Faster Evaluation of Multidimensional Integrals

A. Papageorgiou
J.F. Traub

Department of Computer Science
Columbia University
New York, NY 10027
June 1997

Abstract

In a recent paper Keister proposed two quadrature rules as alternatives to Monte Carlo for certain multidimensional integrals and reported his test results. In earlier work we had shown that the quasi-Monte Carlo method with generalized Faure points is very effective for a variety of high dimensional integrals occuring in mathematical finance. In this paper we report test results of this method on Keister's examples of dimension 9 and 25, and also for examples of dimension 60, 80 and 100.

For the 25 dimensional integral we achieved accuracy of 10^{-2} with less than 500 points while the two methods tested by Keister used more than $220,000$ points. In all of our tests, for n sample points we obtained an empirical convergence rate proportional to n^{-1} rather than the $n^{-1/2}$ of Monte Carlo.

1 Introduction

Keister [1] points out that multi-dimensional integrals arise frequently in many branches of physics. He rules out product rules of one-dimensional methods because the number of integrand evaluations required grows exponentially in the number of dimensions. He observes that although Monte Carlo (MC) methods are desirable in high dimension, a large number, n, of integrand evaluations can be required since the expected error decreases as $n^{-1/2}$.

This motivates Keister to seek non-product rules for a certain class of integrands defined below. He proposes two quadrature rules, one by Mc Namee and Stenger (MS) [2], and a second due to Genz and Patterson (GP) [3],[4], which he tests on a specific example of his class of integrands.

In this paper we report test results on Keister's example using quasi-Monte Carlo (QMC) methods. QMC methods evaluate the integrand at deterministic points in contrast to MC methods which evaluate the integrand at random points. The deterministic points belong to *low discrepancy* sequences which, roughly speaking, are uniformly spread as we will see in the next section. Niederreiter [5] is an authoritative monograph on low discrepancy sequences, their properties, and their applications to multi-dimensional integration.

The Koksma-Hlawka inequality (see the next section for a precise statement) states that low discrepancy sequences yield a worst case error for multivariate integration bounded by a multiple of $(\log n)^d/n$, where n is the number of evaluations and d is the dimension of the integrand. A similar bound on the average error is implied by Woźniakowski's theorem [6]. The proof of this theorem is based on concepts and results from information-based complexity [7].

For d fixed and n large, the error $(\log n)^d/n$ beats the MC error $n^{-1/2}$. But for n fixed and d large, the $(\log n)^d/n$ factor looks ominous. Therefore, it was believed that QMC methods should not be used for high-dimensional problems; $d = 12$ was considered high [8, p. 204]. Traub and a then Ph.D. student, Paskov, decided to test the efficacy of QMC methods for the valuation of financial derivatives. Software construction and testing of QMC methods for financial applications was began in Fall 1992. The first tests were run on a very difficult financial derivative in 360 dimensions, which required 10^5 floating point operations per evaluation. Surprisingly, QMC methods consistently beat MC methods.

The first published announcement was in January 1994 [9]. Details appeared in [10], [11], [12]. Tests by other researchers [13], [14] lead to similar conclusions for the high-dimensional problems of mathematical finance.

These results are empirical. A number of hypotheses have been advanced to explain the observed results. One of these is that, due to the discounted value of money, the financial problems are highly non-isotropic with some dimensions far more important than others. Perhaps the QMC methods take advantage of this. A generally accepted explanation is not yet available.

Since Keister's test integral is isotropic it provides an example which is very different than the examples from mathematical finance. To our surprise the QMC method beat both MC and two other methods tested by Keister by very convincing margins.

The problems in [1] require the computation of a weighted multi-dimensional integral

$$\int_{\mathbb{R}^d} f(x)\rho(x)\,dx, \tag{1}$$

where d is the dimension of the problem, $f : \mathbb{R}^d \to \mathbb{R}$ is a *smooth* function, and the weight $\rho(x)$, $x \in \mathbb{R}^d$ satisfies

$$\rho(x) = \prod_{j=1}^{d} \eta(x_j), \tag{2}$$

with $\eta(-x_j) = \eta(x_j)$, $x_j \in \mathbb{R}$. Thus, the weight is symmetric with respect to permutations and changes of sign of the variables. The example in [1] (see also [15]) is

$$\int_{\mathbb{R}^d} \cos(\|x\|)e^{-\|x\|^2}\,dx, \tag{3}$$

where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^d$.

The integral in (3) can be reduced, via a change of variable, to a one-dimensional integral which can be analytically integrated. As we will see, the QMC method takes advantage of the dependence on the norm *automatically* and provides a numerical solution with error similar to a one-dimensional integral.

The QMC method that we test in this paper uses points from the generalized Faure sequence, which was constructed by Tezuka [16]. We will refer to it as QMC-GF. This sequence has been very successful in solving problems of mathematical finance [12], [14].

The performance of QMC-GF on the integral (3) is most impressive. For example, for the 25-dimensional integral it achieves error 10^{-2} using less than 500 points, far superior to all the other methods. Its error over the range we tested, which was up to 10^6 points, was $c \cdot n^{-1}$, with $c < 110$, $d = 9, 25, 80, 60, 100$. That may be compared with the MC method whose error was proportional to $n^{-1/2}$.

We summarize the remainder of this paper. In the next section we provide a brief introduction to low discrepancy sequences. Test results are given in the third section. A summary of our results and future research concludes the paper.

2 Low Discrepancy Sequences

Discrepancy is a measure of deviation from uniformity of a sequence of real numbers. In particular, the discrepancy of n points $x_1, \ldots, x_n \in [0,1]^d$, $d \geq 1$, is defined by

$$D_n^{(d)} = \sup_E \left| \frac{A(E;n)}{n} - \lambda(E) \right|, \tag{4}$$

where the supremum is taken over all the subsets of $[0,1]^d$ of the form $E = [0,t_1) \times \cdots \times [0,t_d)$, $0 \leq t_j \leq 1$, $1 \leq j \leq d$, λ denotes the Lebesgue measure, and $A(E;n)$ denotes the number of the x_j that are contained in E. A detailed analysis of low discrepancy sequences can be found in [5] and in the references therein.

A sequence $x_1, x_2, \ldots$ of points in $[0,1]^d$ is a low discrepancy sequence iff

$$D_n^{(d)} \leq c(d)\frac{(\log n)^d}{n}, \ \forall n > 1, \tag{5}$$

where the constant $c(d)$ depends only on the dimension d. Neiderreiter, see [5], gives a general method for constructing (t,d)-sequences, $t \geq 0$, which are low discrepancy sequences. The discrepancy of the first n points in a (t,d)-sequence is given by

$$D_n^{(d)} \leq c(t,d,b)\frac{(\log n)^d}{n} + O\left(\frac{(\log n)^{d-1}}{n}\right),$$

where $b \geq 2$ is an integer parameter, upon which the sequence depends, and $c(t,d,b) \approx b^t/d! \cdot (b/2\log b)^d$. Hence, the value $t = 0$ is desirable.

The generalized Faure sequence [16] is a $(0,d)$ sequence and is obtained as follows. For a prime number $b \geq d$ and $n = 0, 1, \ldots$, consider the base b representation of n, i.e.,

$$n = \sum_{i=0}^{\infty} a_i(n) b^i,$$

where $a_i(n) \in [0,b)$ are integers, $i = 0, 1, \ldots$. The j-th coordinate of the point x_n is then given by

$$x_n^{(j)} = \sum_{k=0}^{\infty} x_{nk}^{(j)} b^{-k-1}, \ 1 \leq j \leq d,$$

where

$$x_{nk}^{(j)} = \sum_{s=0}^{\infty} c_{ks}^{(j)} a_s(n).$$

The matrix $C^{(j)} = (c_{ks}^{(j)})$ is called the generator matrix of the sequence and is given by $C^{(j)} = A^{(j)} P^{j-1}$, where $A^{(j)}$ is a nonsingular lower triangular matrix and P^{j-1} denotes the $j-1$ power of the Pascal matrix, $1 \leq j \leq d$.

We conclude this section by stating the Koksma-Hlawka inequality which establishes the relationship between low discrepancy sequences and multivariate integration, see [5]. If f is a real function, defined on $[0,1]^d$, of bounded variation, $V(f)$, in the sense of Hardy and Krause, then for any sequence $x_1, \ldots, x_n \in [0,1)^d$ we have

$$\left| \int_{[0,1]^d} f(x)\, dx - \frac{1}{n}\sum_{i=1}^{n} f(x_i) \right| \leq V(f) D_n^{(d)}.$$

3 Methods and Test Results

We transform the integral (3) to one over the cube $[0,1]^d$. We have

$$I_d(\cos) = \int_{\mathbb{R}^d} \cos(\|x\|) e^{-\|x\|^2}\, dx \ = \ 2^{-d/2} \int_{\mathbb{R}^d} \cos(\|y\|/\sqrt{2}) e^{-\|y\|^2/2}\, dy \tag{6}$$

$$= \pi^{d/2} \int_{\mathbb{R}^d} \cos(\|y\|/\sqrt{2}) \frac{e^{-\|y\|^2/2}}{(2\pi)^{d/2}}\, dy$$

$$= \pi^{d/2} \int_{[0,1]^d} \cos\left(\sqrt{\sum_{j=1}^{d} (\phi^{-1})^2(t_j)/2}\right) dt,$$

where ϕ is the cummulative normal distribution function with mean 0 and variance 1,

$$\phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-s^2/2}\, ds,\ u \in [-\infty, \infty].$$

We obtain the n deterministic sample points $x_i = (x_{i1,\ldots,x_i d}) \in \mathbb{R}^d$, $i = 1, \ldots, n$, by setting $x_{ij} = \phi^{-1}(t_{ij})$, where $t_i = (t_{i1}, \ldots, t_{id}) \in [0,1]^d$, $i = 1, \ldots, n$, are n consecutive terms of a low discrepancy sequence. Our method, $I_{d,n}$, is defined by:

$$I_{d,n}(\cos) = \frac{\pi^{d/2}}{n} \sum_{i=1}^{n} \cos\left(\sqrt{\sum_{j=1}^{d} (\phi^{-1})^2(t_{i,j})/2}\right). \tag{7}$$

Our test problem could be reduced to a one-dimensional integral. We did not do this because we wanted test how QMC methods perform on d-dimensional integrals. As we will see, the empirical rate of convergence of QMC-GF is n^{-1} which suggests that this method takes advantage of the dependence on the norm automatically without a dimension reducing transformation. A method corresponding to (7) can be derived for the more general integration problem (1) with weight function satisfying (2).

We report test results. We used the generalized Faure[1] low discrepancy sequence [16] to derive the sample points for the QMC method. We remind the reader that we call this the QMC-GF method. We compared this method to the McNamee-Stenger (MS) and Genz-Patterson (GP) [2], [3], [4] methods. We also tested using a Monte Carlo method of the form (7), i.e., using randomly generated points $t_{i,j}$. Hence, we use the same change of variable for QMC-GF and MC.

The value, $I_d(\cos)$, of the integral (6) is, see [1], $I_9(\cos) = -71.633234291$ and $I_{25}(\cos) = -1.356914 \cdot 10^6$. We used Mathematica to compute $I_{60}(\cos) = 4.89052986 \cdot 10^{14}$, $I_{80}(\cos) = 6.78878724 \cdot 10^{19}$ and $I_{100}(\cos) = 4.57024396 \cdot 10^{24}$. We measure the accuracy of an approximation by computing its relative error (fractional deviation). We observe the least number of sample points required by an method to achieve and *maintain* a relative error below a specified level, e.g. 10^{-3}, until the end of the simulation. We introduced this more conservative way of assessing the performance of a method in [12]. Thus, we study the error of an method throughout a simulation. We believe that this has advantages over performance reports that are based only on values at the end of a simulation.

We summarize our findings and then provide some details.

- The QMC-GF method outperforms the MS and GP methods for $d = 25$.

[1]The generalized Faure and the Sobol' low discrepancy sequences are included in FINDER, a Columbia University software system, and are available to researchers upon request by writing the authors.

- The MS and GP methods are sensitive to the dimension. They perform quite well for $d = 9$ and very poorly for $d = 25$. For example, for $d = 25$ and for accuracy of the order 10^{-2} these methods use some $220,000$ points while the QMC-GF method uses less than 500 points. Therefore, they should only be used when the dimension is relatively low.

- The QMC-GF method performs well for $d = 9$, 25, 60, 80 and 100.

- The relative error of the QMC-GF method is bounded by

$$c_d \cdot n^{-1},\ c_d < 110,\ n \leq 10^6,\ d = 9, 25, 60, 80, 100. \tag{8}$$

  Note that this is an empirical conclusion. We write c_d to suggest that, in principle, this constant depends on d although we did not see a strong dependence in our tests.

- The QMC-GF method achieves relative error 10^{-2} using about 500 points.

- The relative error of the MC method is bounded by $\beta \cdot n^{-1/2}$ as predicted by the theory.

First we consider the case $d = 9$. The performance of the QMC-GF, and the GP methods is comparable for accuracy less than 10^{-4}. The relative error of the MS method fluctuates about the value 10^{-4} for sample sizes between $36,967$ and $96,745$ points, see [1, Table I], and is slower than the QMC-GF method since it requires at least four times as many function evaluations. (For this level of accuracy the MC method requires more than 10^6 points).

For $d = 25$ the results are striking. The MS and GP methods require about $220,000$ points for accuracy of order 10^{-2} while the QMC-GF method requires less than 500 points. Table I is from [1, Table II] and exhibits the performance of the MS and GP methods.

Method	**Number of Points**	**Relative Error**
GP and MS	$1,251$	2.00
GP	$19,751$	0.40
MS	$20,901$	0.75
GP	$227,001$	0.06
MS	$244,101$	0.07

Table I. Comparison of MS and GP methods, d=25

Table II summarizes the performance of the QMC-GF method.

Method	**Number of Points**	**Relative Error**
QMC-GF	500	10^{-2}
QMC-GF	$1,200$	10^{-3}
QMC-GF	$14,500$	$5 \cdot 10^{-4}$
QMC-GF	$214,000$	$5 \cdot 10^{-5}$

Table II. The Quasi-Monte Carlo method, d=25

As we mentioned above, we are using a very conservative criterion when we report relative error. It takes about 219,000, 490,000, and many more than 10^6 points for the MC method to reach accuracies of 10^{-3}, $5 \cdot 10^{-4}$, and $5 \cdot 10^{-5}$, respectively.

Figure 1 exhibits the relative error of the QMC-GF method for $d = 25$. The horizontal axis shows the sample size n, while the vertical axis shows the relative error. The horizontal lines depict the accuracy.

Figure 2 shows the convergence rate of the QMC-GF method. We plot the logarithm of the relative error as a function of the logarithm of the sample size for $d = 25$ and obtain the linear convergence summarized in (8).

Recently, Keister [17] obtained good results using a public domain version of the Sobol' low discrepancy sequence.

Keister [1] did not perform tests for $d > 25$. We tested the QMC-GF method for $d = 60$, 80 and 100 and we found that its performance is comparable to that of the lower values of d. We did not find evidence suggesting that its performance suffers as the dimension grows. This is shown in the empirical error equation (8) and is further demonstrated in Figure 3, which shows the convergence of QMC-GF for $d = 100$. In particular, in Figure 3 we plot the logarithm of the relative error as a function of the logarithm of the sample size.

4 Summary and Future Research

We have shown that the QMC-GF method beats MC methods and the MS and GP methods by a wide margin for Keister's 25-dimensional example. We have also shown that its good performance is maintained when the dimension takes much higher values. Other high dimensional problems motivated by applications to physics should be tested.

Extensive testing on a variety of high-dimensional integrals which occur in mathematical finance also find QMC methods consistently beating the MC method. Preliminary results from our tests on high-dimensional integrals arising from several very different applications again point to the superiority of QMC over MC.

The results are empirical. There is currently no theory which explains why, for a variety of applications, QMC methods are much better than one would expect from the Koksma-Hlawka inequality or from Woźniakowski's theorem. Finding the theoretical justification for the superiority of QMC methods for certain classes of integrands is a most important direction of future research.

Acknowledgments

We thank Bradley Keister for his comments on a draft of this paper. We are grateful to Richard Palmer for directing us to Bradley Keister's paper, and to Henryk Woźniakowski for his comments on the manuscript.

References

[1] Keister, B.D., Multidimensional Quadrature Algorithms, *Computers in Physics*, 10:20, 119–122, 1996.

[2] Mc Namee, J., and Stenger, F., Construction of Fully Symmetric Numerical Integration Formulas, *Numer. Math.*, 10, 327–344, 1967.

[3] Genz, A., A Lagrange Extrapolation Algorithm for Sequences of Approximations to Multiple Integrals, *SIAM J. Sci. Stat. Comput.*, 3, 160–172, 1982.

[4] Patterson, T.N.L., The Optimum Addition of Points to Quadrature Formulae, *Mathematics of Computation*, 22, 847–856, 1968.

[5] Niederreiter, H., Random Number Generation and Quasi-Monte Carlo Methods, *CBMS-NSF Regional Conference Series in Applied Math. No. 63, SIAM*, 1992.

[6] Woźniakowski, H., Average case complexity of multivariate integration, *Bulletin of the American Mathematical Society*, 24, 185–194, 1991.

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

[8] Bratley, P., Fox, B.L., and Niederreiter, H., Implementation and Tests of Low-Discrepancy Sequences, *ACM Trans. on Modeling and Computer Simulation*, 2:3, 195–213, 1992.

[9] Traub, J.F. and Woźniakowski, H., Breaking Intractability, *Scientific American*, 270, 102–107, 1994.

[10] Paskov, S.H. and Traub, J.F., Faster Valuation of Financial Derivatives, *The Journal of Portfolio Management*, 113–120, Fall 1995.

[11] Paskov, S.H., New Methodologies for Valuing Derivatives, in Mathematics of Derivative Securities, S. Pliska and M. Dempster eds., Isaac Newton Institute, *Cambridge University Press*, Cambridge, UK, 1997.

[12] Papageorgiou, A., and Traub, J.F., Beating Monte Carlo, *Risk*, 9:6, 63–65, 1996.

[13] Joy, C., Boyle, P.P., and Tan, K.S., Quasi-Monte Carlo Methods in Numerical Finance, *working paper, University of Waterloo, Waterloo*, Ontario, Canada N2L 3G1, 1995.

[14] Ninomiya, S., and Tezuka, S., Toward real-time pricing of complex financial derivatives, *Applied Mathematical Finance*, 3, 1–20, 1996.

[15] Capstick, S., and Keister, B.D., Multidimensional quadrature algorithms at higher degree and/or dimension, *Journal of Computational Physics*, 123, 267–273, 1996.

[16] Tezuka, S., Uniform Random Numbers: Theory and Practice, *Kluwer Academic Publishers*, Boston, 1995.

[17] Keister, B.D., Private communication, 1997.

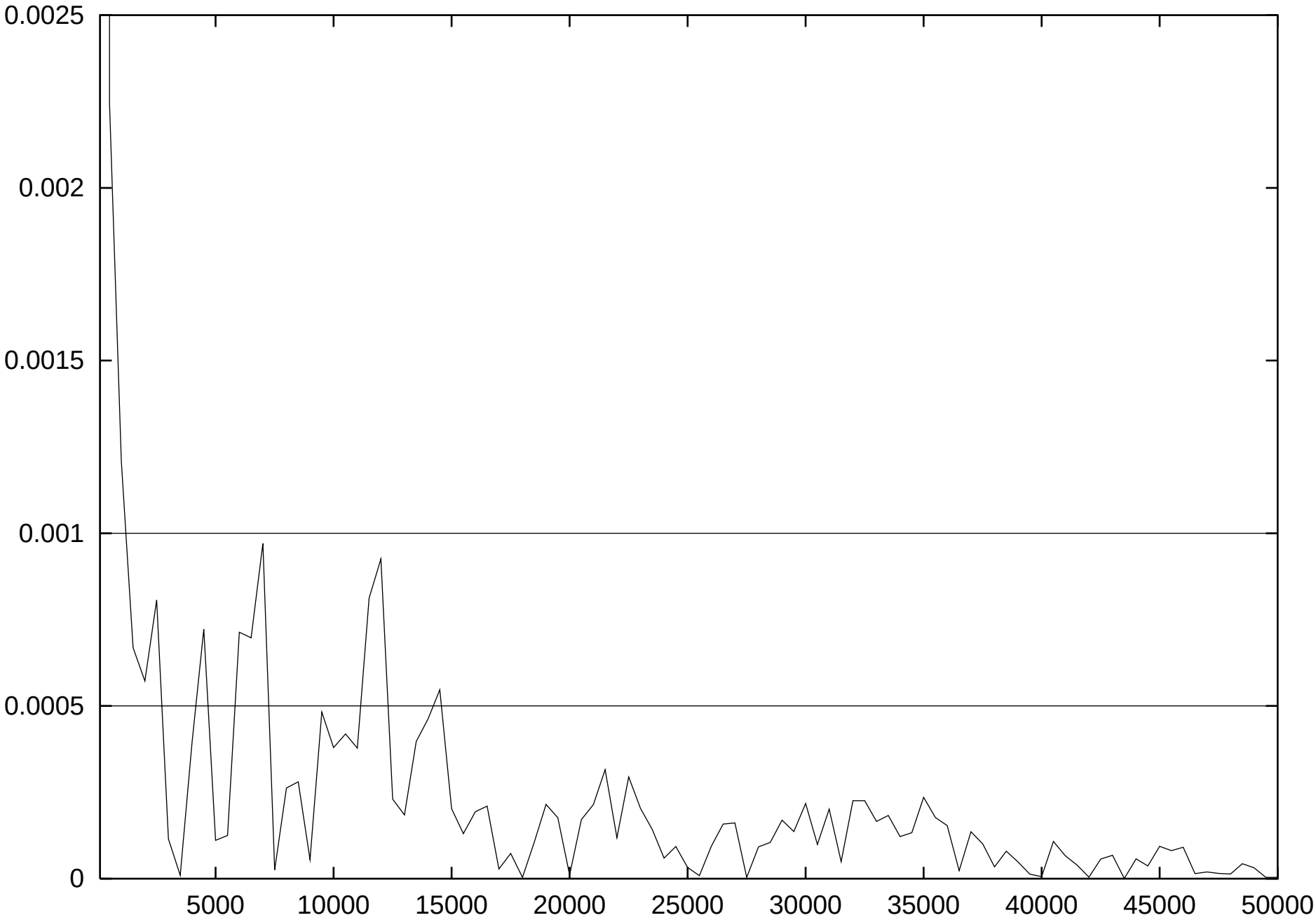

Figure 1. QMC-GF, relative error as a function of the sample size, d=25

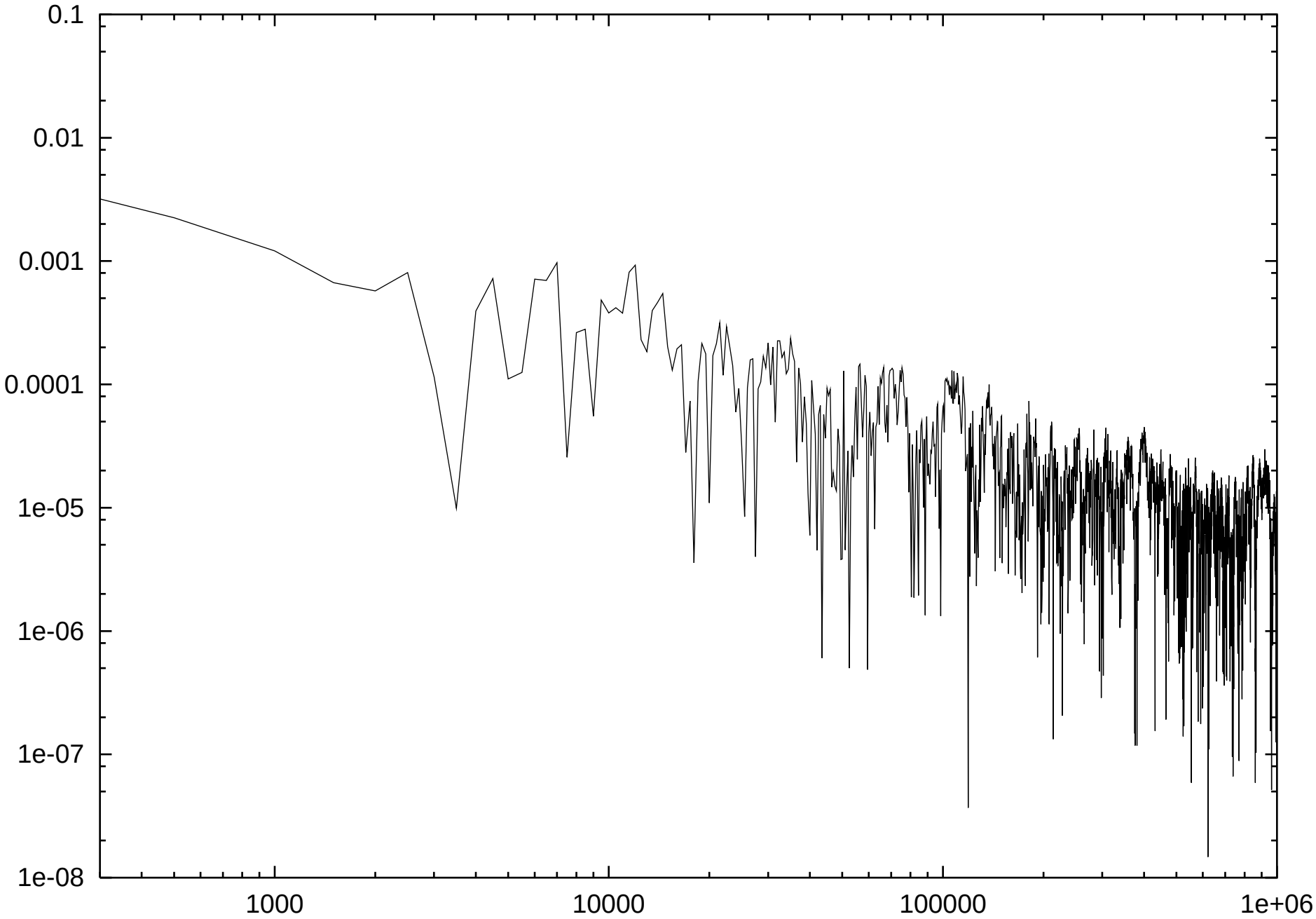

Figure 2. QMC-GF, log(relative error) as a function of log(sample size), d=25

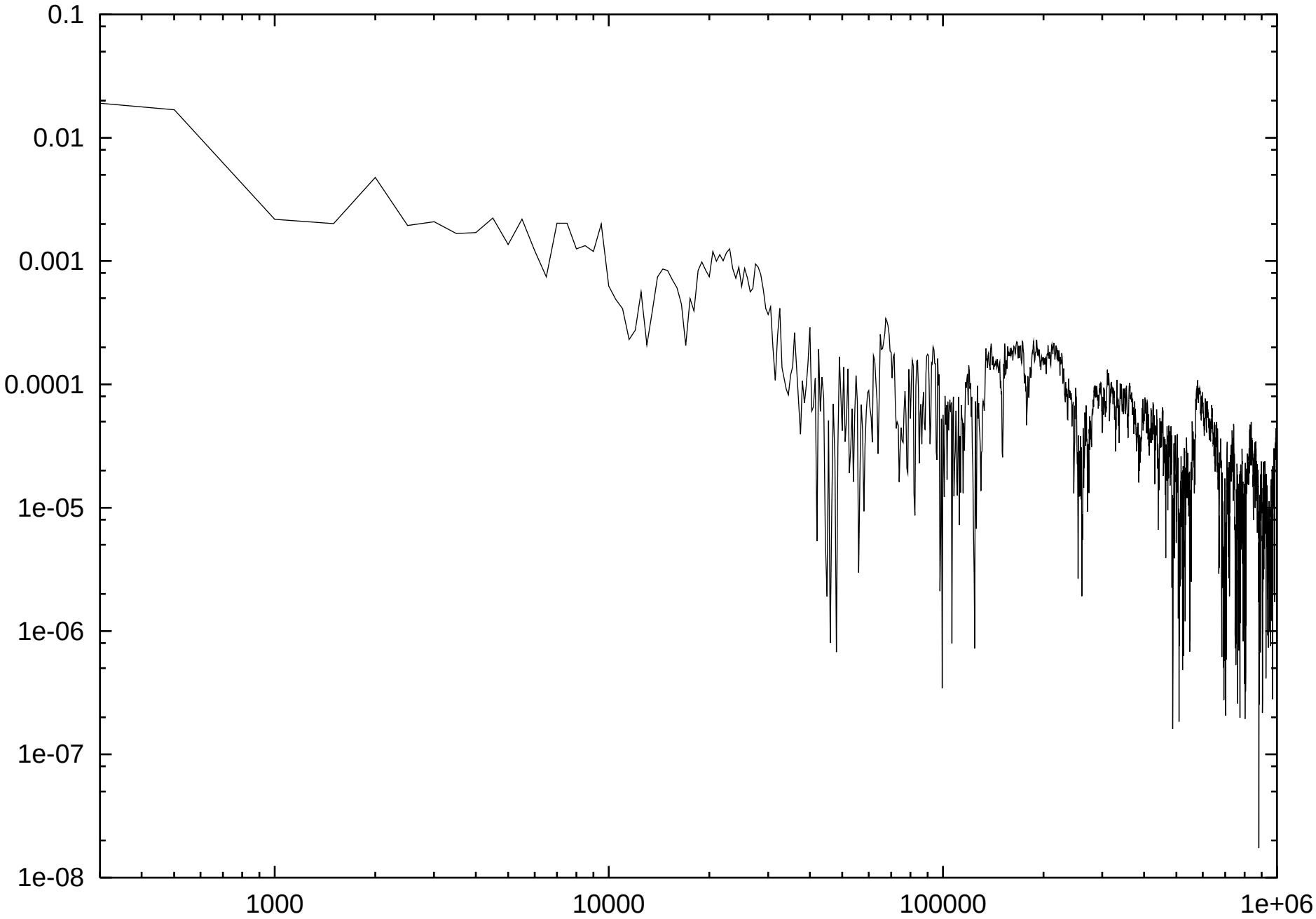

Figure 3. QMC-GF, log(relative error) as a function of log(sample size), d=100