JOURNAL OF APPROXIMATION THEORY **85**, 1–15 (1996)
ARTICLE NO. 0025

On the Existence of Unbiased Monte Carlo Estimators

PETER MATHÉ

Weierstrass Institute for Applied Analysis and Stochastic, Mohrenstrasse 39, D-10117 Berlin, Germany

Communicated by Allan Pinkus

Received August 18, 1994; accepted in revised form February 22, 1995

For many typical instances where Monte Carlo methods are applied attempts were made to find unbiased estimators, since for them the Monte Carlo error reduces to the statistical error. These problems usually take values in the scalar field. If we study vector valued Monte Carlo methods, then we are confronted with the question of whether there can exist unbiased estimators. This problem is apparently new. Below it is settled precisely. Partial answers are given, indicating relations to several classes of linear operators in Banach spaces.

1. INTRODUCTION AND NOTATION

In many practical applications the program designer is confronted with the "curse of dimensionality", an exponential dependence on the dimension, which is inherent in most error estimates provided by classical numerical analysis, see e.g. [TWW88] for a sample of typical numerical problems and the respective error estimates.

Often this can be overcome by choosing Monte Carlo methods, i.e., numerical methods involving random parameters in the computational process, see [HH64] for an excellent, by now classical treatment on the applicability of Monte Carlo methods. Within the classical theory one prefers unbiased Monte Carlo estimators, since they are self-focusing if the numerical simulation is repeated.

The "crude Monte Carlo integration", cf. [HH64, Chapter 5.2], is certainly the most prominent example. Suppose we want to compute the integral $I(f) := \int_\Omega f(\omega)\, d\mu(\omega)$ for some (square integrable) function $f\colon \Omega \to \mathbb{R}$ and probability μ. We may regard the mapping f as a real valued random variable. Then the expectation of this random variable is just $I(f)$. This may be rephrased by stating that $\omega \to f(\omega)$ is an unbiased Monte Carlo method for the functional $I(f)$. The basic result within the classical theory claims, that the sample mean of n independent copies of $\omega \to f(\omega)$

0021-9045/96 $18.00

leads to a variance reduction by a factor of $1/\sqrt{n}$. In other words, the crude Monte Carlo integration just described is an unbiased Monte Carlo error of magnitude $1/\sqrt{n}$ (The precise notion of Monte Carlo error is introduced below). This error behavior is typical as long as we want to approximate a functional of the input data. However, given some functional it may often be hard to find unbiased estimators, see [KW86, ENS89].

Let us also mention recent progress on the existence of unbiased estimators (in statistical sense). While it has been known since 1956 that there cannot exist such unbiased estimates for density estimation, see [Ros56], progress has recently been made by [LB93]. The authors prove the nonexistence of informative unbiased estimators for singular problems, which exhibits the necessity of a bias-variance trade-off as an essential component for such kind of problems, see the discussion in the introduction there. Within our framework, the notion of informativity corresponds to the requirement of finite errors. On the other hand, all problems studied below are not singular in their sense. The authors emphasize in Note 1 that their results are independent of the choice of the norm in the target space. This phenomenon is no longer true within the present context of Monte Carlo methods and is one of the reasons that our arguments must be completely different, emphasizing geometric properties of the target space, where the error is measured.

Below we are concerned with *vector valued Monte Carlo methods* for the randomized approximation of *linear* mappings. The classical results extend easily from Monte Carlo methods for the approximation of linear functionals to a finite number of those, if we gather unbiased estimates for each component. The problem is how far one can proceed by this: Are there infinite dimensional problems, which admit unbiased Monte Carlo estimators?

This problem is apparently new. As an illustration let us briefly and informally introduce the following

EXAMPLE 1. Let f be any periodic function defined on the interval $[0, 2\pi]$, which has a (generalized) derivative in $\tilde{L}_2(0, 2\pi)$, the space of all periodic square integrable functions. It is well known that any such function expands into a Fourier series (converging in $\tilde{L}_2(0, 2\pi)$), which means, that we have

$$f(x) = \sum_{k=1}^{\infty} \gamma_k(f) e^{-ikx}, \qquad x \in [0, 2\pi], \tag{1}$$

with Fourier coefficients

$$\gamma_k(f) := \frac{1}{2\pi} \int_0^{2\pi} f(\xi) e^{-ik\xi} \, d\xi. \tag{2}$$

For such functions an unbiased Monte Carlo estimator (using Fourier coefficients) can be obtained in the following (trivial) way by choosing a member of the Fourier series at random. Precisely, let $p_k := 6/(\pi^2 k^2)$, $k \in \mathbb{N}$, and choose $(\pi^2 k^2/6)\ \gamma_k(f) e^{-ikx}$, $k \in \mathbb{N}$, with probability p_k, respectively. The average performance of this method yields the Fourier series of f, provided the random variable

$$k \to \frac{\pi^2 k^2}{6}\ \gamma_k(f) e^{-ikx}$$

taking values in $\tilde{L}_2(0, 2\pi)$ is integrable. But we even have the stronger square integrability from

$$\sum_{k=1}^{\infty} p_k \frac{\pi^4 k^4}{36} |\gamma_k(f)|^2 \|e^{-ikx}\|_2^2 = \frac{\pi^2}{6} \sum_{k=1}^{\infty} k^2 |\gamma_k(f)|^2 < \infty, \tag{3}$$

for functions with square integrable derivative, see [Pie87].

The above example can be considered as a special instance of approximating a diagonal mapping in the space l_2 of square summable sequences. To see this we switch from the spaces of functions to the respective spaces of Fourier coefficients in the following way. Assign any function f the sequence $x_k := k\gamma_k(f)$, $k \in \mathbb{N}$. Then the approximation of f is replaced by the approximation of $(\gamma_k(f))_{k \in \mathbb{N}}$. The differentiability assumption ensures that $\sum_{k=1}^{\infty} |x_k|^2 < \infty$. This means that we assign every sequence $(x_k)_{k \in \mathbb{N}}$ the sequence $((1/k) x_k)_{k \in \mathbb{N}}$, corresponding to a diagonal mapping $D_\sigma \colon l_2 \to l_2$.

Summarizing, we have switched from the random approximation of functions to the random approximation of a diagonal mapping between Hilbert spaces. Within this framework a theory of stochastic numerical methods is available, see [Mat91, Hei94]. The basic notion will be the notion of a (linear) Monte Carlo method. Throughout the paper we restrict ourselves to linear Monte Carlo methods only. Let X and Y be Banach spaces. Denote by $\mathfrak{L}(X, Y)$ the space of all bounded linear operators and by $\mathfrak{F}(X, Y)$ the subspace of all operators of finite rank, cf. [Pie80, Pie87] for notation from the theory of operators in Banach spaces. By $\mathfrak{F}^k(X, Y)$ we denote the subset of operators of rank at most k. Corresponding to [Mat91, Hei94, Mat94] we propose the following

Definition 1. A triple

$$\mathscr{P} := ([\Omega, \mathscr{F}, P], \mathrm{u}, \mathrm{k})$$

is called a (*linear*) *Monte Carlo method*, if

(1) $[\Omega, \mathscr{F}, P]$ is a probability space.

(2) $\mathrm{u}\colon \Omega \to \mathfrak{F}(X, Y)$ is such that the mapping $\Phi\colon X_0 \times \Omega \to Y$, defined by

$$\Phi(x, \omega) := (\mathrm{u}(\omega))(x), \qquad x \in X,\ \omega \in \Omega,$$

is product measurable into Y and the set $\{(\mathrm{u}(\omega))(x),\ x \in X,\ \omega \in \Omega\}$ is a separable subset in Y.

(3) The cardinality function $\mathrm{k}\colon \Omega \to \mathbb{N}$ is a measurable natural number, for which

$$\mathrm{u}_\omega := \mathrm{u}(\omega) \in \mathfrak{F}^{\mathrm{k}(\omega)}(X, Y), \qquad \omega \in \Omega.$$

Remark 1. For linear Monte Carlo methods as introduced above we could directly assign $\mathrm{k}(\omega) := \mathrm{rank}(\mathrm{u}(\omega))$, since this would result in an appropriate measurable cardinality function. However, for general classes of methods such an assignment would not be meaningful, so we kept the definition with the more general choice of cardinality function, see [Hei94, Mat94].

For such Monte Carlo methods we can assign the cardinality

$$\text{MC-card}(\mathscr{P}) := \int_\Omega \mathrm{k}(\omega)\, dP(\omega).$$

By definition, for every $x \in X$ the mapping

$$\omega \to \mathrm{u}_\omega(x)$$

is a Radon random variable in Y, see [LT91] for spaces of Banach space valued random variables. Therefore the error of any Monte Carlo method $\mathscr{P} := ([\Omega, \mathscr{F}, P], \mathrm{u}, \mathrm{k})$ for an operator $S\colon X \to Y$ at input $x \in X$ may be defined as

$$e(S, \mathscr{P}, x) := \left(\int_\Omega \|S(x) - \mathrm{u}_\omega(x)\|_Y^2\, dP(\omega) \right)^{1/2},$$

while the overall performance is given by

$$e(S, \mathscr{P}) := \sup_{\|x\|_X \leqslant 1} e(S, \mathscr{P}, x).$$

We agree to denote by

$$a_n^{\mathrm{mc}}(S) := \inf\left\{ \sup_{\|x\|_X \leqslant 1} \left(\int_\Omega \|S(x) - \mathrm{u}_\omega(x)\|_Y^2 \, dP(\omega) \right)^{1/2}, \ \text{MC-card}(\mathscr{P}) < n \right\}$$

the *nth Monte Carlo approximation number* of the linear operator $S\colon X \to Y$, see [Mat91] for more information on that topic. We shall make use of the following submultiplicativity property

$$a_n^{\mathrm{mc}}(RST) \leqslant \|R\| \, a_n^{\mathrm{mc}}(S) \, \|T\|, \tag{4}$$

whenever the product is correctly defined.

If a Monte Carlo method $\mathscr{P}$ has a finite error for some linear operator S, i.e., $e(S, \mathscr{P}) < \infty$, then the function

$$S_{\mathscr{P}}(x) := \int_\Omega \mathrm{u}_\omega(x) \, dP(\omega)$$

exists and denotes the respective expectation. From now on we make the assumption that

$$\sup_{\|x\|_X \leqslant 1} \int_\Omega \|\mathrm{u}_\omega(x)\|_Y^2 \, dP(\omega) < \infty,$$

ensuring that $e(0, \mathscr{P}) < \infty$, where 0 denotes the zero operator. It is the aim of this paper to study properties of $S_{\mathscr{P}}$.

DEFINITION 2. An operator $S \in \mathfrak{L}(X, Y)$ *admits an unbiased Monte Carlo method* if there is a Monte Carlo method $\mathscr{P}$ with

(1) $$\text{MC-card}(\mathscr{P}) < \infty,$$

(2) $$\sup_{\|x\|_X \leqslant 1} \int_\Omega \|\mathrm{u}_\omega(x)\|_Y^2 \, dP(\omega) < \infty,$$

and

(3) $$S(x) = S_{\mathscr{P}}(x), \qquad x \in X.$$

If $S \in \mathfrak{L}(X, Y)$ admits an unbiased Monte Carlo method then we let

$$u(S) := \inf\left\{ \sup_{\|x\|_X \leqslant 1} \left(\int_\Omega \|\mathrm{u}_\omega(x)\|_Y^2 \, dP(\omega) \right)^{1/2}, \ \mathscr{P} \text{ is unbiased for } S \right\}. \tag{5}$$

This turns into a norm and we have $\|S: X \to Y\| \leqslant u(S)$. As explained above, property (2) is equivalent to the statement that $\mathscr{P}$ has a finite error for some bounded operator acting between X and Y. Within this framework it is immediate that any finite rank operator L admits an unbiased Monte Carlo method by letting $\mathscr{P}$ be choosing L with probability 1. However, there are different ways of representing a given finite dimensional mapping, leading to different Monte Carlo methods.

EXAMPLE 2. Let $\mathrm{Id}^m: \mathbb{R}^m \to \mathbb{R}^m$ be the identity and e_j, $j=1, ..., m$, denote the unit vector basis in $\mathbb{R}^m$.

1. (*trivial*) *representation*:

$$\mathrm{Id}^m(x) = \frac{1}{m} \sum_{j=1}^{m} (m x_j) e_j, \qquad x = (x_1, ..., x_m).$$

2. (*nuclear*) *representation*:

$$\mathrm{Id}^m(x) = \frac{1}{2^m} \sum_{\varepsilon_1, ..., \varepsilon_m = \pm 1} \left(\sum_{j=1}^{m} \varepsilon_j x_j \right) \left(\sum_{j=1}^{m} \varepsilon_j e_j \right), \qquad x = (x_1, ..., x_m).$$

Observe that it is much more elaborate to find unbiased Monte Carlo methods with prescribed properties. We will not consider that problem.

2. DIAGONAL MAPPINGS $D_\sigma: l_2 \to l_q$, $1 \leqslant q \leqslant \infty$

We are going to study diagonal mappings in sequence spaces. To make things precise we need to introduce the Lorentz sequence spaces, cf. [Pie87, 2.1]. For any sequence $x = (\xi_j)_{j \in \mathbb{N}}$ of real numbers, which is convergent to 0 we assign with $(s_n(x))_{n \in \mathbb{N}}$ the non-increasing rearrangement (in modulus) of x. The (real) Lorentz space $l_{r, w}$ consists of all sequences x for which the sequence $(n^{1/r - 1/w} s_n(x))_{n \in \mathbb{N}}$ belongs to l_w equipped with the norm arising from this identification. Explicitly, if $1 \leqslant w < \infty$ then we let

$$\|x\|_{r, w} := \left(\sum_{j=1}^{\infty} [n^{1/r - 1/w} s_n(x)]^w \right)^{1/w}$$

while for $w = \infty$ we put

$$\|x\|_{r, \infty} := \sup_{n \in \mathbb{N}} \{ n^{1/r} s_n(x) \}.$$

Observe that we can identify for any $1 \leqslant p < \infty$ the classical sequence spaces l_p with $l_{p,p}$, while we denote by c_0 the space $l_{\infty,\infty}$ and by l_∞ the space of all bounded sequences equipped with the supremum norm.

We begin our study by considering a specific class of diagonal operators. Given any sequence $\sigma = (\sigma_j)_{j \in \mathbb{N}}$ of real numbers, let us consider the mapping $x = (x_j)_{j \in \mathbb{N}} \to (\sigma_j x_j)_{j \in \mathbb{N}}$. Denote this mapping by D_σ. For choices of $1 \leqslant p, q \leqslant \infty$ the operator D_σ acts continuously from l_p to l_q if and only if the diagonal σ belongs to l_r with $1/r = \max\{0, 1/q - 1/2\}$.

The simple construction outlined in the introductory section implies the existence of unbiased Monte Carlo methods for $D_\sigma\colon l_2 \to l_2$ whenever $\sigma \in l_2$, which is more than required for continuity. This is typical and we shall derive necessary conditions later. However, the class of operators admitting an unbiased Monte Carlo method can be enlarged for other spaces, employing a more involved construction. This is provided in the following

THEOREM 1. *A diagonal mapping $D_\sigma\colon l_2 \to l_q$, $q < \infty$, admits an unbiased Monte Carlo method if $\sigma \in l_q$. Moreover we have $u(D_\sigma\colon l_2 \to l_q) \leqslant \|\sigma\|_q$.*

Before proving the theorem we need a preparatory lemma. Let ε^m denote the uniform distribution on the extreme points of B_∞^m, i.e., all vectors $\omega \in \{+1, -1\}^m$. Thus the basic probability space is $[\{+1, -1\}^m, \mathscr{F}^m, \varepsilon^m]$. For computations involving ε^m the following lemma is useful.

LEMMA 1. *Let $m > 1$.*

(1) *For all $1 \leqslant j, k \leqslant m$ we have*

$$\int_{\{+1,-1\}^m} \langle \omega, e_j \rangle \langle \omega, e_k \rangle \, d\varepsilon^m(\omega) = \begin{cases} 1, & \text{if } \; j = k \\ 0, & \text{else.} \end{cases} \tag{6}$$

(2) *For all $a \in \mathbb{R}^m$ we have*

$$\left(\int_{\{+1,-1\}^m} |\langle \omega, a \rangle|^2 \, d\varepsilon^m(\omega) \right)^{1/2} = \|a\|_2.$$

Proof. The second statement is a consequence of the first one. To prove the first statement, let us observe that for $\omega \in \{+1, -1\}^m$ and $1 \leqslant j \leqslant m$ the number $\langle \omega, e_j \rangle$ is ± 1. This implies $\int_{\{+1,-1\}} \langle \omega, e_j \rangle^2 \, d\varepsilon^m(\omega) = 1$. If $j \neq k$ then

$$\begin{aligned} &\operatorname{card}\{\omega \in \{+1, -1\}^m, \langle \omega, e_j \rangle \langle \omega, e_k \rangle = 1\} \\ &\quad = \operatorname{card}\{\omega \in \{+1, -1\}^m, \langle \omega, e_j \rangle \langle \omega, e_k \rangle = -1\}, \end{aligned}$$

which is equal to 2^{m-1}, and from which the proof can be completed. ∎

Remark 2. The above lemma is well known. The reader familiar with probability theory will recognize that the distribution of $\langle \omega, a\rangle$ is the distribution of the sum $\sum_{j=1}^{m} \varepsilon_j \langle e_j, a\rangle$, where the ε_j are independent numbers taking values $+1$ and -1 with equal probability, i.e., they form a Bernoulli sequence.

Proof of Theorem 1. The proof is constructive. We are going to design a concrete Monte Carlo method of cardinality 1. Let $\Omega^m := \{+1, -1\}^m$ equipped with ε^m, introduced above as probability. Denote by $\Omega := \bigcup_{m \in \mathbb{N}} \Omega^m$ the disjoint union, equipped with the σ-algebra $\mathscr{F}$, generated from the sequence $\mathfrak{F}^m$. A probability P will be given as a mixture of ε^m in the following way.

Observe, that for $\varepsilon > 0$ we can find a decreasing sequence $(\tau_m)_{m \in \mathbb{N}}$, such that $\tau_1 = 1$ and $\sigma_m = \tau_m \beta_m$, $m \in \mathbb{N}$, $\lim_{m \to \infty} \tau_m = 0$ and $\sum_{j=1}^{\infty} |\beta_j|^q \leqslant (1+\varepsilon)^q \sum_{j=1}^{\infty} |\sigma_j|^q$. For a proof of this fact we refer to [Pie80, 8.6.4]. Let $p_m := \tau_m - \tau_{m+1}$, $m \in \mathbb{N}$. Then we have $p_m \geqslant 0$ and $\sum_{m=1}^{\infty} p_m = \tau_1 = 1$, which means that this sequence gives rise to a probability P by letting $P := \sum_{m=1}^{\infty} p_m \varepsilon^m$. (We implicitly extended the probabilities ε^m to all Ω.) So far we have defined a probability space $[\Omega, \mathscr{F}, P]$. Now, define a mapping $\mathrm{u}\colon \Omega \to \mathfrak{F}(l_2, l_q)$ by

$$\mathrm{u}_\omega(x) := \left(\sum_{j=1}^{m} \omega_j \langle x, e_j\rangle\right)\left(\sum_{j=1}^{m} \omega_j \beta_j e_j\right), \qquad \omega = (\omega_1, ..., \omega_m), \quad x \in l_2.$$

It is readily checked that $\mathscr{P} = ([\Omega, \mathscr{F}, P], \mathrm{u}, 1)$ is a linear Monte Carlo method of cardinality 1. Moreover, for $x \in l_2$ we have by Lemma 1

$$\begin{aligned}
\int_\Omega \|\mathrm{u}_\omega(x)\|_q^2 \, dP(\omega) &= \sum_{m=1}^{\infty} p_m \int_{\Omega^m} \|\mathrm{u}_\omega(x)\|_q^2 \, d\varepsilon^m(\omega) \\
&= \sum_{m=1}^{\infty} p_m \int_{\Omega^m} \left|\sum_{j=1}^{m} \omega_j \langle x, e_j\rangle\right|^2 \left\|\sum_{j=1}^{m} \omega_j \beta_j e_j\right\|_q^2 d\varepsilon^m(\omega) \\
&\leqslant \sum_{m=1}^{\infty} p_m \|x\|_2^2 \left(\sum_{j=1}^{m} |\beta_j|^q\right)^{2/q} \\
&\leqslant (1+\varepsilon)^2 \|x\|_2^2 \left(\sum_{j=1}^{\infty} |\sigma_j|^q\right)^{2/q},
\end{aligned}$$

such that $u(D_\sigma\colon l_2 \to l_q) \leqslant (\sum_{j=1}^{\infty} |\sigma_j|^q)^{1/q}$, provided $\mathscr{P}$ was unbiased for D_σ. But this is true, since

$$T(x) := \int_\Omega \mathrm{u}_\omega(x) \, dP(\omega), \qquad x \in l_2,$$

is well-defined and equal to

$$
\begin{aligned}
T(x) &= \sum_{m=1}^{\infty} p_m \int_{\Omega^m} \left(\sum_{j=1}^{m} \omega_j \langle x, e_j \rangle \right) \left(\sum_{j=1}^{m} \omega_j \beta_j e_j \right) d\varepsilon^m(\omega) \\
&= \sum_{m=1}^{m} p_m \sum_{j=1}^{m} \beta_j \langle x, e_j \rangle e_j \\
&= \sum_{j=1}^{\infty} \beta_j \left(\sum_{m=j}^{m} p_m \right) \langle x, e_j \rangle e_j \\
&= \sum_{j=1}^{\infty} \beta_j \tau_j \langle x, e_j \rangle e_j \\
&= D_\sigma(x),
\end{aligned}
$$

again, by Lemma 1. The proof of the theorem is complete. ∎

Analogously we could prove

COROLLARY 1. *A diagonal mapping* $D_\sigma \colon l_2 \to l_\infty$ *admits an unbiased Monte Carlo method provided the diagonal tends to* 0.

All operators considered so far have been compact. So the question arises, whether this is typical and leads to the following

Open Problem. Does the identity $\mathrm{Id} \colon l_2 \to l_\infty$ *admit an unbiased Monte Carlo method*?

Below we turn to the question of whether there are necessary conditions for the existence of unbiased Monte Carlo methods. Such conditions will be expressed in terms of the Monte Carlo approximation numbers, introduced above. Therefore we are looking for a reformulation of Theorem 1. To do this we need the following result, which points at the relation between the ordinary approximation numbers and their Monte Carlo counterpart, see [Mat91, Lemma 5]. Topics (2) and (3), below provide typical instances where the Monte Carlo and the (ordinary) approximation numbers differ.

PROPOSITION 1. *There is a constant* $C < \infty$ *such that for all* $m, n \in \mathbb{N}$ *we have*

(1)

$$a_n^{\mathrm{mc}}(\mathrm{Id} \colon l_2^m \to l_q^m) \leqslant C \frac{m^{1/q}}{n^{1/q'}},$$

if $1 \leqslant q \leqslant 2$ *and* $1/q' = 1 - 1/q$,

(2)

$$a_n^{\mathrm{mc}}(\mathrm{Id}\colon l_2^m \to l_q^m) \leqslant C\,\frac{m^{1/q}}{n^{1/2}},$$

if $2 \leqslant q < \infty$, *and*

(3)

$$a_n^{\mathrm{mc}}(\mathrm{Id}\colon l_2^m \to l_\infty^m) \leqslant C\,\sqrt{\frac{\log(1+m)}{n}}.$$

For the convenience of the reader we shall provide a proof, different from the one given previously in [Mat91, Hei94], thereby using the construction outlined above. We need a geometric property of Banach spaces.

DEFINITION 3. A Banach space Y has type p, $1 \leqslant p \leqslant 2$, if there is a constant $C < \infty$, such that for all probability spaces $[\Omega, \mathscr{F}, P]$, $k \in \mathbb{N}$ and independent random elements $\Upsilon_1, ..., \Upsilon_k$ in $L_2(\Omega, \mathscr{F}, P, Y)$, for which $\int_\Omega \Upsilon_j(\omega)\, dP(\omega) = 0$, $j = 1, ..., k$, we have

$$\left(\int_\Omega \left\| \sum_{j=1}^k \Upsilon_j(\omega) \right\|_Y^2 dP(\omega)\right)^{1/2} \leqslant C \left(\sum_{j=1}^k \int_\Omega \|\Upsilon_j(\omega)\|_Y^p\, dP(\omega)\right)^{1/p}. \qquad (7)$$

The smallest constant satisfying the above inequality shall be called the *type-p-constant* and is denoted by $T_p(Y)$.

Remark 3. The notion of a type of a Banach space was originally introduced in [MP76], where this definition was given in terms of Rademacher sequences. But, as can be seen easily, it can be extended to the above situation, albeit the type-p-constant is different by a factor of at most 2, see [LT91, Chapter 9.2], but also [HJ74]. It can be seen that $T_2(\mathbb{R}) = 1$. Moreover, every finite-dimensional Banach space has type p, $1 \leqslant p \leqslant 2$, although the respective constant may depend on the dimension, cf. [TJ88, Chapter 1, §4]. However, we have

$$T_q(l_q^m) \leqslant C, \qquad \text{if} \quad 1 \leqslant q \leqslant 2$$

and

$$T_2(l_q^m) \leqslant C \begin{cases} 1 & \text{if} \quad 2 \leqslant q < \infty \\ \sqrt{\log(1+m)}, & \text{if} \quad q = \infty \end{cases}$$

for some universal constant $C < \infty$.

Now we are prepared to provide the proof of Proposition 1, following the one given in [Mat94].

Proof of Proposition 1. We shall carry out the estimates only for $2 \leqslant q < \infty$. The other cases follow the same lines. By Theorem 1 there is an unbiased Monte Carlo method $\mathscr{P} = ([\Omega, \mathscr{F}, P], u, 1)$ leading to $u(\mathrm{Id}\colon l_2^m \to l_q^m) \leqslant m^{1/q}$. The sample mean

$$\mathscr{P}^n = ([\Omega^n, \mathscr{F}^n, P^n], \mathrm{v}, n)$$

of n independent copies of $\mathscr{P}$, which is defined by

$$\mathrm{v}_{\omega^n}(x) := \frac{1}{n} \sum_{j=1}^{n} \mathrm{u}_{\omega_j}(x), \qquad x \in l_2^m, \quad \omega^n := (\omega_1, ..., \omega_n) \in \Omega^n,$$

provides another unbiased Monte Carlo method, this time of cardinality n, which has an error at $x \in l_2^m$, $\|x\|_2 \leqslant 1$, of

$$e(\mathrm{Id}\colon l_2^m \to l_q^m, \mathscr{P}^n, x) = \left(\int_\Omega \left\| x - \frac{1}{n} \sum_{j=1}^{n} \mathrm{u}_{\omega_j}(x) \right\|_q^2 dP^n(\omega^n) \right)^{1/2} \tag{8}$$

$$= \frac{1}{n} \left(\int_\Omega \left\| \sum_{j=1}^{n} (x - \mathrm{u}_{\omega_j}(x)) \right\|_q^2 dP^n(\omega^n) \right)^{1/2} \tag{9}$$

$$\leqslant \frac{1}{\sqrt{n}} T_2(l_q^m)\, e(\mathrm{Id}\colon l_2^m \to l_q^m, \mathscr{P}). \tag{10}$$

Since

$$e(\mathrm{Id}\colon l_2^m \to l_q^m, \mathscr{P}) \leqslant \|\mathrm{Id}\colon l_2^m \to l_q^m\| + u(\mathrm{Id}\colon l_2^m \to l_q^m),$$

the proof can be completed in case $2 \leqslant q < \infty$. ∎

Now we are able to provide the reformulation of Theorem 1 in terms of the respective Monte Carlo approximation numbers $a_n^{\mathrm{mc}}(D_\sigma\colon l_2 \to l_q)$, as promised above.

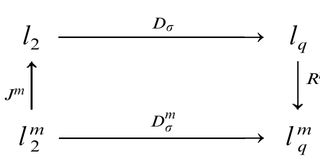
Corollary 2. *Let* $1 \leqslant q < \infty$ *and put* $r := \max\{2, q'\}$. *If*

$$(a_n^{\mathrm{mc}}(D_\sigma\colon l_2 \to l_q))_{n \in \mathbb{N}} \in l_{r,q}$$

then the mapping $D_\sigma\colon l_2 \to l_q$ *admits an unbiased Monte Carlo method.*

Proof. To simplify the proof we shall assume that the diagonal is nonnegative and nonincreasing.

Since the following diagram is commutative (where R^m assigns every vector in l_q the vector with the first m coordinates, J^m denotes the natural embedding and D_σ^m is the corresponding restriction),



and $\|J^m\| \leqslant 1$, $\|R^m\| \leqslant 1$, we have by inequality (4) the estimate

$$a_n^{\mathrm{mc}}(D_\sigma^m) \leqslant a_n^{\mathrm{mc}}(D_\sigma).$$

Moreover, we have

$$\sigma_m a_n^{\mathrm{mc}}(\mathrm{Id}_{2,q}^m \colon l_2^m \to l_q^m) \leqslant a_n(D_\sigma^m), \tag{11}$$

which can be seen as follows, cf. also [Pie87, 2.9.3]. Without loss of generality we may assume $\sigma_m > 0$. Since

$$\mathrm{Id}_{2,q}^m = (D_\sigma^m \colon l_2^m \to l_q^m) \cdot (D_\sigma^m \colon l_2^m \to l_2^m)^{-1}$$

and

$$\|(D_\sigma^m \colon l_2^m \to l_2^m)^{-1}\| = \sigma_m^{-1}$$

the estimate (11) follows immediately.

Letting $m = 2n - 1$ and inserting the results from Proposition 1 we see that the sequence $(\sigma_{2n-1})_{n \in \mathbb{N}}$ belongs to l_q whenever the assumptions from Corollary 2 are fulfilled. But this is equivalent to $(\sigma_n)_{n \in \mathbb{N}} \in l_q$, see e.g. [Pie87, Prop. 2.1.9]. ∎

We turn to the question whether there are non-trivial necessary conditions to be imposed on an operator in order to admit an unbiased Monte Carlo method. In terms of the Monte Carlo approximation numbers a fairly general condition can be given, provided the target space has some type p. Indeed, the sample mean construction from the proof of Proposition 1 implies

THEOREM 2. *Suppose the Banach space Y has type p, $p > 1$. If an operator $S \colon X \to Y$ admits an unbiased Monte Carlo method then necessarily $(a_n^{\mathrm{mc}}(S))_{n \in \mathbb{N}} \in l_{p', \infty}$.*

Taking into account the behavior of the Monte Carlo approximation numbers of the diagonal mappings and applying a technique similar to the one of Corollary 2 this transfers to

COROLLARY 3. *Let $1 \leqslant q < \infty$. If $D_\sigma \colon l_2 \to l_q$ admits an unbiased Monte Carlo method then $\sigma \in l_{q, \infty}$.*

Remark 4. A look at the necessary and sufficient conditions proves that there is only a small gap left. Since the arguments to prove the necessary conditions are very rough we conjecture that the sufficient conditions are sharp.

On the class of operators acting between Hilbert spaces we immediately obtain

THEOREM 3. *Any Hilbert–Schmidt operator (acting between Hilbert spaces) admits an unbiased Monte Carlo method. Conversely, if an operator between Hilbert spaces admits an unbiased Monte Carlo method then necessarily the sequence of singular numbers belongs to* $l_{2,\infty}$

Indeed, The norm u, as defined in (5), in easily seen to be unitarily invariant, such that the result for diagonal operators implies the respective result for arbitrary ones using the Schmidt-representation, see [Pie87, 2.11.4]. The class of Hilbert–Schmidt operators corresponds to the class of operators having square summable singular numbers, see [Pie80, 15.5.5], these corresponding to diagonal mappings having square summable diagonal.

3. SOME RELATIONS TO OPERATOR IDEALS

As could be seen in Theorem 3, Hilbert–Schmidt operators admit an unbiased Monte Carlo method. The class of such operators is closely related to the theory of operator ideals as developed in [Pie80]. The first important (though simple) observation is

THEOREM 4. *The class of all operators admitting an unbiased Monte Carlo method turns into a normed operator ideal by letting* $S \to u(S)$ *be the norm as defined in equation* (5).

This immediately implies

COROLLARY 4. *Nuclear operators admit an unbiased Monte Carlo method.*

We shall introduce this class of operators below. The assertion of the corollary follows from the fact that the ideal of the nuclear operators is the smallest normed ideal. However, a direct proof of the corollary could also be given using the trivial unbiased representation known from the introductory example.

Next we shall see that the class of operators admitting an unbiased Monte Carlo method can be enlarged considerably in many situations. We shall examplify this by introducing the ideal of (r, p, q)-nuclear operators.

Let $0 < r \leqslant \infty$, $1 \leqslant p$, $q \leqslant \infty$ with $1 + 1/r \geqslant 1/p + 1/q$. Denote by $\mathfrak{N}_{(r,p,q)}$ the ideal of (r, p, q)-nuclear operators, i.e., an operator $S \in \mathfrak{L}(X, Y)$ is (r, p, q)-nuclear, if there is a representation

$$Sx = \sum_{j=1}^{\infty} \sigma_j \langle x, a_j \rangle \, y_j, \qquad x \in X,$$

with $a_j \in X'$ and $y_j \in Y$ satisfying

$$\|\sigma\|_r < \infty,$$

$$w_{q'}((a_j)_{j \in \mathbb{N}}) := \sup_{\|x''\| \leqslant 1} \left(\sum_{j=1}^{\infty} |\langle x'', a_j \rangle|^{q'} \right)^{1/q'} < \infty$$

and

$$w_{p'}((y_j)_{j \in \mathbb{N}}) := \sup_{\|b\| \leqslant 1} \left(\sum_{j=1}^{\infty} |\langle b, y_j \rangle|^{p'} \right)^{1/p'} < \infty.$$

Denote the respective quasi norm by

$$n_{(r, p, q)}(S) := \inf \|\sigma\|_r \, w_{q'}((a_j)_{j \in \mathbb{N}}) \, w_{p'}((y_j)_{j \in \mathbb{N}}),$$

where the infimum is taken over all possible representations, see [Pie80, 18.1] for more details. In particular we obtain as $\mathfrak{N}_{(1, 1, 1)}$ the ideal of nuclear operators mentioned above. The ideal of (r, p, q)-nuclear operators can be characterized by a factorization property, see [Pie80, 18.1.3]. In fact every $(q, q, 2)$-nuclear operator factors through a diagonal mapping from l_2 to l_q with diagonal belonging to l_q. (In case $q = \infty$ the diagonal converges to 0.) Together with Theorem 1 and Corollary 1 this provides

COROLLARY 5. *Let $1 \leqslant q \leqslant \infty$. Every $(q, q, 2)$-nuclear operator S admits an unbiased Monte Carlo method. Moreover we have*

$$u(S) \leqslant n_{(q, q, 2)}(S).$$

The largest ideal of this type which admits unbiased Monte Carlo methods is obtained as $\mathfrak{N}_{(\infty, \infty, 2)}$.

For operators acting between Hilbert spaces H and K it is readily checked that

$$\mathfrak{N}_{(\infty, \infty, 2)}(H, K) = \mathfrak{N}_{(2, 2, 2)}(H, K)$$

and we obtain the Hilbert–Schmidt operators once more, see [Pie80, 18.5.4].

REFERENCES

[ENS89] S. M. ERMAKOV, W. W. NEKRUTKIN, AND A. S. SIPIN, "Random Processes for Classical Equations of Mathematical Physics," Mathematics and its Applications (Soviet Series), Vol. 34, Reidel, Dordrecht, 1989.

[Hei94] S. Heinrich, Random approximation in numerical analysis, *in* "Functional Analysis: Proceedings of the Essen Conference" (K. D. Bierstedt *et al.*, Eds.), Lecture Notes in Pure and Appl. Math., Vol. 150, pp. 123–171, Dekker, New York/Basel/Hong Kong, 1994.

[H64] J. M. Hammersley and D. C. Handscomb, "Monte Carlo Methods," Methuen, London, 1964.

[HJ74] J. Hoffmann-Jørgensen, Sums of independent Banach space valued random variables, *Studia Math.* **52** (1974), 159–186.

[KW86] M. H. Kalos and P. A. Whitlock, "Monte Carlo Methods," Vol. 1, Wiley, New York, 1986.

[LB89] R. C. Liu and L. D. Brown, Nonexistence of informative unbiased estimators in singular problems, *Ann. Statist.* **21**, No. 1 (1993), 1–13.

[LT91] M. Ledoux and M. Talagrand, "Probability in Banach Spaces," *in* "Ergebnisse der Mathematik und ihrer Grenzgebiete, *3. Folge*," Vol. 23, Springer-Verlag, Berlin/Heidelberg/New York, 1991.

[Mat91] P. Mathé, Random approximation of Sobolev embeddings, *J. Complexity* **7** (1991), 261–281.

[Mat94] P. Mathé, "Approximation Theory of Monte Carlo Methods," Habilitation thesis, 1994.

[MP76] B. Maurey and G. Pisier, Séries de variables aléatoires vectorielles indépendantes et géométrie des espaces de Banach, *Studia Math.* **58** (1976), 45–90.

[Pie80] A. Pietsch, "Operator Ideals," North-Holland, Amsterdam/New York/Oxford, 1980.

[Pie87] A. Pietsch, "Eigenvalues and *s*-Numbers," Mathematik und ihre Anwendungen in Physik und Technik, Vol. 43, Geest & Portig, Leipzig, 1987.

[Ros56] M. Rosenblatt, Remarks on nonparametric estimates of a density function, *Ann. Math. Statist.* **27** (1956), 832–837.

[TJ88] N. Tomczak-Jaegermann, "Banach-Mauar Distances and Finite-Dimensional Operator Ideals," Pitman Monographs and Surveys in Pure and Applied Mathematics, Vol. 38, Longman, Essex, 1988.

[TWW88] J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski, "Information-Based Complexity," Academic Press, New York, 1988.