Contributions to the Theory of Monte Carlo and Quasi-Monte Carlo Methods

by

Giray Ökten

ISBN: 1-58112-041-9

DISSERTATION.COM

1999

ISBN: 1-58112-041-9

Dissertation.com
1999

www.dissertation.com/library/1120419a.htm

CONTRIBUTIONS TO THE THEORY OF MONTE CARLO AND QUASI-MONTE CARLO METHODS

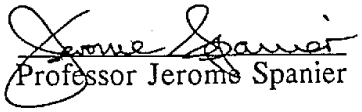
BY

GİRAY ÖKTEN

A Dissertation submitted to the Faculty of Claremont Graduate University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate Faculty of Mathematics

Claremont, California
1997

Approved by:



Professor Jerome Spanier

We, the undersigned, certify that we have read this dissertation and approve it as adequate in scope and quality for the degree of Doctor of Philosophy.

Dissertation Committee:

Professor Jerome Spanier, Chair

Professor John Angus, Member

Professor Henry Krieger, Member

Professor Francis Su, Visiting Examiner

DEDICATED TO MY MOTHER, HATİCE FİKRET SİPAHİ-ÖKTEN, AND MY FATHER, MY LIFE TIME TUTOR, MUZAFFER ÖKTEN

Abstract of the Dissertation

CONTRIBUTIONS TO THE THEORY OF MONTE CARLO AND QUASI-MONTE CARLO METHODS

by

Giray Ökten

The Claremont Graduate School: 1997

Quasi-Monte Carlo methods, which are often described as deterministic versions of Monte Carlo methods, were introduced in the 1950s by number theoreticians. They improve several deficiencies of Monte Carlo methods; such as providing estimates with deterministic bounds and avoiding the paradoxical difficulty of generating random numbers in a computer. However, they have their own drawbacks. First, although they provide faster convergence than Monte Carlo methods asymptotically, the advantage may not be practical to obtain in "high" dimensional problems. Second, there is not a practical way to measure the error of a quasi-Monte Carlo simulation. Finally, unlike Monte Carlo methods, there is a scarcity of error reduction techniques for these methods.

In this dissertation, we attempt to provide remedies for the disadvantages of quasi-Monte Carlo methods mentioned above. In the first part of the dissertation, a hybrid-Monte Carlo sequence designed to obtain error reduction in high dimensions is studied. Probabilistic results on the discrepancy of this sequence as well as results obtained by applying the sequence to problems

from numerical integration and mathematical finance are presented. In the second part of the dissertation, a new hybrid-Monte Carlo method is introduced, in an attempt to obtain a practical statistical error analysis using low-discrepancy sequences. It is applied to problems from mathematical finance and particle transport theory to compare its effectiveness with the conventional methods. In the last part of the dissertation, a generalized quasi-Monte Carlo integration rule is introduced. A Koksma-Hlawka type inequality for the rule is proved, using a new concept for the variation of a function. As a consequence of the rule, error reduction techniques and in particular an "importance sampling" type statement are derived.

Problems from different disciplines are used as practical tests for our methods. The numerical results obtained in favor of the methods suggest the practical advantages that can be realized by their use in a wide variety of applications.

ACKNOWLEDGMENTS

I would like to thank the Mathematics Faculty of Claremont Graduate University for the support they have provided in this long journey. I thank Professor Jerome Spanier, Professor John Angus, Professor Henry Krieger and Professor Francis Su for serving as my committee members. A particular acknowledgment is due Professor Jerome Spanier and Professor John Angus, whose suggestions improved the quality of the dissertation.

My gratitude to my thesis advisor, Professor Jerome Spanier, who introduced me to Monte Carlo and quasi-Monte Carlo methods. He has been helpful in both mathematical and non-mathematical issues. I am indebted for his constant support, encouragement and friendship.

I have been most fortunate for having a loving and supporting family and I thank every member of them. In particular, I acknowledge my beloved wife, Aşkım Cebeci-Ökten, for her love, support and the joy she has brought to my life.

Contents

Chapter 1

Introduction

1.1 Monte Carlo Methods

Monte Carlo methods were mainly developed in the 1940s, by mathematicians and scientists, including Von Neumann, Fermi, Ulam and Metropolis, working on the development of nuclear weapons in Los Alamos during World War II. The term "Monte Carlo" was coined for these methods as a code word, by Von Neumann and Ulam, suggesting the "probabilistic" nature of these methods. Indeed, these methods were first applied to problems with apparent probabilistic structure. The first documented application of Monte Carlo methods appears in a paper by G. Comte de Buffon in 1777. He described an experiment in which a needle of length L is thrown randomly onto a floor ruled with straight lines a distance d apart. He then suggested estimating the probability P that the needle will intersect a line, by throwing the needle many times and calculating the ratio of the number of throws hitting a line to the total number of throws. He also carried out the mathematical analysis and showed that

$$P = \frac{2L}{\pi d}.$$

Later, Laplace suggested that this idea could be used to evaluate π from the throws of the needle.

Monte Carlo methods can be used not only in "naturally" stochastic problems, but also in deterministic problems. In general, any quantity that can be written as the expected value of a random variable defined on a probability space, can be estimated by these methods. The theoretical basis is the strong law of large numbers, which states that the sequence of sample means of a random variable converges to its expected value almost surely, provided the expected value is finite. A simple application is to numerical integration, where one uses the averages

$$\frac{1}{N}\sum_{n=1}^{N} f(\alpha_n)$$

to estimate the integral

$$\int_{[0,1]^s} f(x)dx.$$

Here, $\{\alpha_n\}_{n=1}^{\infty}$ is a "random" sequence of vectors in $[0,1)^s$, which is needed to make sure that the samples $f(\alpha_1), ..., f(\alpha_N)$ are independent and identically distributed realizations of the random variable f, defined on the space $[0,1]^s$ with the Lebesgue measure as the probability measure. The concept of "randomness", however, is a matter of debate among mathematicians, as well as philosophers. When we think of a random process, we often think of events whose outcomes cannot be determined, such as physical outcomes of the "scattering" of an electron by an atom, or simply the outcome of flipping a fair coin. However, the fact that we cannot determine the outcome of flipping a fair coin is not a property of the physical process itself, but merely an indication of our understanding of the process, making it difficult to define "uncertainty", or "randomness" in an objective way. In spite of this difficulty (or maybe, because of it), several mathematicians and philosophers, including Kolmogorov, Martin-Löf, Chaitin and Popper, have suggested various definitions of

"random" numbers.

Whatever our understanding of a random event or a random sequence of numbers might be, the challenge is to find a practical way to generate random numbers, which itself causes other problems. Today, we use mainly arithmetical operations to generate numbers that we call "pseudorandom" numbers, which are hoped to imitate "random" numbers. The success of the imitation is determined by the use of statistical tests, often called "tests of randomness".

> "Any one who considers arithmetical methods of producing random digits is, of course, in a state of sin. For, as has been pointed out several times, there is no such thing as a random number-there are only methods to produce random numbers, and a strict arithmetical procedure is of course not such a method. (It is true that a problem we suspect of being solvable by random methods may be solvable by some rigorously defined sequence, but this is a deeper mathematical question than we can go into now.)" [von Neumann]

Perhaps what von Neumann stated as a deeper mathematical question was solved, at least partly, by the theory of uniform distribution, which claims that the reason our "random" methods work does not have anything to do with their "randomness", but "uniformity". This brings us to the notion of "quasi-Monte Carlo" methods or "number-theoretic" methods.

1.2 Quasi-Monte Carlo Methods

These methods are often described as deterministic versions of Monte Carlo methods, although in practice, Monte Carlo methods are also strictly speaking, deterministic. Quasi-Monte Carlo methods emphasize "uniformity" of a number sequence, instead of "randomness" of it.

> "It is suggested that, instead of clinging to vague concepts of randomness, it might be better to aim at working with sequences making no pretence of random origin, but so devised as to give the best possible guarantee of accuracy in computations. Such methods of computation could be described as quasi-Monte Carlo." [Zaremba]

The theoretical basis of these methods is the Koksma-Hlawka inequality, which gives a "deterministic" upper bound for the absolute value of the error in approximating an integral, $\int_{[0,1]^s} f(x)dx$, by the average value of the integrand, $\frac{1}{N}\sum_{n=1}^{N} f(q_n)$, evaluated on a sequence of "deterministic" vectors $\{q_n\}_{n=1}^{\infty}$. The error is bounded by a product of two terms; the variation of the integrand, which measures the "smoothness" of it, and the discrepancy of the underlying number sequence $\{q_n\}$, which measures the deviation of that sequence from uniform distribution.

An advantage of quasi-Monte Carlo methods is that they avoid the difficulties associated with "randomness". Moreover, they provide deterministic error bounds, as given by the Koksma-Hlawka inequality, with a deterministic order $O(N^{-1}(\log N)^s)$ (as $N \to \infty$). For Monte Carlo methods, however, the error bounds are probabilistic with order $O\left(N^{-1/2}\right)$. In Chapter 3, we will investigate the advantages and disadvantages of these methods further. It will then be clear that both methods (Monte Carlo and quasi-Monte Carlo) have some desirable properties, which leads us to the study of hybrid-Monte Carlo methods.

1.3 Hybrid-Monte Carlo Methods

The objective of these methods is to combine the best features of Monte Carlo and quasi-Monte Carlo methods. Mainly, they have been used to obtain both improved error bounds and practical error analysis in the domain of low-discrepancy sequences. Chapters 2 and 3 of this thesis are devoted to these methods.

The "success" of these methods comes at the expense of borrowing the not-so-clear concepts of Monte Carlo methods. We will, for example, try to establish a statistical analysis of the results of quasi-Monte Carlo estimates in Chapter 3, which is a desirable feature of Monte Carlo estimates. However, performing statistical calculations on Monte Carlo estimates is itself open to criticism.

> "Such [pseudorandom] sequences are perfectly determined, and the results of computations carried out with their help are equally determined. Consequently, it does not make sense to talk, for instance, of the variance of the results of such computations." [Zaremba]

We will not try to justify Monte Carlo calculations in this thesis. Instead, we will follow the literature in treating pseudorandom numbers as "random" numbers and using probability theory to analyze them. The reason that this can be done, borrowing von Neumann's words, is a deep mathematical question left unanswered.

In Chapter 2, we will study theoretical properties of a hybrid sequence, called a mixed sequence, and applications of this sequence to numerical integration and mathematical finance. In Chapter 3 we will introduce a hybrid-Monte Carlo method whose estimates satisfy deterministic error bounds, as in quasi-Monte Carlo theory, yet statistical analysis can be applied to measure accuracy, as in Monte Carlo theory. Numerical results obtained upon application of this method to problems from particle transport theory and mathematical finance will be presented. Finally, in Chapter 4, we will introduce a generalized quasi-Monte Carlo integration rule and prove a Koksma-Hlawka type inequality for the rule using a recently introduced concept, "variation in measure". Error reduction techniques, and in particular "importance sampling", will be derived as consequences of this rule.

Chapter 2

Simulation in High Dimensions

2.1 Introduction

The so called "mixed" sequences have been first introduced by Spanier [1] as a hybrid-Monte Carlo method to generate Markovian random walks in order to solve some particle transport problems by simulation. Briefly, a mixed sequence is an $s-$dimensional sequence whose elements are vectors obtained by concatenating $d-$dimensional vectors from a low-discrepancy sequence with $(s-d)-$dimensional random vectors. A formal definition will be given in the next section. In the last few years, a good amount of numerical work has been done by several authors applying these sequences to particle transport problems (see [2], [3]). It has been observed that mixed sequences perform consistently better than pseudorandom and quasirandom implementations in "difficult" problems. The difficulty of a particle transport problem, in which random walks of arbitrary length must be simulated, is determined by the rate of convergence of the Neumann series to the solution of the problem. The strategy in applying these sequences, in general, is to choose an "optimal" dimension for the quasirandom part of the mixed sequence in a heuristic way and then fill out the remaining "slots" with pseudorandom vectors.

In Section 2.2, a probabilistic result that relates the discrepancy of the mixed sequence to the discrepancy of its quasirandom part will be given. To compare the effectiveness of mixed sequences with that of low-discrepancy and pseudorandom sequences, we will apply them to problems from numerical integration in Section 2.3 and problems from mathematical finance in Section 2.4.

We review some of the basic definitions.

Definition 1 *Let* $\{x_n\}_{n=1}^{\infty}$ *be a sequence in* $I^s = [0,1)^s$. *For* $\alpha = (\alpha_1, ..., \alpha_s) \in \bar{I}^s = [0,1]^s$ *and* $N > 0$ *the local discrepancy is defined as*

$$g(\alpha; x_n) = \frac{1}{N} \sum_{n=1}^{N} 1_{[0,\alpha)}(x_n) - \alpha_1 \cdots \alpha_s,$$

where $[0,\alpha) = \prod_{i=1}^{s} [0,\alpha_i)$ *and the function* 1_R *is the characteristic function of the set* R. *The star discrepancy of the sequence* x_n *is then defined by*

$$D_N^*(x_n) = \sup_{\alpha \in \bar{I}^s} |g(\alpha; x_n)| .$$

Note that $D_N^*(x_n) \leq 1$ *for all* N.

2.2 Mixed(s,d) Sequences

Definition 2 *For a given* $d-$*dimensional sequence* $\{q_n\}_{n=1}^{\infty}$, *a mixed* (s,d) *sequence is an* $s-$*dimensional sequence* $(s > d)$ *described by*

$$m_n = (q_n, U_n) \qquad n = 1, 2, ...,$$

where $U_1, U_2, ...$ *are* $(s-d)$ *-dimensional i.i.d. vectors drawn from the uniform distribution on* $[0,1]^{s-d}$.

Lemma 1 *Let $m_n = (q_n, U_n)$ be a mixed (s,d) sequence. For $\alpha = (\alpha_1, ..., \alpha_s) \in \bar{I}^s$, put $\alpha' = (\alpha_1, ..., \alpha_d)$.*

1. For a fixed α, the expected value of $g(\alpha; m_n)$ is given by

$$E[g(\alpha; m_n)] = \alpha_{d+1} \cdots \alpha_s \, g(\alpha'; q_n).$$

2. For any α, the variance of $g(\alpha; m_n)$ satisfies

$$Var[g(\alpha; m_n)] \leq \frac{1}{4N} \left(D_N^*(q_n) + 1 \right).$$

Proof. First let's observe that

$$1_{[0,\alpha)}(m_n) = \begin{cases} 1_{[0,\alpha')}(q_n) & \textit{with probability } \alpha_{d+1} \cdots \alpha_s \\ 0 & \textit{with probability } 1 - \alpha_{d+1} \cdots \alpha_s \end{cases}.$$

Then, $E[1_{[0,\alpha)}(m_n)] = \alpha_{d+1} \cdots \alpha_s \, 1_{[0,\alpha')}(q_n)$. We have

$$\begin{aligned} E[g(\alpha; m_n)] &= \frac{1}{N} \sum_{n=1}^{N} E[1_{[0,\alpha)}(m_n)] - \alpha_1 \cdots \alpha_s \\ &= \frac{1}{N} \sum_{n=1}^{N} \alpha_{d+1} \cdots \alpha_s \, 1_{[0,\alpha')}(q_n) - \alpha_1 \cdots \alpha_s \\ &= \alpha_{d+1} \cdots \alpha_s \left(\frac{1}{N} \sum_{n=1}^{N} 1_{[0,\alpha')}(q_n) - \alpha_1 \cdots \alpha_d \right) \\ &= \alpha_{d+1} \cdots \alpha_s \, g(\alpha'; q_n). \end{aligned}$$

This completes the proof of the first part of the lemma. To prove the second part, we need the following facts: First we observe that

$$Var[1_{[0,\alpha)}(m_n)] = E[1_{[0,\alpha)}(m_n)^2] - E[1_{[0,\alpha)}(m_n)]^2$$

$$
\begin{aligned}
&= \alpha_{d+1} \cdots \alpha_s \, 1_{[0,\alpha')}(q_n) - \alpha_{d+1}^2 \cdots \alpha_s^2 \, 1_{[0,\alpha')}(q_n) \\
&= \alpha_{d+1} \cdots \alpha_s \, (1 - \alpha_{d+1} \cdots \alpha_s) 1_{[0,\alpha')}(q_n).
\end{aligned}
$$

Furthermore, from

$$
1_{[0,\alpha)}(m_i) \cdot 1_{[0,\alpha)}(m_j) = \begin{cases} 1_{[0,\alpha')}(q_i) \cdot 1_{[0,\alpha')}(q_j) & \textit{with probability } \alpha_{d+1}^2 \cdots \alpha_s^2 \\ 0 & \textit{with probability } 1 - \alpha_{d+1}^2 \cdots \alpha_s^2 \end{cases},
$$

we get

$$
E[1_{[0,\alpha)}(m_i) \cdot 1_{[0,\alpha)}(m_j)] = \alpha_{d+1}^2 \cdots \alpha_s^2 \cdot 1_{[0,\alpha')}(q_i) \cdot 1_{[0,\alpha')}(q_j),
$$

and

$$
E[1_{[0,\alpha)}(m_i)] E[1_{[0,\alpha)}(m_j)] = \alpha_{d+1} \cdots \alpha_s \, 1_{[0,\alpha')}(q_i) \cdot \alpha_{d+1} \cdots \alpha_s \, 1_{[0,\alpha')}(q_j).
$$

Thus

$$
Cov[1_{[0,\alpha)}(m_i), 1_{[0,\alpha)}(m_j)] = 0.
$$

Using these results, we obtain:

$$
\begin{aligned}
Var[g(\alpha; m_n)] &= \frac{1}{N^2} \sum_{n=1}^{N} Var[1_{[0,\alpha)}(m_n)] \\
&= \frac{1}{N^2} \sum_{n=1}^{N} \alpha_{d+1} \cdots \alpha_s \, (1 - \alpha_{d+1} \cdots \alpha_s) 1_{[0,\alpha')}(q_n) \\
&= \frac{\alpha_{d+1} \cdots \alpha_s (1 - \alpha_{d+1} \cdots \alpha_s)}{N} \left(\frac{1}{N} \sum_{n=1}^{N} 1_{[0,\alpha')}(q_n) \right) \\
&\leq \frac{1}{4N} \left(\frac{1}{N} \sum_{n=1}^{N} 1_{[0,\alpha')}(q_n) - \alpha_1 \cdots \alpha_d + \alpha_1 \cdots \alpha_d \right) \\
&= \frac{1}{4N} \left(g(\alpha'; q_n) + \alpha_1 \cdots \alpha_d \right)
\end{aligned}
$$

$$\leq \quad \frac{1}{4N}\left(D_N^*(q_n)+1\right). \blacksquare$$

Theorem 1 *Let q_n be a $d-$dimensional sequence in $[0,1)^d$ and m_n be the corresponding mixed (s,d) sequence. Then for any $\varepsilon > 0$*

$$D_N^*(m_n) \leq \varepsilon + D_N^*(q_n),$$

with probability greater than or equal to

$$1 - \frac{1}{\varepsilon^2}\frac{1}{4N}\left(D_N^*(q_n)+1\right)$$

(this is a probability for sufficiently large N).

Proof. Let $\varepsilon > 0$. From Chebyshev's inequality, we have

$$P\left(-\varepsilon + E[g(\alpha;m_n)] < g(\alpha;m_n) < \varepsilon + E[g(\alpha;m_n)]\right) \geq 1 - \frac{1}{\varepsilon^2}Var[g(\alpha;m_n)] \quad (2.1)$$

for a fixed $\alpha \in \bar{I}^s$. From the lemma, we can rewrite (2.1) as

$$\begin{aligned} P\left(-\varepsilon + E[g(\alpha;m_n)] < g(\alpha;m_n) < \varepsilon + E[g(\alpha;m_n)]\right) & \\ \geq \quad 1 - \frac{1}{\varepsilon^2}Var[g(\alpha;m_n)] & \\ \geq \quad 1 - \frac{1}{\varepsilon^2}\frac{1}{4N}\left(D_N^*(q_n)+1\right), & \end{aligned}$$

in other words, $g(\alpha;m_n)$ is in the interval $(-\varepsilon + E[g(\alpha;m_n)], \varepsilon + E[g(\alpha;m_n)])$ with probability greater than or equal to $p = 1 - \frac{1}{\varepsilon^2}\frac{1}{4N}\left(D_N^*(q_n)+1\right)$. Each statement in the following chain of implications is true with probability greater than or equal to p :

$$-\varepsilon + E[g(\alpha; m_n)] < g(\alpha; m_n) < \varepsilon + E[g(\alpha; m_n)]$$

$$\begin{aligned}
&\Rightarrow \quad |g(\alpha; m_n)| \leq \max\left(|-\varepsilon + E[g(\alpha; m_n)]|, |\varepsilon + E[g(\alpha; m_n)]|\right) \\
&\Rightarrow \quad |g(\alpha; m_n)| \leq \sup_{\alpha} \max\left(|-\varepsilon + E[g(\alpha; m_n)]|, |\varepsilon + E[g(\alpha; m_n)]|\right) \\
&\Rightarrow \quad |g(\alpha; m_n)| \leq \max\left(\sup_{\alpha} |-\varepsilon + E[g(\alpha; m_n)]|, \sup_{\alpha} |\varepsilon + E[g(\alpha; m_n)]|\right).
\end{aligned}$$

Using the first part of the lemma, we can find an upper bound for the right hand side of the above inequality. Indeed

$$\begin{aligned}
\sup_{\alpha} |-\varepsilon + E[g(\alpha; m_n)]| &\leq \varepsilon + \sup_{\alpha} |E[g(\alpha; m_n)]| \\
&= \varepsilon + \sup_{\alpha} |\alpha_{d+1} \cdots \alpha_s \, g(\alpha'; q_n)| \\
&= \varepsilon + D_N^*(q_n),
\end{aligned}$$

and similarly

$$\sup_{\alpha} |\varepsilon + E[g(\alpha; m_n)]| \leq \varepsilon + D_N^*(q_n).$$

Therefore, the last inequality of the above chain of inequalities implies $|g(\alpha; m_n)| \leq \varepsilon + D_N^*(q_n)$ with probability p.

Neither the probability p nor the right hand side of the above inequality depends on α; hence we can take the supremum of the left hand side of the inequality to conclude

$$D_N^*(m_n) \leq \varepsilon + D_N^*(q_n),$$

with probability greater than or equal to

$$1 - \frac{1}{\varepsilon^2} \frac{1}{4N} \left(D_N^*(q_n) + 1\right). \blacksquare$$

Corollary 1 *Put* $\varepsilon := (\varepsilon_N) = (\frac{1}{N^{a/2}})$, $0 < a < 1$, *in the above theorem. Then*

$$D_N^*(m_n) \leq \frac{1}{N^{a/2}} + D_N^*(q_n), \tag{2.2}$$

with probability greater than or equal to

$$p_N = 1 - \frac{1}{4N^{1-a}} \left(D_N^*(q_n) + 1 \right).$$

Now let q_n be a $d-$dimensional low-discrepancy sequence with[1]

$$D_N^*(q_n) \leq c_d \frac{(\log N)^d}{N} + O\left(\frac{(\log N)^{d-1}}{N} \right). \tag{2.3}$$

Then, from (2.2), we have

$$D_N^*(m_n) \leq \frac{1}{N^{a/2}} + c_d \frac{(\log N)^d}{N} + O\left(\frac{(\log N)^{d-1}}{N} \right), \tag{2.4}$$

with probability greater than or equal to

$$\begin{aligned} p_N &= 1 - \frac{1}{4N^{1-a}} \left(D_N^*(q_n) + 1 \right) \\ &\geq 1 - \frac{1}{4N^{1-a}} \left(c_d \frac{(\log N)^d}{N} + O\left(\frac{(\log N)^{d-1}}{N} \right) + 1 \right) \\ &= 1 - \frac{1}{4N^{1-a}} - \frac{c_d}{4} \frac{(\log N)^d}{N^{2-a}} - O\left(\frac{(\log N)^{d-1}}{N^{2-a}} \right). \end{aligned} \tag{2.5}$$

Hence, as $N \to \infty$

$$D_N^*(m_n) \to 0 \quad as \quad \frac{1}{N^{a/2}} + c_d \frac{(\log N)^d}{N} + O\left(\frac{(\log N)^{d-1}}{N} \right) \to 0,$$

[1] For example, q_n can be chosen as a (t, s) sequence (see [4]) in dimension d.

and the corresponding probabilities p_N satisfy

$$p_N \to 1 \ as \quad \frac{1}{4N^{1-a}} + \frac{c_d}{4}\frac{(\log N)^d}{N^{2-a}} + O\left(\frac{(\log N)^{d-1}}{N^{2-a}}\right) \to 0,$$

where the exponents satisfy $0 < 1 - a < 1$ and $1 < 2 - a < 2$.

Remark 1 *We should emphasize the fact that the mixed sequence m_n consists of $(s-d)$ - dimensional i.i.d. uniform vectors $U_1, U_2, ...$ and a deterministic $d-$dimensional vector sequence q_n. Although the error bound given in Corollary 1 shows the dependence of $D_N^*(m_n)$ on $D_N^*(q_n)$, it does not reflect the additional error introduced upon generating the i.i.d. vectors $U_1, U_2, ...$ using say, a pseudorandom vector algorithm. An error bound incorporating this additional error component will include a term that depends on $s - d$; the dimension of the pseudorandom component of m_n. Exactly how the incorporation of this component of error will affect the right hand side of equation (2.2) is an open question. Therefore one should not draw the conclusion that in a practical implementation making use of m_n, the discrepancy of the sequence is minimized by minimizing the dimension d. In that case, we would not use the deterministic vector sequence q_n at all, since we already have a "random" sequence available! In a practical implementation, the rule of thumb is to choose d to be the maximum dimension for which one is comfortable using a low-discrepancy sequence, unless it is possible to quantify and balance the two components of error that result as the parameter d varies between 0 and s.*

It might be illustrative to calculate the upper and lower bounds of equations (2.3), (2.4) and (2.5), to compare a low-discrepancy sequence q_n in dimension s with the corresponding mixed (s, d) sequence. We choose q_n to be the $(t, s)-$sequence as constructed by Niederreiter in [4]. For $2 \leq s \leq 20$, Niederreiter calculates the values c_s that appear in (2.3) by optimizing the choice of the base of the sequence. Let m_n

be the corresponding mixed (s, d) sequence. Omitting lower order terms, let

$$A_1 = c_s \frac{(\log N)^s}{N}$$

be the upper bound of the discrepancy of q_n (see equation (2.3) when dimension is s),

$$A_2 = \frac{1}{N^{a/2}} + c_d \frac{(\log N)^d}{N}$$

be the probabilistic upper bound of the discrepancy of m_n (see equation (2.4)), and

$$A_3 = 1 - \frac{1}{4N^{1-a}} - \frac{c_d}{4} \frac{(\log N)^d}{N^{2-a}}$$

be the lower bound for p_N (see equation (2.5)). In the following table, A_1, A_2 and A_3 are calculated for $s = 4, 6, ..., 20$. We take $d = \frac{s}{2}$, $a = 0.9$ and $N = 10^6$.

s	A_1	A_2	A_3
4	0.0031	0.0020	0.937
6	0.1293	0.0024	0.937
8	3.9683	0.0051	0.937
10	108.421	0.0144	0.936
12	2707.64	0.1313	0.929
14	202107	0.3968	0.912
16	1.4×10^6	3.9703	0.688
18	2.8×10^7	11.095	0.241
20	2.1×10^9	108.423	-5.87

Although these numbers are of little practical use (for example, we know that discrepancy is always less than or equal to 1), they nevertheless suggest improvements that may be attained upon using mixed sequences in large dimensions. They also

show that although the upper bound given by (2.4) is of $O(N^{-a/2})$ asymptotically, for practical values of N ($N \sim 10^6, 10^7$) and large values of s ($s \sim 20, 40$), it is indeed smaller than the upper bound A_1 established for the corresponding low-discrepancy sequence. In the next sections, we will see numerical evidence in support of the improvements furnished by mixed sequences.

2.3 Applications to Numerical Integration

In this section, we will use the Sobol, Faure and Niederreiter (base 2) sequences (see [5], [6] and [7] for the implementations of these sequences), their mixed versions and pseudorandom numbers[2] to numerically calculate the integrals of some test functions. These test functions have been used by several authors to compare performances of low-discrepancy and pseudorandom sequences. Using them, we want to investigate numerically the possibility that mixed sequences exhibit superior behavior in higher dimensions. In other words, the advantage of these sequences is expected to be realized when the performance of the low-discrepancy sequences starts degrading because of the large dimension of the particular integrand whose integral is being estimated. Of course, the dimension at which a mixed sequence gives better estimates than the corresponding low-discrepancy sequence depends on the integrand.

The test functions have known integrals so that we know the exact error of numerical integration. In the following graphs, the absolute value of error produced by the underlying sequence is plotted against the number of samples N, where $N = 20000, 40000, ..., 10^6$.

1. $f_1(x) = \prod_{i=1}^{s} |4x_i - 2|$

   The integral of the function on the $s-$dimensional unit cube is 1. When $s = 20$,

[2]The numbers are generated using the built-in pseudorandom number generator of Microsoft FORTRAN PowerStation, version 1.0, in which language the codes used in this chapter are written.