_1^m) = \max\left\{\left(\frac{h-n}{\sum_{j=1}^h \sigma_j^{-2}}\right)^{1/2}, \quad h = n+1, \dots, m\right\}.$$

(2) *As a special case we have*

$$e^{mc}(S^m, \mathfrak{M}_n(\mathbb{R}^m), B_1^m) = \sqrt{\frac{m-n}{m}}.$$

Proof: We will prove the equality in (1) by a reduction of the above summation problem to the case of the example treated in [Mat92]. The reader is referred to the computation carried out there. We know from [Mat92, Thm. 5] that a least favorable distribution μ for the average-case may be chosen to sit at the extreme points $\{\pm e_1, \dots, \pm e_m\}$ of B_1^m, where $e_1, \dots, e_m$ denotes the unit vector basis in $\mathbb{R}^m$ i.e., μ is of the form

$$\mu = \sum_{j=1}^m (\alpha_j \delta_{e_j} + \beta_j \delta_{-e_j}),$$

for some $\alpha_j, \beta_j \ge 0$, $\sum_{j=1}^m (\alpha_j + \beta_j) = 1$. For such μ and any method

$$u(x) = \sum_{j=1}^m \chi_I(j) c_j x_j, \quad x \in \mathbb{R}^m$$

we can compute the error

$$\begin{aligned} e(S_\sigma, u, \mu)^2 &= \int |S_\sigma^m(x) - u(x)|^2 d\mu(x) \\ &= \sum_{j=1}^m (\alpha_j + \beta_j)|\sigma_j - \chi_I(j)c_j|^2 \\ &= \sum_{j=1}^m \gamma_j |\sigma_j - \chi_I(j)c_j|^2 \end{aligned}$$

with $\gamma_j = \alpha_j + \beta_j, \quad j = 1, \dots, m$. It is easy to see that for any $(\gamma_j)_{j=1}^m$ and choice of I the above expression is minimized with respect to $(c_j)_{j\in I}$ if $c_j := \sigma_j, \quad j \in I$ hence

$$\begin{aligned} e(S_\sigma^m, \mathfrak{M}_n(\mathbb{R}^m), \mu)^2 &= \inf\left\{\sum_{j=1}^m \gamma_j\sigma_j^2(1-\chi_I(j)), \quad \mathrm{card}(I) = n\right\} \\ &= \inf\left\{\sum_{j=1}^m \gamma_j\sigma_j^2(1-\xi_j), \quad \xi_j \in \{0,1\}, \sum_{j=1}^m \xi_j = n\right\} \qquad (2) \\ &= \tilde{a}_{n+1}^2(D_\sigma, \mu) \end{aligned}$$

with $\tilde{a}_{n+1}^2(D_\sigma, \mu)$ as in [Mat92, equation 21]. The shift from $n \to n+1$ results from the different definition of the quantities. Thus we have carried out the reduction process and can rely on the result proved in [Mat92]. □

Remark 1. The second statement in the theorem can be proven directly, proceeding after (2). Moreover, an optimal random method is obtained by a uniform choice of n-sets I out of $\{1, \dots, m\}$ and equal weights $c_1 = \dots = c_m = 1$. A least favorable distribution μ is the uniform distribution on $\{\pm e_1, \dots, \pm e_m\}$.

Now we are prepared for the main result of this section. Considerations will henceforth be restricted to the case S^m, since no solution is available for the general case.

Theorem 2. *For any* $1 \le n < m$ *we have*

(1) $e_n^{mc}(S^m, \mathfrak{M}_n(\mathbb{R}^m), B_\infty^m) = \frac{m}{1+\sqrt{\frac{n(m-1)}{m-n}}}$.

(2) *An optimal Monte Carlo method for the above problem is obtained by a uniform choice of n-sets I out of* $\{1, \dots, m\}$ *and equal weights*

$$c_\infty = c_1 = \dots = c_m = \frac{m}{n + \sqrt{\frac{n(m-n)}{m-1}}}.$$

Proof: We shall prove the upper bound by computing the error of the method described in (2). The case $n = 1$ shall be omitted since it is easier to handle. So let us fix $2 \le n < m$. The uniform choice of n-sets I can be characterized as a probability

ρ on $\mathfrak{P}(n)$, the collection of all n-subsets of $\{1,\dots,m\}$, i.e., $\rho = \frac{1}{\binom{m}{n}}\sum_{I\in\mathfrak{P}(n)}\delta_I$. Let us consider the error of the random method P, obtained from ρ by

$$P(u_{I,c}) = \begin{cases}\rho(I) & \text{,if } c = c_\infty\\ 0 & \text{,if } c \neq c_\infty\end{cases}.$$

Given any $x \in B_\infty^m$ we calculate

$$\begin{aligned}
e(S^m,P,x)^2 &= \frac{1}{\binom{m}{n}}\sum_{I\in\mathfrak{P}(n)}|\sum_{j=1}^m x_j - c_\infty\sum_{j\in I}x_j|^2\\
&= \frac{1}{\binom{m}{n}}\sum_{I\in\mathfrak{P}(n)}|\sum_{j=1}^m x_j(1-c_\infty\chi_I(j))|^2\\
&= \frac{1}{\binom{m}{n}}\sum_{I\in\mathfrak{P}(n)}\sum_{k,l=1}^m x_kx_l(1-c_\infty\chi_I(k))(1-c_\infty\chi_I(l))\\
&= \sum_{k=1}^m x_k^2\frac{1}{\binom{m}{n}}\sum_{I\in\mathfrak{P}(n)}(1-c_\infty\chi_I(k))^2\\
&+ \sum_{k\neq l}x_kx_l\frac{1}{\binom{m}{n}}\sum_{I\in\mathfrak{P}(n)}(1-c_\infty\chi_I(k))(1-c_\infty\chi_I(l))\\
&= \sum_{k=1}^m x_k^2\left(1-2c_\infty\frac{\binom{m-1}{n-1}}{\binom{m}{n}}+c_\infty^2\frac{\binom{m-1}{n-1}}{\binom{m}{n}}\right)\\
&+ \sum_{k\neq l}x_kx_l\left(1-2c_\infty\frac{\binom{m-1}{n-1}}{\binom{m}{n}}+c_\infty^2\frac{\binom{m-2}{n-2}}{\binom{m}{n}}\right)
\end{aligned}$$

$$(3)\qquad = \|x\|_2^2\left(1-2c_\infty\frac{n}{m}+c_\infty^2\frac{n}{m}\right)+\sum_{k\neq l}x_kx_l\left(1-2c_\infty\frac{n}{m}+c_\infty^2\frac{n(n-1)}{m(m-1)}\right).$$

It is easy to check that $1-2c_\infty\frac{n}{m}+c_\infty^2\frac{n(n-1)}{m(m-1)} = 0$ and $1-2c_\infty\frac{n}{m}+c_\infty^2\frac{n}{m} = \frac{m}{\left(1+\sqrt{\frac{n(m-1)}{m-n}}\right)^2}$. Thus we obtain

$$(4)\qquad e(S^m,P,B_\infty^m) = \frac{m^{1/2}}{1+\sqrt{\frac{n(m-1)}{m-n}}}\sup_{\|x\|_\infty=1}\|x\|_2$$

$$(5)\qquad = \frac{m}{1+\sqrt{\frac{n(m-1)}{m-n}}},$$

which completes the proof of the upper estimate.

We turn to the lower estimate and shall construct a probability μ on $\{+1,-1\}^m \in B_\infty^m$, which provides the μ-average-case error

$$e(S^m,\mathfrak{M}_n(\mathbb{R}^m),\mu) = \frac{m}{1+\sqrt{\frac{n(m-1)}{m-n}}}.$$

This probability will turn out to be a mixture of probabilities $\mu_l^m, l = 0, \dots, m$ which themselves are uniform distributions on all vectors $x \in \{+1, -1\}^m$, satisfying $\sum_{j=1}^m x_j = m - 2l$, i.e., exactly l of their components are equal to -1. The following lemma describes properties of these measures and can be proven by straightforward computation.

Lemma 1. *For any choice of m and $l = 0, \dots, m$ we have*

(1)
$$\int x_j d\mu_l^m(x) = \frac{m-2l}{m}, \quad j = 1, \dots, m,$$

(2)
$$\int x_j x_k d\mu_l^m(x) = \begin{cases} 1 & , \text{if } j = k \\ \frac{(m-2l)^2 - m}{m(m-1)} & , \text{if } j \neq k \end{cases}.$$

(3) *For any method $u(x) = \sum_{j \in I} c_j x_j, x \in \mathbb{R}^m$ we have*

$$\begin{aligned} e(S^m, u, \mu_l^m)^2 &= (m-2l)^2 - \frac{2(m-2l)^2}{m} \sum_{j \in I} c_j + \frac{(m-2l)^2 - m}{m(m-1)} \left(\sum_{j \in I} c_j \right)^2 \\ &+ \left(1 - \frac{(m-2l)^2 - m}{m(m-1)} \right) \sum_{j \in I} c_j^2. \end{aligned}$$

Observe that $e(S^m, u, \mu_l^m)^2 = e(S^m, u, \mu_{m-l}^m)^2$. Using the above identities we construct the desired μ as follows. Put $\beta := \sqrt{\frac{n(m-1)}{m-n}}$ and $w := \frac{m^2}{1+\beta}$. Let $(w_l)_{l=0}^m, w_l \geq 0, \sum_{l=0}^m w_l = 1$ be chosen such that

$$\sum_{l=0}^m w_l (m-2l)^2 = w.$$

This is possible since $0 < w < m^2$. Having fixed such choice of $(w_l)_{l=0}^m$ let μ be defined as

$$\mu := \sum_{l=0}^m w_l \mu_l^m.$$

Let us now fix any method $u_{I,c} \in \mathfrak{M}_n(\mathbb{R}^m)$ and put $s := \sum_{j \in I} c_j$. We conclude, using topic (3) of Lemma 1 that

$$\begin{aligned} e(S^m, u, \mu)^2 &= \sum_{l=0}^m w_l e(S^m, u, \mu_l^m)^2 \\ &= w - \frac{2w}{m} s + \frac{s^2(w-m)}{m(m-1)} + \frac{m^2 - w}{m(m-1)} \sum_{j \in I} c_j^2. \end{aligned}$$

To minimize this with respect to the choice of $u \in \mathfrak{M}_n(\mathbb{R}^m)$ we can sequentially minimize over $(c_j)_{j \in I}$ for fixed s and I, choices of $I, \text{card}(I) \leq n$ and choices of s

and arrive at

$$\begin{aligned}
e(S^m, \mathfrak{M}_n(\mathbb{R}^m), \mu)^2 &\geq \inf_s \left\{ w - \frac{2ws}{m} + \frac{s^2}{mn(m-1)} (w(n-1) + m(m-n)) \right\} \\
&= w \inf_s \left\{ 1 - \frac{2}{m} s + \frac{m(n-1) + (m-n)(1+\beta)}{m^2 n(m-1)} s^2 \right\} \\
&= w \left(1 - \frac{n(m-1)}{m(n-1) + (m-n)(1+\beta)} \right) \\
&= \frac{w}{1+\beta} \left(1 + \beta - \frac{n(m-1)}{(m-n) + \frac{m(n-1)}{1+\beta}} \right).
\end{aligned}$$

Since $\beta^2 - 1 = \frac{m(n-1)}{m-n}$ we obtain

$$e(S^m, \mathfrak{M}_n(\mathbb{R}^m), \mu)^2 \geq \frac{w}{1+\beta},$$

which proves the lower bound. Moreover we can see, that the lower bound is attained for any method $u(x) = c \sum_{j \in I} x_j$ where $c = \frac{s}{n} = \frac{m}{n + \sqrt{\frac{n(m-n)}{m-1}}}$. □

Another look at the proof of the foregoing Theorem 2 gives rise to

Corollary 1. *For* $2 \leq q \leq \infty$ *we have for* $1 \leq n < m$

$$e^{mc}(S^m, \mathfrak{M}_n(\mathbb{R}^m), B_q^m) = \frac{m^{1/q'}}{1 + \sqrt{\frac{n(m-1)}{m-n}}},$$

where this bound is attained with the same method as described in topic (2) of Theorem 2.

Proof: A look at the proof of the upper bound in Theorem 2, see equation (3) gives an individual error for the method P_∞

$$e(S^m, P_\infty, x) = \|x\|_2 \left(\frac{m^{1/2}}{1 + \sqrt{\frac{n(m-1)}{m-n}}} \right).$$

Thus, analogously to (4) we have

$$e(S^m, P_\infty, B_q^m) = \frac{m^{1/2}}{1 + \sqrt{\frac{n(m-1)}{m-n}}} \max_{\|x\|_q = 1} \|x\|_2 = \frac{m^{1/q'}}{1 + \sqrt{\frac{n(m-1)}{m-n}}}.$$

To prove the lower bound define $T_q : \mathbb{R}^m \to \mathbb{R}^m$ as

$$T_q(x) := m^{-1/q} x, \quad x \in \mathbb{R}^m.$$

It is easily seen that $\|T_q : l_\infty^m \to l_q^m\| = 1$. Let μ be any least favorable distribution for B_∞^m and put $\mu_q := \mu \circ T_q^{-1}$, be the image distribution of μ under T_q. Hence $\mu(B_q^m) = 1$ and we have for any method $u \in \mathfrak{M}_n(\mathbb{R}^m)$

$$e(S^m, u, \mu_q) = m^{-1/q} e(S^m, u, \mu),$$

from which the proof can be completed. □

We turn to the optimal error for S^m on balls B_p^m, $1 < p < 2$. We are, however, not able to compute the exact error, as done before, but we can provide an estimate which reflects the right asymptotics with respect to $m, n \in \mathbb{N}$. The inherent difficulty in case $1 < p < 2$ can be seen from equation (3) and

$$\sup_{\|x\|_p=1} \|x\|_2 = \begin{cases} 1 & , 1 \le p \le 2 \\ m^{1/2-1/p} & , 2 \le p \le \infty \end{cases}.$$

Given two functions $f, g : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ the asymptotics $f \asymp g$ means that there are constants $0 < c < C < \infty$, such that $cf(m,n) \le g(m,n) \le Cf(m,n), m, n \in \mathbb{N}$.

Theorem 3. *Let $1 \le p \le 2$. Then we have for $1 \le n \le m$*

(1)

$$e^{mc}(S^m, \mathfrak{M}_n(\mathbb{R}^m), B_p^m) \asymp \sqrt{\frac{m-n}{n + (m-n)m^{-2/p'}}}.$$

(2) *An almost optimal Monte Carlo method P_p is obtained from uniform sampling of $I \in \mathfrak{P}(n)$ with weight $c_p = \frac{m}{n+(m-n)m^{-2/p'}}$.*

Remark 2. The extremal cases $p = 1, 2$ can be compared with the results in (2) of Theorem 1. and Corollary 1.

Proof (of the theorem): We shall start with the individual error, computed in (3), which represents the error of any method P, consisting of uniform sampling and weighting with equal weight c. Hence

$$\begin{aligned} e(S^m, P, B_p^m)^2 &= \sup_{\|x\|_p=1} \left\{ \|x\|_2^2 \frac{n(m-n)}{m(m-1)} + \left(\sum_{j=1}^m x_j \right)^2 \left(1 - 2c\frac{n}{m} + c^2 \frac{n(n-1)}{m(m-1)} \right) \right\} \\ &\asymp \sup_{\|x\|_p=1} \left\{ \|x\|_2^2 \frac{n(m-n)}{m^2} + \left(\sum_{j=1}^m x_j \right)^2 \left(1 - c\frac{n}{m} \right)^2 \right\}. \end{aligned}$$

If we substitute the value of c_p and observe that $|\sum_{j=1}^m x_j| \le m^{1/p'}$, $x \in B_p^m$ we obtain the upper estimate.

To prove the lower estimate we have to choose an *almost* least favorable distribution on B_p^m, appropriately. So, let μ_1 be the uniform distribution on $\{\pm e_1, \dots, \pm e_m\}$ and μ_2 be the two-point distribution on $e = m^{-1/p}(1, \dots, 1)$ and $-e$. Thus μ_1 and μ_2 are concentrated on B_p^m. Further, let $w := \frac{m^2}{m^2 + (m^2 - m^{2/p})}$, hence $0 < w \le 1$. Finally define $\mu := w\mu_1 + (1-w)\mu_2$.

To estimate the error with respect to this measure μ choose any method $u(x) = \sum_{j\in I} c_j x_j$ and let $s := \sum_{j\in I} c_j$. We have

$$\begin{aligned} e(S^m, u, \mu)^2 &= we(S^m, u, \mu_1)^2 + (1-w)e(S^m, u, \mu_2)^2 \\ &= \frac{w}{m}\left(n - 2s + \sum_{j\in I} c_j^2\right) + \frac{w(m-n)}{m} + \frac{1-w}{m^{2/p}}\left(m^2 - 2sm + s^2\right). \end{aligned}$$

Minimizing sequentially over $(c_j)_{j\in I}$ for fixed s and I, choices of $I, \mathrm{card}(I) \leq n$ and choices of s we arrive at

$$e(S^m, u, \mu)^2 \geq \inf_s \left\{ \frac{wn}{m}\left(1 - \frac{s}{n}\right)^2 + w\frac{m-n}{m} + \frac{1-w}{m^{2/p}}(m-s)^2 \right\}.$$

This is minimized at $s = \frac{mn}{n+(m-n)m^{-2/p'}}$, hence $c_p := \frac{s}{n}$ as stated in (2) in the formulation of the theorem. The choice of such c_p, denote the respective method by u_p for a moment, provides the error

$$e(S^m, u_p, \mu)^2 = \frac{m-n}{(2 - m^{-2/p'})(n + (m-n)m^{-2/p'})} \geq \frac{1}{2}\frac{m-n}{(n + (m-n)m^{-2/p'})}. \quad \square$$

3. Applications to Monte Carlo integration

In this section we are going to apply the results of Section 2 to the simplest integration problem $\mathrm{Int} : C(0,1) \to \mathbb{R}$ with $\mathrm{Int}(f) := \int_0^1 f(t)dt, \quad f \in C(0,1)$. This is natural, since one can expect that for large m the functionals $\frac{1}{m}S^m$ tend to Int. The class of possible (deterministic) methods shall be

$$\mathfrak{M}_n(C(0,1)) := \left\{ \sum_{j=1}^n c_j \delta_{\xi_j}, \quad c_j \in \mathbb{R}, \xi_j \in (0,1) \right\},$$

which means, that any $u \in \mathfrak{M}_n(C(0,1))$, which is of the form $u = \sum_{j=1}^n c_j \delta_{\xi_j}$ acts as $u(f) = \sum_{j=1}^n c_j f(\xi_j), \quad f \in C(0,1)$. One can prove that $\mathfrak{M}_n(C(0,1))$ is a Borel set in the space of finite measures on $(0,1)$, equipped with the weak topology. According to the definition in [Mat92] any probability P on $\mathfrak{M}_n(C(0,1))$ will be called a Monte Carlo method, the error of which on $B(0,1)$, the unit ball of $C(0,1)$ in the supremum norm $\|f\|_\infty := \max_{t\in[0,1]} |f(t)|$ will be defined as

$$e(\mathrm{Int}, P, B(0,1)) := \sup_{\|f\|_\infty = 1} \left(\int |\mathrm{Int}(f) - u(f)|^2 dP(u) \right)^{1/2}$$

and

$$e^{mc}(\mathrm{Int}, \mathfrak{M}_n(C(0,1)), B(0,1)) := \inf\left\{ e(\mathrm{Int}, P, B(0,1)), \quad P(\mathfrak{M}_n(C(0,1))) = 1 \right\}.$$

With the above notation we can state

Theorem 4. *For every $n \in \mathbb{N}$ we have*

$$e^{mc}(\mathrm{Int}, \mathfrak{M}_n(C(0,1)), B(0,1)) = \frac{1}{1+\sqrt{n}}.$$

An optimal Monte Carlo method can be chosen to be of the form $P_n := c_n \sum_{j=1}^{n} \delta_{\xi_j}$, where $(\xi_j)_{j=1}^{n}$ is an independent uniformly distributed sample in $(0,1)$ and $c_n := \frac{1}{n+\sqrt{n}}$.

Proof: P_n is clearly a Monte Carlo method on $\mathfrak{M}_n(C(0,1))$. It provides an error at $f \in B(0,1)$

$$\begin{aligned} e(\mathrm{Int}, P_n, f)^2 &= \int \cdots \int |\,\mathrm{Int}(f) - c_n \sum_{j=1}^{n} f(\xi_j)|^2 d\xi_1 \ldots \xi_n \\ &= (1 - 2nc_n + n(n-1)c_n^2)\,|\,\mathrm{Int}(f)|^2 + nc_n^2 \,\mathrm{Int}(f^2). \end{aligned} \tag{6}$$

Since c_n is chosen satisfying $1 - 2nc_n + n(n-1)c_n^2 = 0$ we arrive at

$$e(\mathrm{Int}, P_n, B(0,1)) = \sqrt{n}c_n = \frac{1}{1+\sqrt{n}}.$$

It remains to prove $\frac{1}{1+\sqrt{n}}$ to be a lower bound. Roughly one could think of $\mathbb{R}^m$ to be mapped to a space of step functions, $x \in \mathbb{R}^m \to \sum_{j=1}^{m} x_j \chi_{[\frac{j-1}{m}, \frac{j}{m})}$. Such a mapping would easily lead to the desired lower bound. However, since step functions are not continuous we shall make use of a pertubation argument. To this end, given any $0 < \varepsilon < \frac{1}{2}$ let $\varphi_\varepsilon \in C(0,1)$ be a piecewise linear continuous function, defined as

$$\varphi_\varepsilon(t) = \begin{cases} \frac{t}{\varepsilon} & ,0 \le t < \varepsilon \\ 1 & ,\varepsilon \le t < 1-\varepsilon \\ \frac{1-t}{\varepsilon} & ,1-\varepsilon \le t \le 1 \end{cases}.$$

Given $m \in \mathbb{N}$ let

$$f_j(t) := \varphi(mt - j + 1), \quad t \in [0,1], j = 1, \ldots, m$$

be a family of m continuous functions with almost disjoint support, $\mathrm{supp}\, f_j = [\frac{j-1}{m}, \frac{j}{m}]$. Further we define $T_\varepsilon^m : \mathbb{R}^m \to C(0,1)$ by

$$T_\varepsilon^m(x) = \sum_{j=1}^{m} x_j f_j, \quad x \in \mathbb{R}^m.$$

It is clear that T_ε^m is linear and $\|T_\varepsilon^m : l_\infty^m \to C(0,1)\| = 1$. Moreover one has

$$\mathrm{Int} \circ T_\varepsilon^m = \frac{1-\varepsilon}{m} S^m,$$

and, if $u \in \mathfrak{M}_n(C(0,1))$, then $u \circ T_\varepsilon^m$ meets at most n coordinates, thus $u \circ T_\varepsilon^m \in \mathfrak{M}_n(\mathbb{R}^m)$.

Now, given any distribution ν^m on B^m_∞ put $\mu^m_\varepsilon := \nu^m \circ (T^m_\varepsilon)^{-1}$, the image of ν^m under T^m_ε. The above arguments imply, given any method $u \in \mathfrak{M}_n(C(0,1))$,

$$\begin{aligned} e(\mathrm{Int}, u, \mu^m_\varepsilon) &= e(\mathrm{Int} \circ T^m_\varepsilon, u \circ T^m_\varepsilon, \nu^m) \\ &= e(\frac{1-\varepsilon}{m} S^m, u \circ T^m_\varepsilon, \nu^m). \end{aligned}$$

Consequently,

$$\begin{aligned} e^{avg}(\mathrm{Int}, \mathfrak{M}_n(C(0,1)), B(0,1)) &\geq e^{avg}(\frac{1-\varepsilon}{m} S^m, \mathfrak{M}_n(\mathbb{R}^m), B^m_\infty) \\ &= \frac{1-\varepsilon}{m} e^{avg}(S^m, \mathfrak{M}_n(\mathbb{R}^m), B^m_\infty), \end{aligned}$$

for any m and $\varepsilon < \frac{1}{2}$. Thus

$$\begin{aligned} e^{avg}(\mathrm{Int}, \mathfrak{M}_n(C(0,1)), B(0,1)) &\geq \sup_m \frac{1}{m} e^{avg}(S^m, \mathfrak{M}_n(\mathbb{R}^m), B^m_\infty) \\ &= \sup_m \frac{1}{1 + \sqrt{\frac{n(m-1)}{m-n}}} \\ &= \frac{1}{1+\sqrt{n}}, \end{aligned}$$

which completes the proof of the theorem. □

It seems worth to discuss some of the properties of the Monte Carlo method P_n, constructed above.

First, notice that the sequence of distributions μ^m_ε can always be chosen to be symmetric. This implies that the linear method constructed in the proof above is optimal among all *affine* methods u of the form $u(x) := a + \sum_{j \in I} c_j f(\xi_j)$, $f \in B(0,1)$. To see this, take any u of the above form, $u = a + u_0, u_0 \in \mathfrak{M}_n(C(0,1))$ and compute

$$\begin{aligned} \int |\mathrm{Int}(f) - u(f)|^2 d\mu^m_\varepsilon(f) &= \int |\mathrm{Int}(f) - u_0(f) - a|^2 d\mu^m_\varepsilon(f) \\ &= \int |\mathrm{Int}(-f) - u_0(-f) - a|^2 d\mu^m_\varepsilon(f) \\ &= \int \frac{|\mathrm{Int}(f) - u_0(f) - a|^2 + |\mathrm{Int}(-f) - u_0(-f) - a|^2}{2} d\mu^m_\varepsilon(f) \\ &= \int |\mathrm{Int}(f) - u_0(f)|^2 d\mu^m_\varepsilon(f) + a^2, \end{aligned}$$

which is minimized for $a = 0$. Secondly, a look at the optimal Monte Carlo method P_n described in Theorem 4. shows that it is biased. Precisely,

$$\int u(f) dP_n(u) = nc_n \mathrm{Int}(f) = \frac{n}{n+\sqrt{n}} \mathrm{Int}(f).$$

This is a "slight" underestimation, while the usual "crude" Monte Carlo method (with $\tilde{c} = \frac{1}{n}$) is unbiased.

Finally, let us mention, that the same optimal Monte Carlo error as stated in Theorem 4 is obtained for continuous functions on the s-dimensional unit cube $[0,1]^s$, thus reproving the independence of the dimension.

References

[Erm71] S.M. Ermakov, *The Monte Carlo method and related problems*, Nauka, Moscow, 1971 (russian).

[Mat92] P. Mathé, *A minimax principle for the optimal error of Monte Carlo methods*, Constr. Approx. (1992), to appear.

[TWW88] J.F. Traub, G.W. Wasilkowski, and H. Woźniakowski, *Information-based complexity*, Academic Press, New York, 1988.

Institut für Angewandte Analysis und Stochastik, Mohrenstrasse 39, O-1086 Berlin, Federal Republic of Germany

Veröffentlichungen des Instituts für Angewandte Analysis und Stochastik

Preprints 1992

1. D.A. Dawson and J. Gärtner: Multilevel large deviations.
2. H. Gajewski: On uniqueness of solutions to the drift–diffusion–model of semiconductor devices.
3. J. Fuhrmann: On the convergence of algebraically defined multigrid methods.
4. A. Bovier and J.-M. Ghez: Spectral properties of one–dimensional Schrödinger operators with potentials generated by substitutions.
5. D.A. Dawson and K. Fleischmann: A super–Brownian motion with a single point catalyst.
6. A. Bovier, V. Gayrard: The thermodynamics of the Curie–Weiss model with random couplings.
7. W. Dahmen, S. Prößdorf, R. Schneider: Wavelet approximation methods for pseudodifferential equations I: stability and convergence.
8. A. Rathsfeld: Piecewise polynomial collocation for the double layer potential equation over polyhedral boundaries. Part I: The wedge, Part II: The cube.
9. G. Schmidt: Boundary element discretization of Poincaré–Steklov operators.
10. K. Fleischmann, F. I. Kaj: Large deviation probability for some rescaled superprocesses.