

Provided for non-commercial research and educational use.
Not for reproduction, distribution or commercial use.

Mathematica Balkanica

Mathematical Society of South-Eastern Europe
A quarterly published by
the Bulgarian Academy of Sciences – National Committee for Mathematics

The attached copy is furnished for non-commercial research and education use only. Authors are permitted to post this version of the article to their personal websites or institutional repositories and to share with other researchers in the form of electronic reprints.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to third party websites are prohibited.

For further information on Mathematica Balkanica visit the website of the journal
<http://www.mathbalkanica.info>

or contact:

Mathematica Balkanica - Editorial Office;
Acad. G. Bonchev str., Bl. 25A, 1113 Sofia, Bulgaria
Phone: +359-2-979-6311, Fax: +359-2-870-7273,
E-mail: balmat@bas.bg

Adaptive Monte Carlo Methods for Numerical Integration ¹

Aneta Karaivanova

Presented by Bl. Sendov

A new adaptive techniques of Monte Carlo integration is proposed and studied. An error analysis is realized. A superconvergent adaptive method is presented. The method combines the idea of separation of the domain into *uniformly small* subdomains with the Kahn approach of *important sampling*. An estimation of the probable error for functions with bounded derivative is proved. This estimate improves the existing results. A simple adaptive Monte Carlo method is also considered. It is shown that for large dimensions d the convergence of the superconvergent adaptive Monte Carlo method goes asymptotically to $O(n^{1/2})$, which corresponds to the convergence of the simple adaptive method.

Both adaptive methods - superconvergent and simple - are applied for calculating multidimensional integrals. Numerical tests are performed on the supercomputer CRAY Y-MP C92A. It is shown that for low dimensions (up to $d = 5$) the superconvergent adaptive method gives better results than the simple adaptive method. When the dimension increases the simple adaptive method becomes better. One needs several seconds for evaluating 30-d integrals by simple adaptive method, while the solution of the same integral by Gaussian quadrature will need more than 10^6 billion years if CRAY Y-MP C92A will be used.

1. Introduction

Monte Carlo methods are a powerful tool in many fields of mathematics, physics and engineering. It is known, that these methods give statistical estimates for the functional of the solution by performing random sampling of a certain chance variable whose mathematical expectation is the desired functional.

¹Partially supported by the Ministry of Education and Science of Bulgaria, under grants I 501/95 and MM 449/94.

Monte Carlo methods have proved to be very efficient in solving multidimensional integrals in composite domains [So73], [DT93a], [Ha66]. The problem of evaluating integrals of high dimension is very important since it appears in many applications of control theory, statistical physics and mathematical economics. For example, one of numerical approaches for solution stochastic systems in control theory leads to a large number of multi-dimensional integrals with dimensionality up to $d = 30$.

It is also shown that for some problems (including one-dimensional ones) in the corresponding functional spaces Monte Carlo methods have better convergence rate than the optimal deterministic methods in such functional spaces [Ni88], [SAK94], [Ba64], [DT89], [DT93].

An important advantage of these methods is that Monte Carlo ones allow to find directly the unknown functional of the problem solution with a number of operations, necessary to solve the problem in one point of the domain [So73], [DT93].

It is well known that Monte Carlo methods are very efficient when parallel processors or parallel computers are available. This is because these methods are inherently parallel and have loose dependencies. In addition, they are well vectorizable by using powerful vector computers like CRAY Y-MP C92A.

2. General description of the Monte Carlo method

Usually Monte Carlo method reduces the problems to the approximate calculation of mathematical expectation values. Let the scalar variable J be the desired solution of the integral

$$(1) \quad J = \int_D f(x)p(x) dx,$$

where $f(x)$ is an integrable function, $x \in D \subset \mathbf{R}^d$ and $p(x) \geq 0$ is a probability density function, such that $\int_D p(x)dx = 1$.

Consider a random point $\xi \in D$ with a density $p(x)$ and let there be n realizations of the random point $\xi_i (i = 1, 2, \dots, n)$. Let a random variable $\theta(\xi)$ be defined in D , such that

$$(2) \quad E\theta(\xi) = J.$$

By definition $E\theta$ exists if and only if $E|\theta|$ exists. It is easy to see, that (2) is fulfilled if $\theta(\xi) = f(\xi)$. The computational problem consists of calculating repeated realizations of θ and of combining them into an appropriate statistical estimator of J . As approximate value of J ,

$$(3) \quad J \approx \frac{1}{n} \sum_{s=1}^n (\theta)_s = \hat{\theta}_n,$$

is set up, where $(\theta)_s$ is the s -th realization of the random variable θ .

Suppose that the random variable θ has a finite dispersion

$$(4) \quad D\theta = \sigma^2(\theta) = E(\theta - E\theta)^2 = E\theta^2 - (E\theta)^2,$$

where $\sigma(\theta)$ is the standard deviation.

The probable error for the usual Monte Carlo method (which does not use any a priori information about the smoothness of $f(x)$) [So73] is defined as:

$$(5) \quad r_n = c_{0.5}\sigma(\theta)n^{-1/2},$$

where $c_{0.5} \approx 0.6745$.

An superconvergent Monte Carlo method is a method for which

$$(6) \quad r_n = cn^{-1/2-\varepsilon(d)},$$

where c is a constant and $\varepsilon(d) > 0$, see [So73], [DT89].

A method of this type is proposed by Dupach [Du56]. The idea of the method consists in the following: the domain D is separated (divided) into subdomains D_j that are uniformly small according both to the probability and to the sizes. That means that there exist constants c_1 and c_2 such that:

$$(7) \quad D = \sum_{j=1}^m D_j, \quad D_i \cap D_j = 0, \quad i \neq j;$$

$$(8) \quad p_j = \int_{D_j} p(x) dx \leq c_1/n;$$

$$(9) \quad \sup_{x_1, x_2 \in D_j} |x_1 - x_2| \leq c_2/n^{1/d}.$$

One random point is generated in each subdomain D_j . Based on this point a corresponding realization of the random variable is evaluated. In case of higher smoothness of the solution this procedure leads to Monte Carlo method with higher rate of convergence, i.e. the probable error can be expressed by the formula (6). This problem was considered by Dimov and Tonev [DT89] with implementation of the τ -moduli [SP88]. Their result is the estimation

$$(10) \quad r_n \leq \text{const} \times \tau(f; d)_{L_2} n^{-1/2-1/d}.$$

The result [DT89] was improved by Takev [Ta92]. His result is:

$$(11) \quad r_n = O(\omega(f, n^{-1})_{L_2})n^{-1/2}.$$

Another degree of quality of the Monte Carlo method is the dispersion of the random variable θ , whose mathematical expectation coincides with J . Let θ be a random variable in the usual Monte Carlo method [So73] such that

$$(12) \quad J = E\theta.$$

Let $\hat{\theta}$ be another random variable for which

$$(13) \quad J = E\hat{\theta}$$

and the conditions providing the existence and the finiteness of the dispersion $D\hat{\theta}$ be fulfilled.

The method for which

$$(14) \quad D\hat{\theta} < D\theta$$

is called "efficient Monte Carlo method" [Ka50], [Mi87], [Di91], [DT93]. A method of this type is proposed by Kahn [Ka50] (for evaluation of integrals) and by Mikhailov and Dimov (for evaluation of integral equations) [Mi87], [Di91].

This work deals with adaptive Monte Carlo methods, which use a priori and a posteriori information obtained during calculations. Both approaches - superconvergent adaptive approach and simple adaptive approach are applied. The works having studied superconvergent Monte Carlo methods show that the separation of the domain into *uniformly small* subdomains brings to an increasing of the rate of convergence. But this separation does not use any a priori information about parameters of the problem. The *Kahn approach* and the approach in [Mi87], [Di91] use the information about the problem parameters, but do not use the idea of *separation*. In this paper a superconvergent method which uses both the idea of *separation* and the idea of *important sampling* is presented. This method is called *superconvergent adaptive Monte Carlo* (SAMC) method. A *simple adaptive Monte Carlo* (AMC) method is also presented and studied.

The paper is organized as follows. Section 2 contains the general description of Monte Carlo methods and some basic facts about superconvergent and efficient Monte Carlo methods. Section 3 contains theoretical results connected to SAMC methods. This method uses both the idea of *separation* and the idea of *important sampling*. In fact, the proposed method is an superconvergent Monte Carlo method with the probable error of type $c \times n^{-1/2-\epsilon(d)}$, but the constant before $N^{-1/2-\epsilon(d)}$ is smaller than the constant of the usual superconvergent Monte Carlo method. In section 4 two adaptive Monte Carlo methods - simple and superconvergent - are implemented. This section contains numerical results

of evaluating multi-dimensional integrals. It is shown that for low dimensions (up to $d = 4$) the superconvergent adaptive Monte Carlo method gives better results than the simple adaptive method.

3. Superconvergent adaptive Monte Carlo method and error estimation

The following problem arises:

Problem 1. *Is it possible to combine the idea of separation of the domain into "uniformly small" subdomains with the Kahn approach of "important sampling"?*

First, consider the one-dimensional problem of evaluation integrals:

$$(15) \quad J = \int_D f(x)p(x) dx, \quad D \equiv [0, 1],$$

where $f(x) \in H(1, L)_{[0,1]}$ and $\int_D p(x)dx = 1$.

Partition $[0,1]$ into m subdomains such that $m \leq n$:

$$(16) \quad x_0 = 0; \quad x_m = 1;$$

$$(17) \quad C_i = 1/2[f(x_{i-1}) + f(1)](1 - x_{i-1}),$$

$$(18) \quad i = 1, \dots, m - 1;$$

$$(19) \quad x_i = \frac{C_i}{f(x_{i-1})(n - i + 1)}, \quad i = 1, \dots, m - 1$$

$$(20) \quad D_i \equiv [x_{i-1}, x_i]$$

The scheme (19) gives us an important separation of the domain $D \equiv [0, 1]$ (see Fig.1). We have:

$$(21) \quad J = \int_0^1 f(x)p(x) dx$$

$$(22) \quad = \sum_{i=1}^m \int_{x_{i-1}}^{x_i} f(x)p(x) dx.$$

Denote by p_i and J_i the following expressions:

$$(23) \quad p_i = \int_{x_{i-1}}^{x_i} p(x) dx$$

and

$$(24) \quad J_i = \int_{x_{i-1}}^{x_i} f(x)p(x) dx.$$

Obviously,

$$(25) \quad \sum_{i=1}^m p_i = 1; \quad \sum_{i=1}^m J_i = J.$$

Consider now a random variable $\xi^{(i)} \in D_i$ with a density function $p(x)/p_i$, where $D_i \equiv [x_{i-1}, x_i]$. In this case

$$(26) \quad J_i = E(p_i f(\xi^{(i)})).$$

Let n_i be a number of random points in D_i ($\sum_{i=1}^m n_i = n$).

It is easy to show that

$$(27) \quad J_i = E \left[\frac{p_i}{n_i} \sum_{s=1}^{n_i} f(\xi_s^{(i)}) \right] = E\theta_{n_i};$$

$$(28) \quad J = E \left[\sum_{i=1}^m \frac{p_i}{n_i} \sum_{s=1}^{n_i} f(\xi_s^{(i)}) \right] = E\theta_n^*.$$

Let $n_i = 1$ (so, $m = n$). Since $f(x) \in H(1, L)_{[0,1]}$, there exist constants L_i , such that

$$(29) \quad L_i \geq \left| \frac{\partial f}{\partial x} \right| \text{ for any } x \in D_i.$$

Moreover, for our scheme there exist constants c_1 , and c_2 , such that

$$(30) \quad p_i = \int_{D_i} p(x) dx \leq c_1/n, \quad i = 1, \dots, n$$

and

$$(31) \quad \sup_{x_1, x_2 \in D_i} |x_1 - x_2| \leq c_2/n, \quad i = 1, \dots, n.$$

We shall say, that the conditions (29)-(31) define an *important separation* of D .

Theorem 1. Let $f(x) \in H(1, L)_{[0,1]}$ and $m = n$. Then for the *important separation* of D

$$r_n \leq \sqrt{2} [1/n \sum_{j=1}^n (L_j c_{1j} c_{2j})^2]^{1/2} n^{-3/2}.$$

Proof. Let D_j be any subdomain of $[0, 1]$. For a fixed point $s_j \in D_j$ we have:

$$(32) \quad f(s_j) = f(s_j) + f'(\eta_j)(\xi_j - s_j),$$

where $\eta_j \in D_j$.

Since $f'(\eta_j) \leq L_j$ we have:

$$(33) \quad Df(\xi_j) \leq Ef^2(\xi_j) \leq L_j^2 E(\xi_j - s_j)^2$$

$$(34) \quad \leq L_j^2 \sup_{x_1, x_2 \in D_j} |x_1 - x_2|^2 \leq L_j^2 c_{2j}^2 / n^2.$$

Now $D\theta_n^*$ can be estimated:

$$(35) \quad D\theta_n^* = \sum_{j=1}^n p_j^2 Df(\xi_j) \leq \sum_{j=1}^n (c_{1j}^2 n^{-2} L_j^2 c_{2j}^2 n^{-2})$$

$$(36) \quad = 1/n \sum_{j=1}^n (L_j c_{1j} c_{2j})^2 n^{-3}.$$

To estimate the probable error one can apply the Chebyshev inequality:

$$(37) \quad \Pr\{|\theta_n^* - E\theta_n^*| < h\} \geq 1 - (D\theta_n^*/h^2),$$

where $h > 0$.

Let

$$(38) \quad h = 1/\varepsilon (D\theta_n^*)^{1/2},$$

where ε is a positive number.

For $\varepsilon = 1/\sqrt{2}$ we have:

$$(39) \quad \Pr\{|\theta_n^* - J| < \sqrt{2}(D\theta_n^*)^{1/2}\} \geq 1/2.$$

The last inequality proves the theorem, since

$$(40) \quad r_n \leq \sqrt{2}(D\theta_n^*)^{1/2}$$

$$(41) \quad \sqrt{2} \left(\frac{1}{n} \sum_{j=1}^n (L_j c_{1j} c_{2j})^2 \right)^{1/2} n^{-3/2}.$$

This result presents a superconvergent adaptive Monte Carlo method. Moreover, the constants $\sqrt{2} \left(\frac{1}{n} \sum_{j=1}^n (L_j c_{1j} c_{2j})^2 \right)^{1/2}$ in (41) is smaller than the constant in the methods of Dupach type [Du56], [DT89].

Now consider multi-dimensional integrals:

$$(42) \quad J = \int_D f(x) p(x) dx, \quad x \in D \subset \mathbf{R}^d,$$

where $f \in H(1, L)_D$.

Assume it is possible to have an important separation of the domain D . Let

$$(43) \quad f(x) \in H(1, L_i)_{D_i}, \quad \text{for any } x \in D_i$$

The following statement is fulfilled:

Theorem 2. *Let there exist an important separation of D such that (43) be fulfilled and $m = n$. Then*

$$r_n \leq \sqrt{2}d \left[\frac{1}{n} \sum_{i=1}^n (L_i c_{1i} c_{2i})^2 \right]^{1/2} n^{-1/2-1/d}.$$

The proof of this theorem follows the same techniques as the proof of Theorem 1.

4. Implementation of adaptive Monte Carlo methods. Numerical tests

The algorithms are realized on CRAY Y-MP C92A machine with two vector processors. Two Monte Carlo methods are considered.

1. Simple adaptive Monte Carlo method

This approach does not use any a priori information about the smoothness of the integrand. It deals with n uniformly distributed random points $x_i \in [0, 1]^d$, $i = 1, \dots, n$ into d -dimensional cube $[0, 1]^d$. For generating of any point, d uniformly distributed random numbers into interval $[0, 1]$ are produced. The algorithm is adaptive: it starts with a relatively small number n , which is given as an input data. During the calculations the variance on each dimension coordinate is estimated. The above mentioned information is used for increasing the density of the new generated points. This approach leads to the following estimation

$$(44) \quad \varepsilon \leq c \frac{1}{n^{1/2}}$$

instead of standard estimation for the probable error

$$(45) \quad \varepsilon \leq 0.6745\sigma(\theta) \frac{1}{n^{1/2}},$$

where

$$(46) \quad c \leq 0.6745\sigma(\theta).$$

2. Superconvergent adaptive Monte Carlo method

As a first step of this method the domain of integration is separated into subdomains with identical volume. For every subdomain the integral J_j is evaluated and an a posteriori information for the dispersion is also received. After that an approximation for the integral $J = \sum_j J_j$ and the total dispersion is obtained. The total dispersion is compared with local a posteriori estimations for the dispersion in every subdomain. The obtained information is used for the next refinement of the domain and for increasing the density of the random points.

The probable error for this approach has the following form

$$(47) \quad \varepsilon \leq cn^{-\frac{1}{2}-\frac{1}{d}}.$$

Obviously, for a large dimension d , the convergence goes asymptotically to $O(n^{-1/2})$.

Since, the approach realized SAMC method is more time consuming, for large dimensions d it is better to use AMC method. This was observed during the performed calculations. But for relatively small d ($d = 1, \dots, 5$) and for "bad" functions (say, functions with bounded first derivative) SAMC method successfully competes with the standard Gaussian quadratures.

The both methods have been performed for evaluating integrals of different dimensions and integrands (both with large and small variance, as well as high and low smoothness).

Here some results of numerical experiments are given. We will present three examples (the exact values of integrals are presented to be able to compare the real error with the a posteriori estimated error).

Example 1. $n = 4$

$$(48) \quad J_1 = \int_0^1 \int_0^1 \int_0^1 \int_0^1 \frac{4x_1 x_3^2 \exp 2x_1 x_2}{(1 + x_2 + x_4)^2} dx_1 dx_2 dx_3 dx_4 = 0.5753.$$

Example 2. $n = 25$

$$(49) \quad J_2 = \int_0^1 \dots \int_0^1 \frac{4x_1 x_3^2 \exp 2x_1 x_3}{(1 + x_2 + x_4)^2} \exp(x_5 + \dots + x_{20})$$

$$(50) \quad \times x_{21} x_{22} \dots x_{25} dx_1 \dots dx_{25} = 103.8.$$

Example 3. $n = 30$

$$(51) \quad J_3 = \int_0^1 \dots \int_0^1 \frac{4x_1 x_3^2 \exp 2x_1 x_3}{(1 + x_2 + x_4)^2} \exp(x_5 + \dots + x_{20})$$

$$(52) \quad \times x_{21} x_{22} \dots x_{30} dx_1 \dots dx_{30} = 3.244.$$

Some results are presented in Table 1.

Table 1. Results of numerical experiments performed on CRAY Y-MP C92A.

<i>Dim.</i> <i>d</i>	<i>Exact</i> <i>sol.</i>	<i>Method</i>	<i>Calcul.</i> <i>solution</i>	<i>CP - time,</i> <i>s</i>	<i>Estim.</i> <i>error</i>	<i>Rel.</i> <i>error</i>	<i>Num.of</i> <i>points</i>
4	0.5753	AMC	0.5763	0.184	0.025	0.0017	20000
4	0.5753	SAMC	0.5755	0.025	$0.82 \cdot 10^{-2}$	0.0003	1728
25	103.8	AMC	107.66	3.338	0.05	0.036	10^5
25	103.6	AMC	104.6	17.2	0.024	0.0077	$5 \cdot 10^5$
25	103.8	AMC	103.1	54.9	0.017	0.0069	10^6
30	3.244	AMC	3.365	4.07	0.099	0.037	10^5
30	3.244	AMC	3.551	0.879	0.23	0.095	$2 \cdot 10^4$
30	3.244	AMC	2.380	0.096	0.43	0.266	$2 \cdot 10^3$

Table 1 contains an information about the dimension of the integral, the exact solution, the applied Monte Carlo approach, calculated solution, CP-time, estimated error, relative error really obtained and number of random points.

Some of numerical results are presented on Fig. 2–7. Fig. 2 shows the results of implementation of SAMC method for solving 5-d integral. The dependence of the calculated values and the error from the number of random points is presented. In this case about 100 random points are sufficient for obtaining a relatively good accuracy. On Fig. 3 results for SAMC and AMC for 5-d integral are compared. One can see that the error of SAMC is much smaller than the error of AMC. Fig. 4 presents results of calculated values and the error for 10-d integral for different numbers of random points. Fig. 5 expresses the strip of estimated error. One can see that the calculated values are close to the exact solution which is 14.808 and lie inside of the error strip. Fig. 6 shows the calculated values and estimated errors for 20-d integral for different number of random points for AMC method. Some numerical results for implementation of AMC method for 30-d integral are presented on Fig. 7. It is clear that in this case the estimated error can not be very small. Nevertheless, such accuracy is sufficient for applications which are considered in control theory.

5. Conclusion

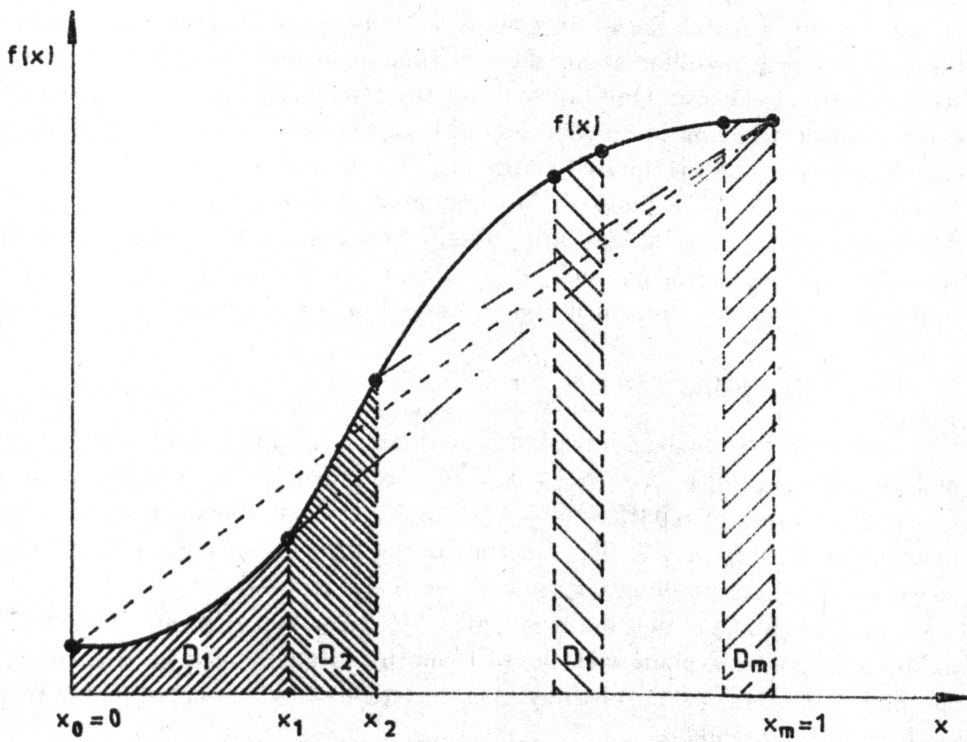
1. For low dimensions SAMC method gives better results than AMC method. For example, for $d = 4$, SAMC method needs 1728 realizations for reaching an error of 0.03%, while AMC method needs 20000 realizations for reaching an error of 0.17%. The CP-time of the SAMC method is about 7 times less than the corresponding CP-time of the AMC method.

2. When the dimension d is high AMC gives better results than SAMC method. It can be explain with the fact that for large dimensions d the error of the SAMC method asymptotically goes to $O(n^{-1/2})$ which corresponds to the error of AMC method.

3. It is very important that Monte Carlo method permits to receive the results of numerical integrating of high-dimensional integrals with a relatively good accuracy. For example, the value of 25-d integral can be computed for 54.9 s only with an error, which is less than 1%. The same integral can be evaluated for 3.34 s with an error of 3.6%. For evaluation of 30-d integral with an error of 3.7% CP-time of 4.07 s is needed. For reaching the same accuracy with quadrature formula of Gaussian type one needs at least 10 nodes on each direction, which means that 10^{30} values of the integrand have to be calculated. It means that 10^{23} s or more than 10^6 billion years are needed if supercomputer CRAY Y-MP C92A is used.

6. Acknowledgments

The author thanks Ivan Dimov for the helpful discussions.



Scheme of an important separation of the interval $D \equiv [0,1]$

Fig. 1

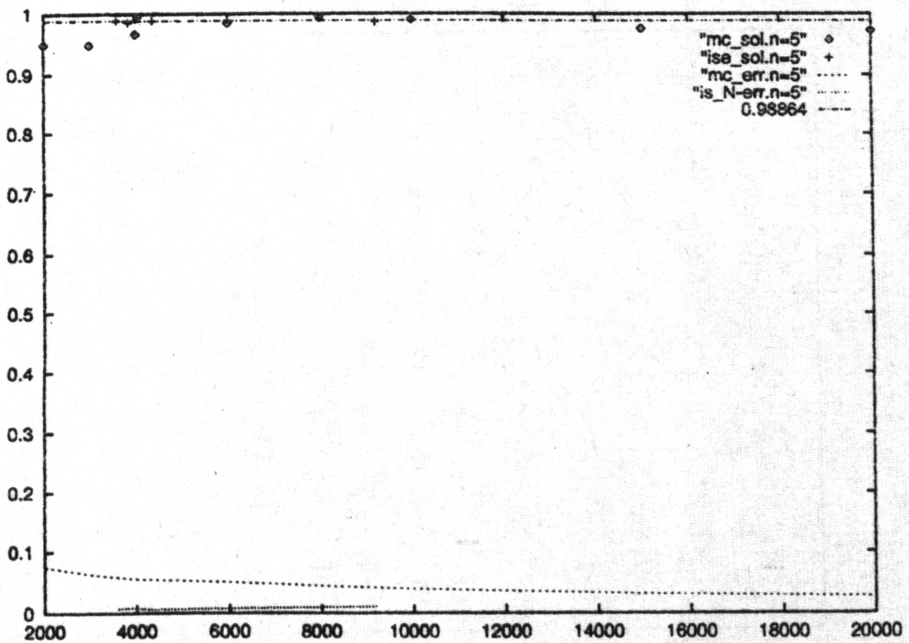
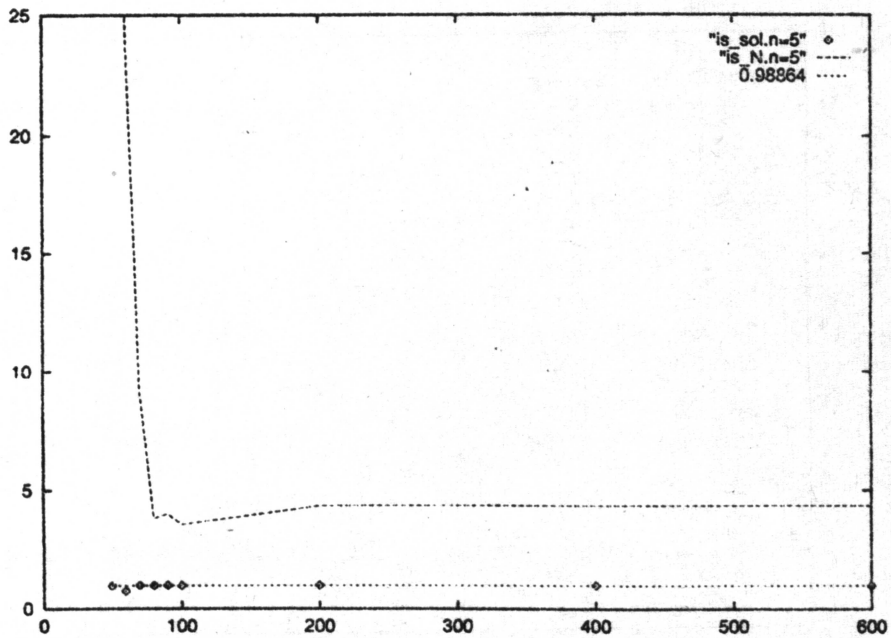


Fig. 2. Fig. 3

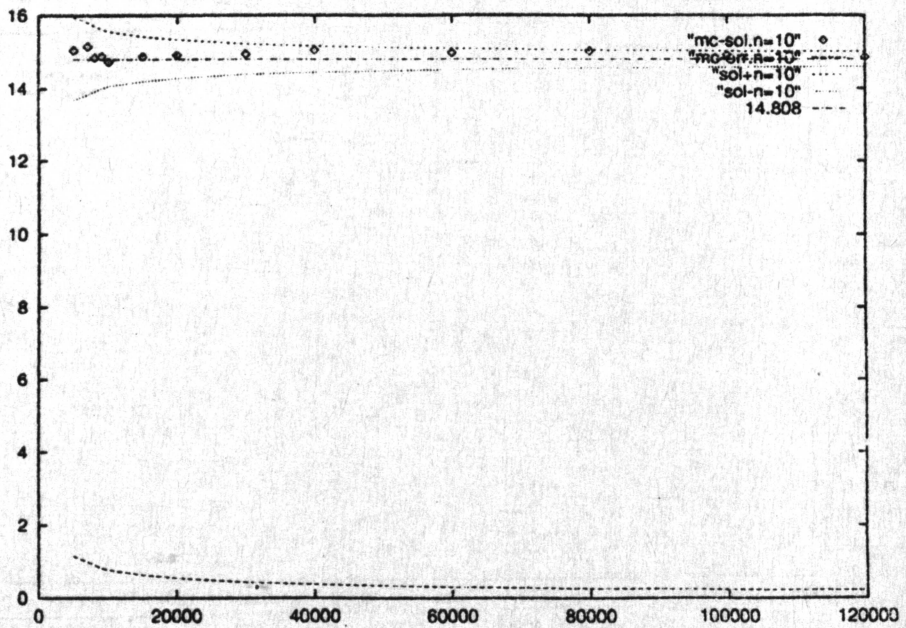
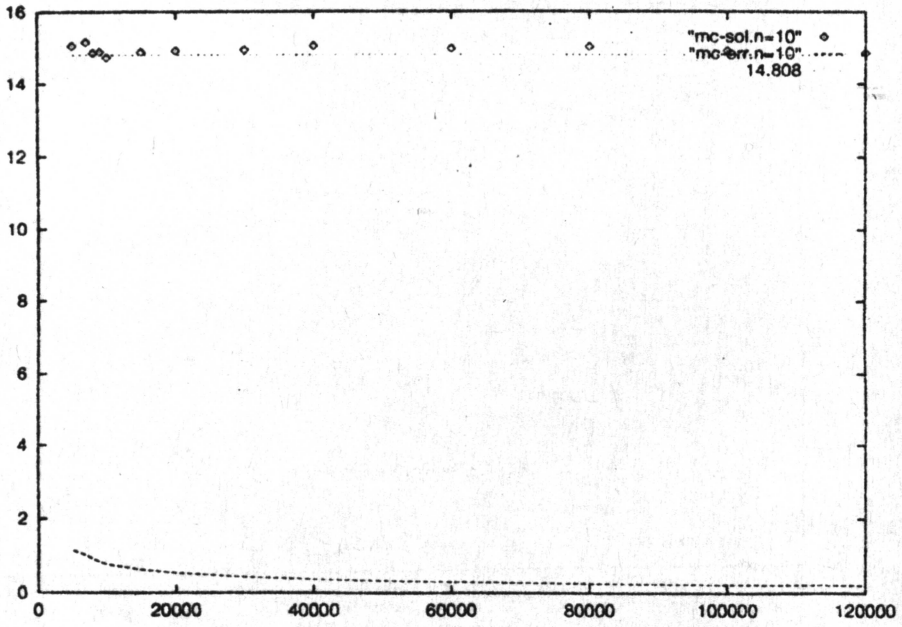


Fig. 4, Fig. 5

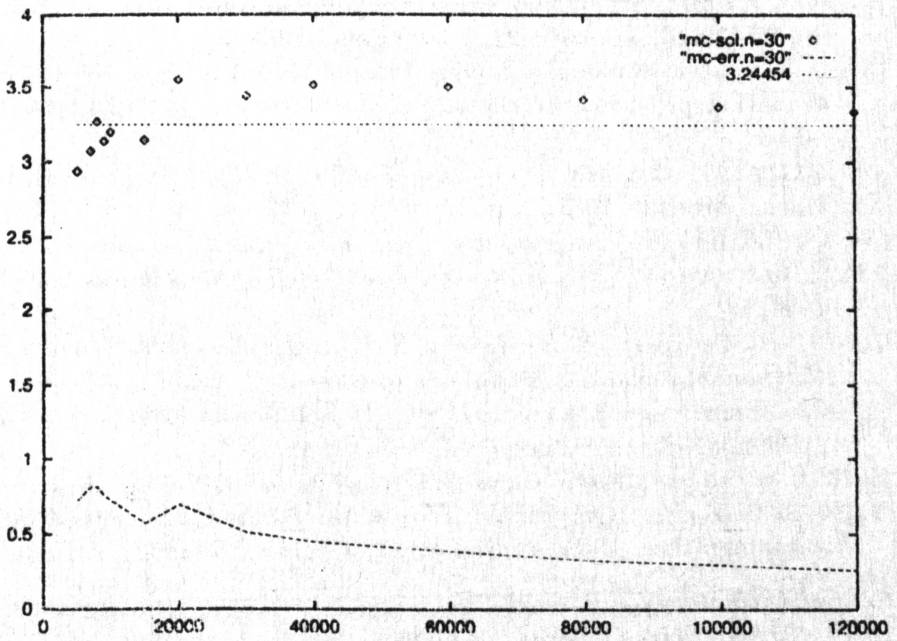
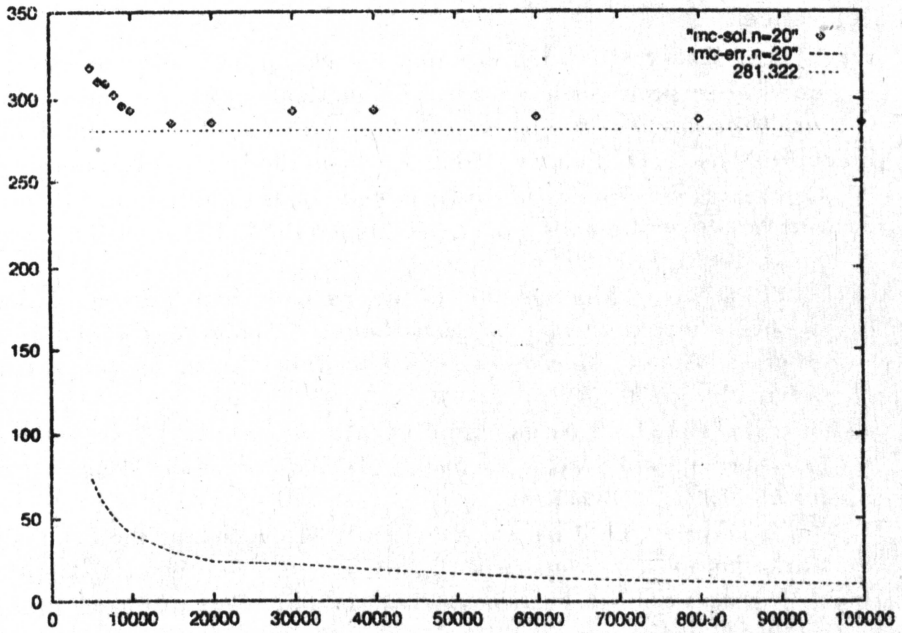


Fig. 6, Fig. 7

References

- [Ba64] N.S. Bahvalov, On the optimal estimations of convergence of the quadrature processes and integration methods, *Numerical Methods for Solving Differential and Integral Equations* (Nauka, Moscow, 1964), 5-63.
- [DT89] I. Dimov, O. Tonev, Monte Carlo methods with overconvergent probable error, in: *Numerical Methods and Applications, Proc. of Intern. Conf on Numerical Methods and Appl., Sofia*, (Publ. House of Bulg. Acad. Sci, Sofia, 1989), 116-120.
- [Di91] I. T. Dimov, Minimization of the probable error for some Monte Carlo methods, in: *Mathematical Modelling and Scientific Computations, Eds. Andreev, Dimov, Markov and Ulrich* (Publ. House of Bulg. Acad. Sci., Sofia, 1991), 159-170.
- [DT93] I. Dimov, O. Tonev, Monte Carlo algorithms: performance analysis for some computer architectures, *J. Computational and Applied Mathematics* **48** (1993), 253-277.
- [DT93a] I. Dimov, O. Tonev, Random walk on distant mesh points Monte Carlo methods, *J. Statistical Physics* **70**, No 5/6 (1993), 1333-1342.
- [Du56] V. Dupach, Stochasticke pocetni metody, *Cas. pro Pest. Mat.* **81**, No 1 (1956), 55-68.
- [Ha66] S. Haber, A modified Monte Carlo quadrature, *Math. of Comput.* **20**, No 95 (1966), 361-368 ; **21**, No 99 (1967), 388-397.
- [Ka50] H. Kahn, Random sampling (Monte Carlo) techniques in neutron attenuation problems, *Nucleonics* **6**, No 5 (1950), 27-33; **6**, No 6 (1950), 60-65.
- [Mi87] G. A. Mikhailov, *Optimization of the "Weight" Monte Carlo Methods*, Nauka, Moscow, 1987.
- [Ni88] S. M. Nikolski, *Quadrature Formulas*, Nauka, Moscow, 1988.
- [SP88] Bl. Sendov, V. Popov, *Averaged Moduli of Smoothness*. Mir, Moscow, 1988.
- [SAK94] Bl. Sendov, A. Andreev, N. Kjurkchiev, *Numerical Solution of Polynomial Equations*, Handbook of Numerical Analysis (General Editors: P.G. Ciarlet and J. L. Lions), Vol. 3), Solution of Equations in R^n (Part 2), North-Holland, Amsterdam, N.Y., 1994.
- [So73] I. M. Sobol, *Monte Carlo Numerical Methods*, Nauka, Moscow, 1973.
- [Ta92] M. D. Takev, On probable error of the Monte Carlo method for numerical integration, *Math. Balkanica (New Series)*, **6** (1992), 231-235.