A Quasi-Monte Carlo Method
for Integration with Improved Convergence⋆

Aneta Karaivanova, Ivan Dimov, and Sofiya Ivanovska

CLPP - Bulgarian Academy of Sciences,
Acad. G. Bonchev St., Bl.25A, 1113 Sofia, Bulgaria
`anet@copern.bas.bg`, `ivdimov@bas.bg`, `sofia@copern.bas.bg`

Abstract. Quasi-Monte Carlo methods are based on the idea that random Monte Carlo techniques can often be improved by replacing the underlying source of random numbers with a more uniformly distributed deterministic sequence. Quasi-Monte Carlo methods often include standard approaches of variance reduction, although such techniques do not necessarily directly translate. In this paper we present a quasi-Monte Carlo method for integration that combines a separation of the domain into *uniformly small* subdomains with the approach of importance sampling. Theoretical estimates for the error bounds and the convergence rate are established. A large number of numerical tests of the proposed method are presented and compared with crude Monte Carlo and weighted uniform sampling. All methods are realized using pseudorandom numbers, and Sobol, Halton and Faure quasirandom sequences. The numerical results confirm the improved convergence of the proposed method when the integrand has bounded derivatives.

1 Introduction

Multidimensional numerical quadratures are of great importance in many practical areas, ranging from atomic physics to finance. The crude Monte Carlo method has rate of convergence $O(N^{-1/2})$ which is independent of the dimension of the integral, and that is why Monte Carlo integration is the only practical method for many high-dimensional problems. Much of the efforts to improve Monte Carlo are in construction of variance reduction methods which speed up the computation.

Quasi-Monte Carlo methods are based on the idea that random Monte Carlo techniques can often be improved by replacing the underlying source of random numbers with a more uniformly distributed deterministic sequence. Quasi-Monte Carlo methods often include standard approaches of variance reduction, although such techniques do not necessarily directly translate. The fundamental feature underlying all quasi-MCMs, however, is the use of a quasi-random sequence. In this paper we study the convergence of a quasi-Monte Carlo method for numerical integration that combines separation of the domain and importance sampling.

2   The Method

Consider the Monte Carlo estimation of the integral:

$$I[f] = \int_D f(x)p(x)\,dx, \tag{1}$$

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

The Monte Carlo integration error is

$$E[\epsilon_N[f]^2]^{1/2} = \sigma[f]N^{-1/2} \tag{2}$$

where

$$\sigma[f] = \left(\int_D (f(x)p(x) - I[f])^2 dx \right)^{1/2}$$

$$\epsilon_N[f] = \int_D f(x)p(x)dx - \frac{1}{N} \sum_{n=1}^{N} f(\xi_n). \tag{3}$$

The error depends on the sequence (factor $N^{-1/2}$) and on the function (factor $\sigma[f]$). All variance reduction methods attack the factor $\sigma[f]$.

The main idea of stratification is as follows. Split the integration region D into N pieces with

$$D = \bigcup_{j=1}^{M} D_j, \ D_i \cap D_j = 0, \ i \neq j; \tag{4}$$

and take N_k random variables in subdomain D_k with

$$\sum_{k=1}^{M} N_k = N. \tag{5}$$

In each subdomain choose points $\xi_n^{(k)}$ distributed with density $p^{(k)}(x)$ such that

$$p^{(k)}(x) = p(x)/\bar{p}_k, \ \ \bar{p}_k = \int_{D_k} p(x)dx. \tag{6}$$

The stratified Monte Carlo formula is:

$$I_N[f] = \sum_{k=1}^{M} \frac{\bar{p}_k}{N_k} \sum_{n=1}^{N_k} f(\xi_n^{(k)}). \tag{7}$$

Stratification always lowers the integration error if the distribution on points is balanced. The resulting error for stratified quadrature is

$$\epsilon_N \approx N^{-1/2}\sigma_s, \ \ \sigma_s^2 = \sum_{k=1}^{M} \sigma^{(k)^2}. \tag{8}$$

Since the variance over a subdomain is always less than the variance over the whole domain, that is $\sigma_s \leq \sigma$, the stratification always lowers the integration error.

In [4] was proposed a combination of stratification with importance sampling. First, consider one-dimensional case. Partition [0,1] into N subintervals:

$$x_0 = 0; \ \ x_N = 1; \ \ D_i \equiv [x_{i-1}, x_i];$$

$$x_i = \frac{C_i}{f(x_{i-1})(N-i+1)}, \ \ i = 1, ..., N-1 \tag{9}$$

where

$$C_i = 1/2[f(x_{i-1}) + f(1)](1 - x_{i-1}), \ i = 1, \ldots, N-1;$$

If $f(x) \in H(1, L)_{[0,1]}$, there exist constants L_i, such that

$$L_i \geq \left| \frac{\partial f}{\partial x} \right| \ \text{ for any } \ x \in D_i. \tag{10}$$

Moreover, for the above scheme there exist constants c_{1_i} and c_{2_i} such that

$$p_i = \int_{D_i} p(x)\, dx \leq c_{1_i}/N, \ i = 1, \ldots, N \tag{11}$$

and

$$\sup_{x_{1_i}, x_{2_i} \in D_i} |x_{1_i} - x_{2_i}| \leq c_{2_i}/N, \ i = 1, \ldots, N. \tag{12}$$

Theorem. Let $f(x) \in H(1, L)_{[0,1]}$. Then for the importance separation (9)-(12) of D

$$\epsilon_N \approx \sqrt{2}[1/N \sum_{j=1}^{N} (L_j c_{1_j} c_{2_j})^2]^{1/2} N^{-3/2}.$$

Now consider the multidimensional case. For the analogous importance separation the following statement is fulfilled ($M = N$):

$$\epsilon_N \approx \sqrt{2}d \left[\frac{1}{N} \sum_{i=1}^{N} (L_i c_{1_i} c_{2_i})^2 \right]^{1/2} N^{-1/2-1/d}.$$

The disadvantage of the above described methods is the increased computational complexity. The accuracy is improved (in fact, importance separation gives the theoretically optimal accuracy, [4]) but the price is increased number of additional computations which makes these methods impractical for large d.

3   Implementation Using Quasirandom Sequences

The use of quasirandom sequences in place of the usual pseudorandom numbers often improves the convergence of the numerical integration. QRNs are constructed to minimize a measure of their deviation from uniformity called discrepancy, which is defined as follows. Consider a set x_i of N points in the d-dimensional unit cube. The discrepancy of this set is

$$D_N = sup_E \left| \frac{\# \text{ of } x_i \in E}{N} - m(E) \right|. \tag{13}$$

Here E is a subrectangle of the unit cube, $m(E)$ is the volume of E, and the sup is taken over all such subrectangles. Most common, the sup is taken only over all subrectangles with one vertex at 0, thus defining the star discrepancy $D^\star$, which is used in the famous Koksma-Hlawka inequality:

Theorem (Koksma-Hlawka). For any sequence x_n and any function with bounded variation, the integration error is bounded as

$$\epsilon_N[f] \leq V[f] D_N^\star, \tag{14}$$

where $V[f]$ is the variation of f in the Hardy-Krausse sense.

A quasirandom, or low-discrepancy, sequence is a sequence which satisfies the condition that

$$D_N \leq C_d \frac{log^d N}{N}, \tag{15}$$

where C_d is a constant for the sequence, independent of N, but which may depend on the dimension d.

There have been many constructions of low discrepancy point sets that have achieved star discrepancies as small as $O(N^{-1}(\log N)^d)$. Most notably there are the constructions of Hammersley, Halton, [3], Soboĺ, [9], Faure, [2], and Niederreiter, [7].

In the presented numerical results, we use multidimensional Halton, Sobol and Faure sequences, [9].

3.1   Importance Separation Using QRNs

Here we consider a modification of the method called *importance separation* and described in section 2. The goal is to have trade-of between the good convergence rate of the method and its computational complexity. We slightly modify the separation of the given domain, which, in the onedimensional case is:

$$x_0 = 0; \;\; x_M = 1; \;\; D_i \equiv [x_{i-1}, x_i];$$

$$x_i = \frac{C_i}{(M - i + 1)}, \;\; i = 1, ..., M - 1 \tag{16}$$

where
$$C_i = 1/2[f(x_{i-1}) + f(1)](1 - x_{i-1}), \ i = 1, \ldots, M - 1.$$

We consider M to be significantly less then N, and M QRNs are generated in each subdomain. The goal is to have in some sense better distribution properties having in mind the behavior of the integrand.

The variance in each subinterval is:

$$V[f] = \int_{x_{i-1}}^{x_i} \left| \frac{df}{dx} dx \right| \le L_i(x_i - x_{i-1}) \le L_i \frac{c'}{M}.$$

Applying Koksma-Hlawka inequality for each subinterval we have

$$\left| \int_{x_{i-1}}^{x_i} f(x)dx - \sum_{j=1}^{N/M} f(x_j) \right| \le L_i \frac{c'}{M} \frac{1}{N/M} = \frac{L_i c'}{N}.$$

Obviously, the error estimation for the whole interval is better than the error estimation of the crude MCM. The multidimensional case is analogous.

4   Numerical Results

A lot of numerical experiments have been done. Here we present the results of solving of two multidimensional integrals, which are used as test examples in [6].

Example 1. The first example is Monte Carlo integration over $I^5 = [0, 1]^5$ of the function

$$f_1(x) = exp \left(\sum_{i=1}^{5} a_i x_i^2 \frac{2 + sin(\sum_{j=1, j \neq i}^{5} x_j)}{2} \right),$$

where $\mathbf{a} = (1, \frac{1}{2}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$.

Example 2. The second example is Monte Carlo integration over $I^7 = [0, 1]^7$ of the function

$$f_2(x) = e^{1 - (\sin^2(\frac{\pi}{2} x_1) + \sin^2(\frac{\pi}{2} x_2) + \sin^2(\frac{\pi}{2} x_3))} \arcsin \left(\sin(1) + \frac{x_1 + \cdots + x_7}{200} \right).$$

In our numerical experiments we compare the results of:

Crude Monte Carlo:
$$I_N[f] = \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{x}_n)$$

Weighted uniform sampling method:
$$I_N[f] = \frac{\sum_{n=1}^{N} f(\mathbf{x}_n)}{\sum_{n=1}^{N} h(\mathbf{x}_n)}$$

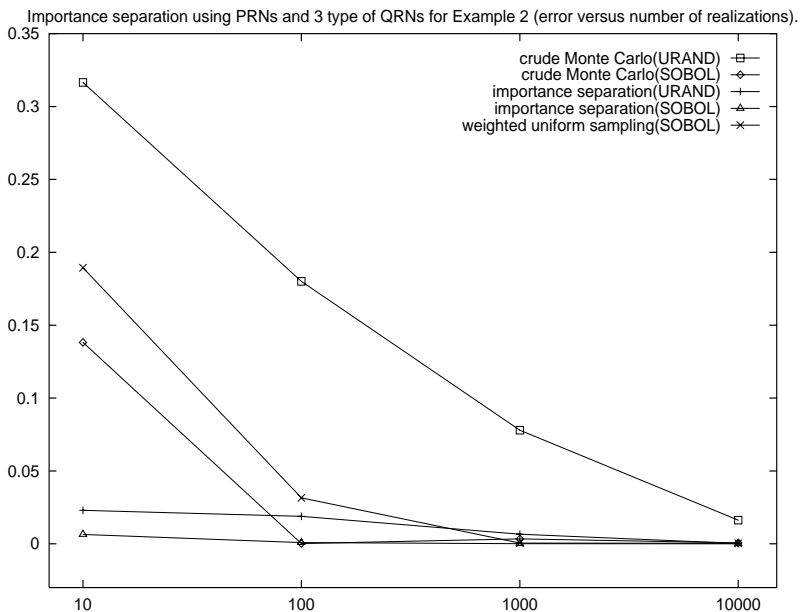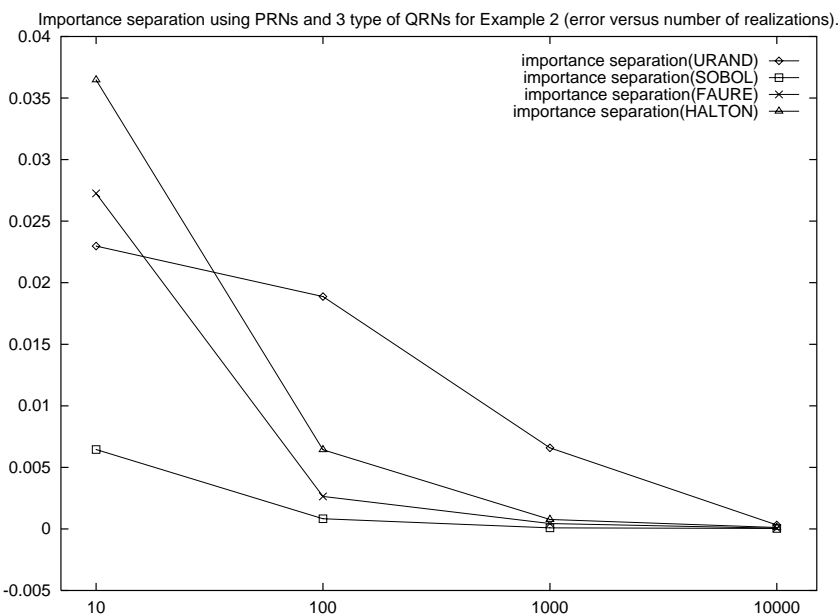Importance separation using PRNs and 3 type of QRNs for Example 2 (error versus number of realizations).

Fig. 1. Crude MC, IS and WUS using PRNs and Sobol QRNs for Example 1 (error versus number of realizations).

Importance separation using PRNs and 3 type of QRNs for Example 2 (error versus number of realizations).

Fig. 2. Importance separation using PRNs and 3 types of QRNs for Example 1 (error versus number of realizations).
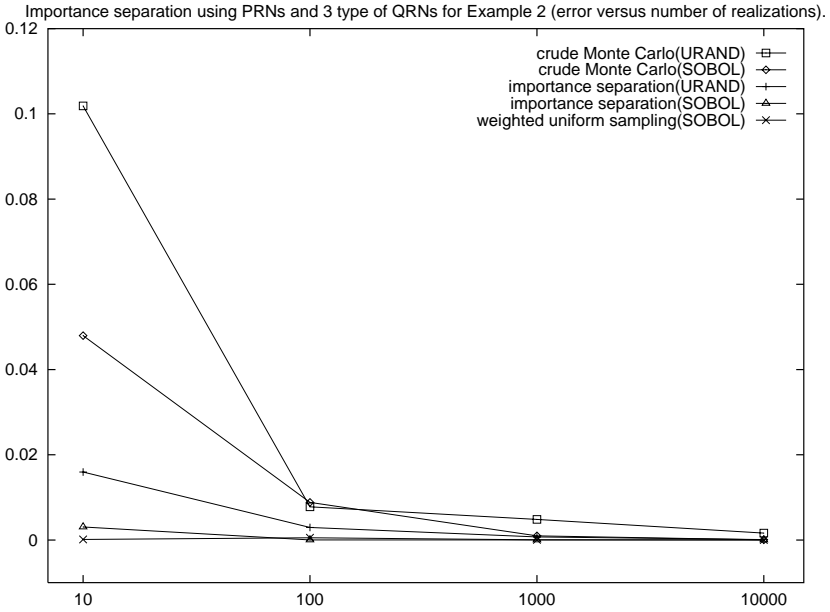
Importance separation using PRNs and 3 type of QRNs for Example 2 (error versus number of realizations).

Fig. 3. Crude MC, IS and WUS using PRNs and Sobol QRNs for Example 2 (error versus number of realizations).

Importance separation using PRNs and 3 type of QRNs for Example 2 (error versus number of realizations).
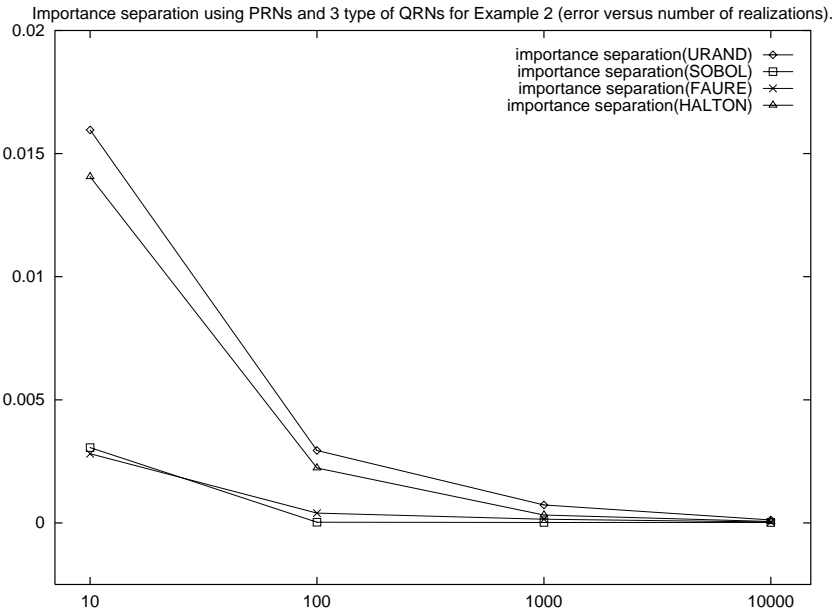

Fig. 4. Importance separation using PRNs and 3 types of QRNs for Example 2 (error versus number of realizations).

Importance separation:

$$I_N[f] = \sum_{k=1}^{M} \frac{\bar{p}_k}{N_k} \sum_{n=1}^{N_k} f(\xi_n^{(k)}).$$

All algorithms are realized with PRNs and QRNs.

The numerical results for the accuracy of the described methods for computing the multidimensional quadratures are presented on Figures 1, 2, 3 and 4. The results are presented as a function of N, number of samples, and as a function of the error, which is computed with respect to the exact solution. Figure 1 and 2 show the results of the crude, weighted uniform sampling and importance separation for both integrals and all methods are performed both with pseudorandom and quasi-random sequences. The importance separation method leads to smaller errors. The most important fact is that using importance separation we have very good accuracy even using small sample. Figure 3 and 4 show the accuracy of importance separation method using different quasirandom sequences. The best results are obtained using the Sobol sequence.

References

1. R. E. Caflisch. Monte Carlo and quasi-Monte Carlo methods, *Acta Numerica*, 7, 1–49, 1998.
2. H. Faure. Discrépance de suites associées à un système de numération (en dimension s), *Acta Arithmetica*, XLI, 337–351, 1992.
3. J. H. Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals, *Numer. Math.*, 2, 84–90, 1960.
4. A. Karaivanova and I. Dimov. Error analysis of an adaptive Monte Carlo method for numerical integration, *Mathematics and Computers in Simulation*, 47, 201–213, 1998.
5. W. Morokoff and R. E. Caflisch. Quasi-Monte Carlo integration, *SIAM J. Comput. Phys.*, 122, 218–231, 1995.
6. B. Moskowitz and R. E. Caflisch. Smoothness and dimension reduction in Quasi-Monte Carlo methods, *J. Math. Comput. Modeling*, 23, 37–54, 1996.
7. H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM: Philadelphia, 1992.
8. A. Owen. Monte Carlo extension of Quasi-Monte Carlo, in D. J. Medieiros, E. F. Watson, M. Manivannan, and J. Carson, (eds.), *Proceedings of WSC'98*, 571–577, 1998.
9. I. M. Soboĺ. The distribution of points in a cube and approximate evaluation of integrals, *Zh. Vychisl. Mat. Mat. Fiz.*, 7, 784–802, 1967. (in Russian).