NEW YORK UNIVERSITY
STERN SCHOOL OF BUSINESS
FINANCE DEPARTMENT

Working Paper Series, 1995

The Use of Low Discrepancy Points in Valuing Complex Financial Instruments.

Lord, Graham, Spassimir Paskov, and Irwin T. Vanderhoof.

FIN-95-27

The Use of Low Discrepancy Points in Valuing Complex Financial Instruments

by Graham Lord, Spassimir Paskov, and Irwin T. Vanderhoof

Abstract: Modern finance has evolved the use of very complex financial instruments. Stock and interest rate options fit this description. Another example of such an instrument would be a mortgage pool involving many tranches and providing relationships between the tranches so that the payoff on one tranche depends upon the amounts paid upon other tranches over the whole history of the pool. Since the valuation of this last instrument would involve a separate probability distribution for each period over the whole period of the pool, the calculation could involve 360 separate probability distributions over the whole period. It would require, then, a multiple integration over all these periods, all 360 of them. Such calculations are generally not possible on an exact basis so that numerical integration must be used. In such an environment only Monte Carlo methods are practical. Certain selected sequences of values called "low discrepancy points" are theoretically more efficient in this kind of calculation than the random numbers usually generated for Monte Carlo calculations. This paper discusses the theoretical basis for such a claim (Niederreiter [8] covers much of the material in a more rigorous fashion.), the calculation of such points, and illustrations of the results of using such methods on several real problems.

The Nature of Low Discrepancy Points

The term "low discrepancy points" is not commonly used in numerical analysis or statistics and a definition of the term is a necessary place to start this discussion. We shall provide a heuristic definition. The term derives from something called the L_2 discrepancy problem.

The L_2 discrepancy problem starts with a unit hypercube - that is a cube of more than three dimensions. Each edge of the cube has a length of one unit. The volume of the cube is therefore 1. Let us assume that there is a large number of points to be distributed within the cube. How can these points be distributed so that if any volume in the cube is selected the proportion of the points within the volume is as "close" as possible to the volume itself?

It seems obvious that such a distribution of points should be possible. It should also be obvious that unless the number of points is very large indeed that the results will not be satisfactory. The discrepancy between the percentage of the total number of points and the volume will differ for a higher portion of all the possible volumes as the number of points decreases. It is not so obvious, but true nonetheless, that a uniformly spaced grid is not the appropriate solution for the problem. The use of a grid leaves large well defined volumes that have no points in them!

Points which provide, on the average, a close fit between these two numbers provide a low discrepancy and are therefore so named.

Such a selection of points was originally of interest in number theory, but the application is not obvious. Suppose, however, that each dimension of the cube is identified as the value of a Cumulative Distribution Function (CDF) of random variates. If the possible values of the variate run from 1 through 10 then the CDF for values of the variate less than 1 is zero and the value of the CDF for values of the variate greater than 10 is 1. If the value of the variate is set at 4 , and the corresponding CDF is .3, this means that for a very large number of selections of the variate the percentage that would be less than 4 would be 30%, or .3 and the proportion over 4 would be .7. Further, if the value of the variate is set at 5 and the value of the CDF is .65 we can say that 65% of the trials would produce a result less than five. We can further say that the probability of a value between 4 and 5 is 35%.

In addition, we can note that for the probability distributions that are common in financial usage, there is a single value of the variate corresponding to each value of the CDF. The CDF of .3 corresponds to a value of 4 of the variate and a value of the CDF of .5 would correspond to a value of the variate between 4 and 5. The exact value could be calculated subject to complete specification of the probability distribution. If we are dealing with a series of independent variates then the dimensions of our hypercube could be interpreted as the separate CDF's corresponding to these variables. A volume within the hypercube would then be interpreted as the measure of the probability of occurrence of an event within the range of the variates corresponding to the values of the CDF's that define the volume.

Let us look at an example. Supposing a volume inside such a cube is defined by .3 to .65, .3 to .65 and 0 to 1. If all three random variates have the distribution described above then the probability of an event between 4 and 5 in the first and second dimensions and any value for the third dimension would be .35 x .35 x 1 = .1225 = 12.25%.

The application of low discrepancy points to Monte Carlo problems can thus be formulated. If we interpret the hypercube as one whose linear dimensions are the values of CDF's of the different random variables we are using in a calculation and these variates can be structured so that they are independent, then each point represents values of all the variables involved. If the variables are interest rates for each month over the next 30 years, then a single point represents a path for interest rates over that whole period. If the number of points is very large we could say that the number of samplings from within each probability volume is proportional to the probability itself. This is one characteristic that we would expect if the numbers were chosen at random. In addition we would expect that there would be the same size probability volume surrounding each of the points. If one more point was added we would expect that there would be a redistribution of the volumes such that the statement would still be true.

The final conclusion is as follows. By using low discrepancy points to define the CDF's of the variables and using the inverse function to solve for the actual variates themselves, we have availed ourselves of the most efficient method of sampling the probability space. This approach should dependably provide the values we expect to get from the use of random numbers in Monte Carlo calculations.

There are, however, several caveats. The results described would be true for very large number of such points. In theory, there should be advantages in the use of low discrepancy points for problems with a low number of dimensions. It is not clear how this advantage would change as the number of dimensions increases. Low discrepancy points should always be at least as efficient as random numbers. However, for there to be adequate reason to change to a new method of implementation of Monte Carlo, there should be evidence that the advantages are real for smaller numbers of calculations in higher dimensions. This is subject to empirical investigation.

Calculation of Low Discrepancy Points for Monte Carlo

The essence of these low discrepancy points is that they are anti-correlated. The earliest of such sequences of numbers seems to be the van der Corput sequence [1].

Construction of a van der Corput sequence would require the initial choice of the number of terms, "N", to be calculated and the base in which numbers were to be expressed. Let the base chosen be "p", a prime number. Express each number, "n"<N in base p.

Any non-negative integer can be expressed :

$n = {}_j\Sigma\ c_j\ p^j\ , j \geq 0$ (For example, $7 = (1x3^0 + 2x3^1)$)

Define the radical-inverse function "P" in base p by

$P(n) = \Sigma\ c_j\ p^{-j-1}$ ($P(7) = 1/3 + 2/9$)

Note that $0 \leq P(n) < 1$. The van der Corput sequence in base p is then;

P(0), P(1), P(2), ...P(n).

Hammersley[2] says "Furthermore, we may justifiably suspect that there exists a deterministic way of chosing the N points in the hypercube that will be better than the Monte Carlo method." He specifically suggested that for a two dimensional problem the van der Corput sequence could be used to develop a series of points that should be efficient for Monte Carlo calculations. His argument was in terms of the low discrepancy characteristic of these points. The coordinates of these points in two dimensions would be:

$$n/N,\ P(n).$$

Hammersley further conjectured that for higher dimensional problems a more complex set of coordinates could be developed. This sequence would involve a series of prime numbers: p_1, p_2, p_3, etc., and the corresponding radical inverse functions P_1, P_2, P_3, etc. The coordinates of these points within the hypercube would be

$$n/N,\ P_1(n),\ P_2(n), P_3(n), \ldots.$$

The set of points from n= 1 to n=N are Hammersley points. Halton[3] showed that the discrepancy for this sequence has a defined upper limit which is less than a power of log N and is a result favorable compared to random numbers. (The points whose coordinates are

$P_1(n)$, $P_2(n)$, $P_3(n)$,.... are referred to as Hammersley-Halton points or Halton points.)

Roth[4] has shown that the L_2 discrepancy must be at least on the order of

$n^{-1}(\log n)^{(d-1)/2}$ and that this result is sharp.

This discrepancy is minimized for Hammersley points with a small and <u>unknown</u> amount added to the first dimensional coordinate.

Woźniakowski[5], and Traub and Woźniakowski[6] [7] (and other publications) established that these "shifted" Hammersley points would be efficient in Monte Carlo calculations.

There are, however, at least two problems with Hammersley points. The first is that the amount by which the first dimension must be shifted is unknown. The theoretical result is therefore not being achieved. The use of classical Hammersley points is possible with some minor loss in efficiency. The second problem is that the construction of Hammersley points is such that if an additional point is to be added, a reconstruction of all previously calculated points is necessary.

For these reasons other sets of low discrepancy points have been investigated and used in Monte Carlo calculations. Faure points and Sobol points have been used with success as recounted later. Niederreiter Points are another variation in this methodology which has not yet been tested in application to financial problems.

Briefly, Faure points are constructed by choosing a value of the prime number "p" higher than the dimensionality of the problem and calculating the values of ${}^1c_j$ from

$n = \Sigma\, {}^1c_j\, p^j$. The values of higher order coefficients are then calculated by recursion.

${}^kc_j(n) = \Sigma\ C(i,j,)\ {}^{k-1}c_i(n)$ (mod p) where the summation is over i to N with $i \geq j$. $C(i,j) = i!/(j!(i-j)!)$

The calculation for Sobol points is more complex and the reader may refer to Niederreiter[8] for more details and a more comprehensive discussion of the various low discrepancy series in Monte Carlo calculations.

Other series are under study. However, the characteristic of all these multidimensional points is that additions to the series fill up gaps left by previous entries. Because of their anti-correlated attribute, they seem to find the empty places in the matrix and fill them up. Figure 1 is a two dimensional graph of points whose coordinates are chosen by standard random number generators. Figures 2 and 3 are respectively 512 and 1024 points whose coordinates are developed from the Faure sequence.

Applications

Three reports are known to exist on the use of low discrepancy points in the Monte Carlo valuation of complex financial instruments. Rupert Brotherton-Ratcliffe of General Re. made a presentation[9] (unpublished) at a meeting of the Society of Actuaries on May 23, 1995. Joy, Boyle, and Tan[10] have presented their findings at a seminar at Columbia. Paskov[11] has also made several presentations of his conclusions.

Before presenting some detailed discussion of each of these studies we will present the conclusions reached by all of the authors. They are as follows:

1) The calculation of the low discrepancy points (quasi-random numbers) is easier and less expensive than the use of the common random number generators.

2) The convergence of the calculations is more rapid by a factor of 5 to 10, and smoother, than that found using common random number generators.

3) The actual values to which the functions converged when the common pseudo-random number generators were used indicated

a bias which seemed to be a dependence upon the starting seed used in the random number generator.

4) The convergence of the low discrepancy point calculations seemed to be considerably better than the theoretical expectation.

The Brotherton-Ratcliffe study concerned the valuation of stock options where the stock price followed a standard Weiner process and the option was based upon the arithmetic or the geometric average price of the stock. Both Faure and Sobol points radically outperformed random numbers for four periods. For 48 periods the Sobol points again outperformed random numbers.

The Joy, Boyle, Tan study centered around the use of Faure points in their calculations. They studied an option on the geometric mean of three assets, an option on a basket of 60% light sweet crude and 40% natural gas, an Asian option involving averaging over 52 periods, and a six month option to enter into a 1 year natural gas swap. They attempted to use realistic models for the prices of the securities.

They conclude that the Faure sequences provide more rapid and regular convergence than do common random number generators for a number of reasons. Perhaps the most important is that the random number generators generate numbers according to some set formulas. In very high dimensions there may be patterns which bias the results. Another possible reason might be that even if the numbers are random we would expect some results to be very bad as well as some being very good - just by chance. Low discrepancy points should give dependable results.

Additionally these calculations do not take into consideration the regularity of the integrand. Financial functions may be especially smooth and therefore well adapted to calculations using low discrepancy points.

The Paskov study is probably the most comprehensive work on this subject to date. The object studied was a Collateralized Mortgage Obligation (CMO) called "Fannie Mae REMIC Trust 1989-23" This is a pool of 30 year mortgages having therefore 360 cash flows. Interest rates were assumed to vary according to a log normal distribution and prepayments were modeled according to an arctangent function (the form was suggested by Goldman Sachs as one which had been used on Wall Street in the past). The CMO had ten tranches including the Z tranche. The distributed results focused upon the A tranche as a representative example of each of these financial interests.

In most cases there were 1,000,000 iterations of the calculations to arrive at an answer. Calculations were done using Halton points, Sobol points and the random number generators Ran1 and Ran2 [12]. Figure 4 shows a result for Tranche A using Ran1

(1988 edition of Numerical Recipes) with four different seeds to start the calculation. The results were similar, though not so striking for Ran2.

This study also made a comparison with the use of antithetic variables as a variance reduction method. The technique of antithetic variables involves the creation of a new function that will have the same expected value as the original function but a lower standard deviation. Since the error bound in Monte Carlo calculations is proportional to the standard deviation of the function and inversely proportional to the square root of the number of points in the calculation, a reduction in the standard deviation will reduce the number of points needed for a given level of expected average accuracy.

In this comparison also, Sobol points provided more rapid and smooth convergence than the antithetic variable approach even using the superior Ran2 random numbers. Figure 5 shows the results of Sobol points, Halton points, and two runs of antithetic variables using Ran2 (1992 edition of Numerical Recipes). It was also found that the Sobol points converged to a result that was the average of twenty separate runs using antithetic variables with different seeds for Ran2. Considering all 10 tranches and 20 runs (200 cases), the Sobol point run was closer to this average than the antithetic variable runs in 131 cases and closer than traditional Monte Carlo in 171 cases. Figure 6 illustrates this result.

Finally, calculations were performed using small numbers of points to test the results under this restriction. The use of Sobol points still provided superior results.

Conclusion

In tests involving different persons, different institutions, different financial instruments, and different methodologies, low discrepancy points have proved more efficient than pseudo-random number generators in speed and accuracy. They must be seriously considered in the future in types of calculation involving complex financial instruments. Two remaining matters need to be addressed. They are the paradox of the marked superiority of these methods over common random number generators, and the other advantages of the use of these techniques over variance reduction methodology.

While there are theoretical demonstrations of the superiority of low discrepancy points over random numbers in low dimensions, this advantage may diminish in high dimensional calculations. However, we have three separate demonstrations that this is not the case. The superiority seems most evident in the high dimensional calculations. The argument that this is somehow due to chance does not seem plausible. If that were the case many of the results would go the other way and the superiority would not be so striking.

The nature of financial calculations is often that the early results are most important and that later events have less impact on the final result due to discounting. If we are dealing with thirty year instruments then the results of the first five or ten years have the greatest impact on present values. The second named author speculates that since Sobol points are more evenly distributed in the early periods of the calculation the results we obtain from these low discrepancy points are more robust than those from pseudo-random numbers.

The last named author speculates that the real reason is that the theoretical results attributed to Monte Carlo assume that there is a source of random numbers. However, no one knows how to produce true random numbers. Any mathematical formula will produce some kind of cycle and therefore some kind of bias. This seems to regularly show up in studies of higher dimensional arrays. A physical process would also have some kind of bias depending upon the actual experimental arrangements. Realistically, no one even wants a truly random process. We want numbers that can be reused so that we can calibrate a new formula to old results and so that we can conveniently improve accuracy by increasing the number of points. Low discrepancy points will be used because they produce the results theoretically predicted for random numbers but are more convenient than actually trying to get random numbers. A comparison with the use of variance reduction techniques must mention the fact that the variance reduction technology requires specific special research for each new problem. The low discrepancy points can be used where ever the CDF's are known.

The last point to mention is the other major advantage in the use of low discrepancy points. That advantage is that they provide a full sampling of the entire space. The first and last named authors have used Faure points in the construction of interest rate scenarios for experiments on asset liability management. They were used rather than pseudo-random scenarios because with Faure points we could be assured that there was reasonable representation of all the possible future according to the assumed interest rate distributions. Random scenarios could not assure that there was reasonable representation of all the possibilities. This last point also applies to the comparison with the effectiveness of variance reduction techniques. Even if such a technique were to produce as rapid a convergence, the essence of variance reduction techniques is the loss of the shape of the function. All higher moments are lost. The use of low discrepancy points preserves all higher moments of the probability distributions and therefore can be used for judgement of risk, as well as expected return, on an asset. Since each of the results has approximately the same probability of occurrence, we have a representation of the CDF for the results of the entire simulation. This is the first known practical application of the low discrepancy points for "reconstruction of a probability surface", a concept often mentioned in the literature.

1.van der Corput, J. C. 1931. Diophantische Ungleichungen. I. Zur Gleichverteilung Modulo Eins. Acta Math. (Stockholm). 52:373-456

van der Corput, J. C. 1932. Diophantische Ungleichungen. II. Rhythmische Systeme. Acta Math. (Stockholm). 59: 209-328

2.Hammersley, J. M. 1960. Monte Carlo Methods for Solving Multivariate Problems. Annals NY Academy of Science. 86.3 844-873

3.Halton, J. H. 1960. On the efficiency of certain quasi-random sequences of points i evaluating multi-dimensional integrals. Numer. Math. 2: 84-90

4.Roth, K. F. 1954. On Irregularities of distribution. Mathematica I: 73-79

Roth, K. F. 1980. On Irregularities of distribution IV. Acta. Arith. 37: 67-75

5.Woźniakowski, H. 1991. Average Case Complexity of Multivariate Integration. Bulletin Amer. Math. Soc. (new Series). 24: 1

6.Traub, J. F. and Woźniakowski, H. 1991. Theory and Applications of Information-Based Complexity. SFI Studies in the Science of Complexity Lect. III. Addison-Wesley

7.Traub, J. F. and Woźniakowski. 1994. Breaking Intractability. Scientific American 270.1: 102-107

8.Niederreiter, H. 1992. Random Number Generation and Quasi-Monte Carlo Methods. SIAM

9.Using Quasi-Random Monte Carlo Simulation to Price Path-Dependent Stock Options.

10.Joy, C., Boyle, P. P., and Tan,Ken Seng. Their paper is titled Quasi- Monte Carlo Methods in Numerical Finance. It has not yet been published. The senior author is located at the Enron Corporation, 1400 Smith Street, Houston Texas. Phelim Boyle is at the School of Accountancy at the University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada.

11.Paskov, Spassimir. 1994. Analysis of Multivariate Problems with Applications to Finance. Unpublished dissertation accepted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the Graduate School of Arts and Sciences of Columbia University

12.Press, W., Teukolsky, S., Vetterling, W., and Flannery, B. 1988 rev. 1992. Numerical Recipes, Cambridge University Press

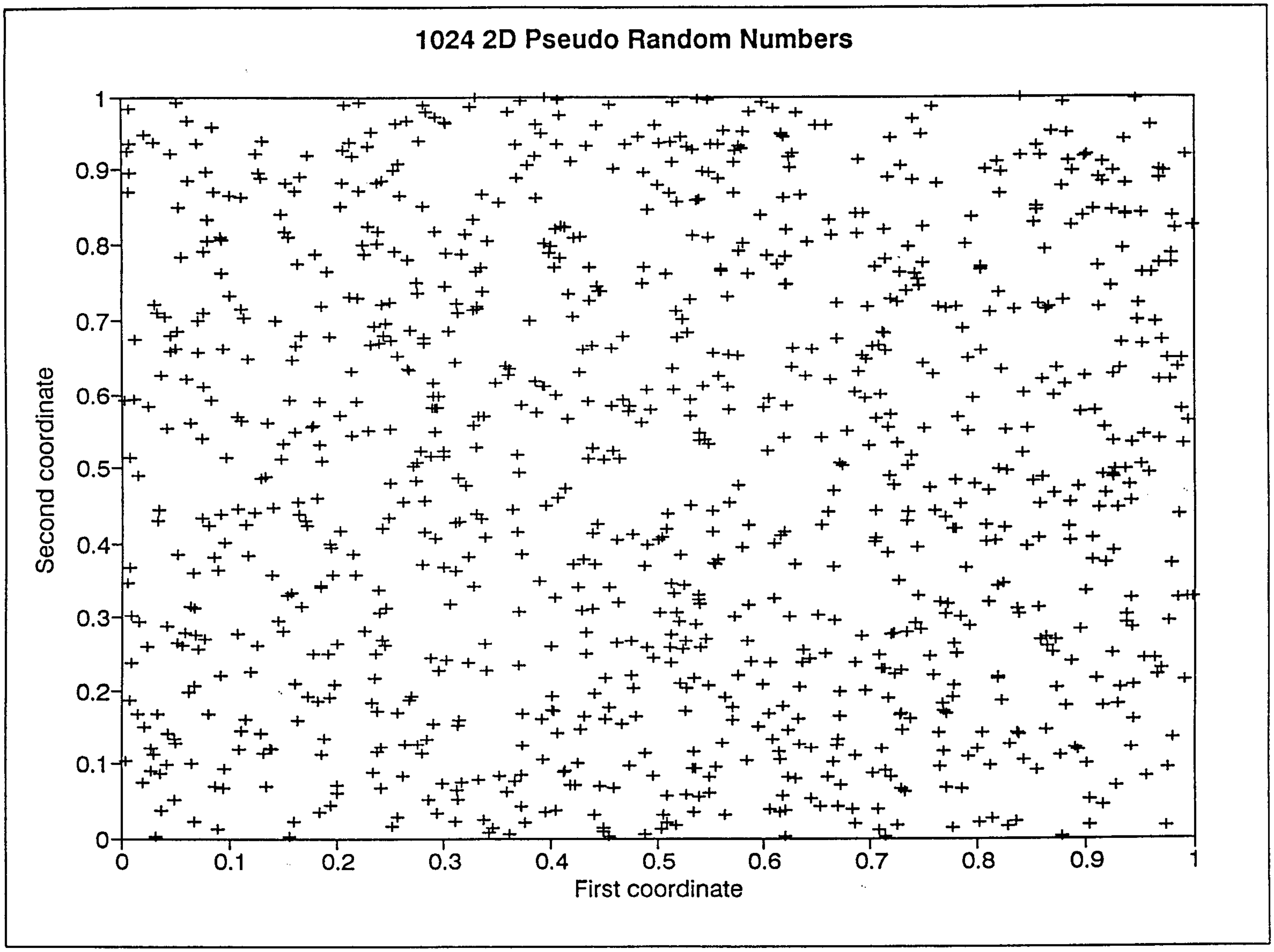
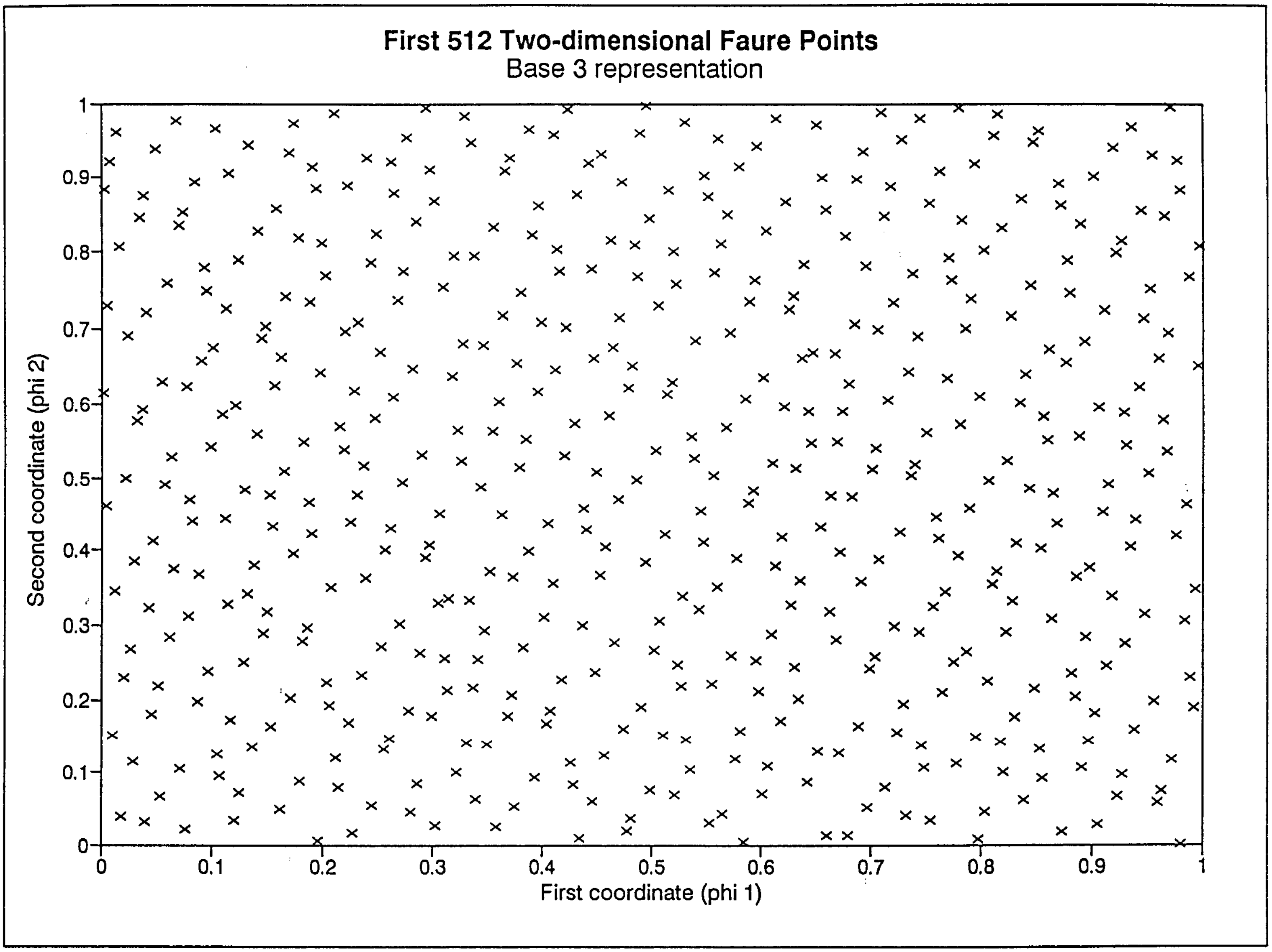

1024 2D Pseudo Random Numbers
Second coordinate
First coordinate
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Fig 1

(10)



First 512 Two-dimensional Faure Points
Base 3 representation
Second coordinate (phi 2)
First coordinate (phi 1)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Fig 2

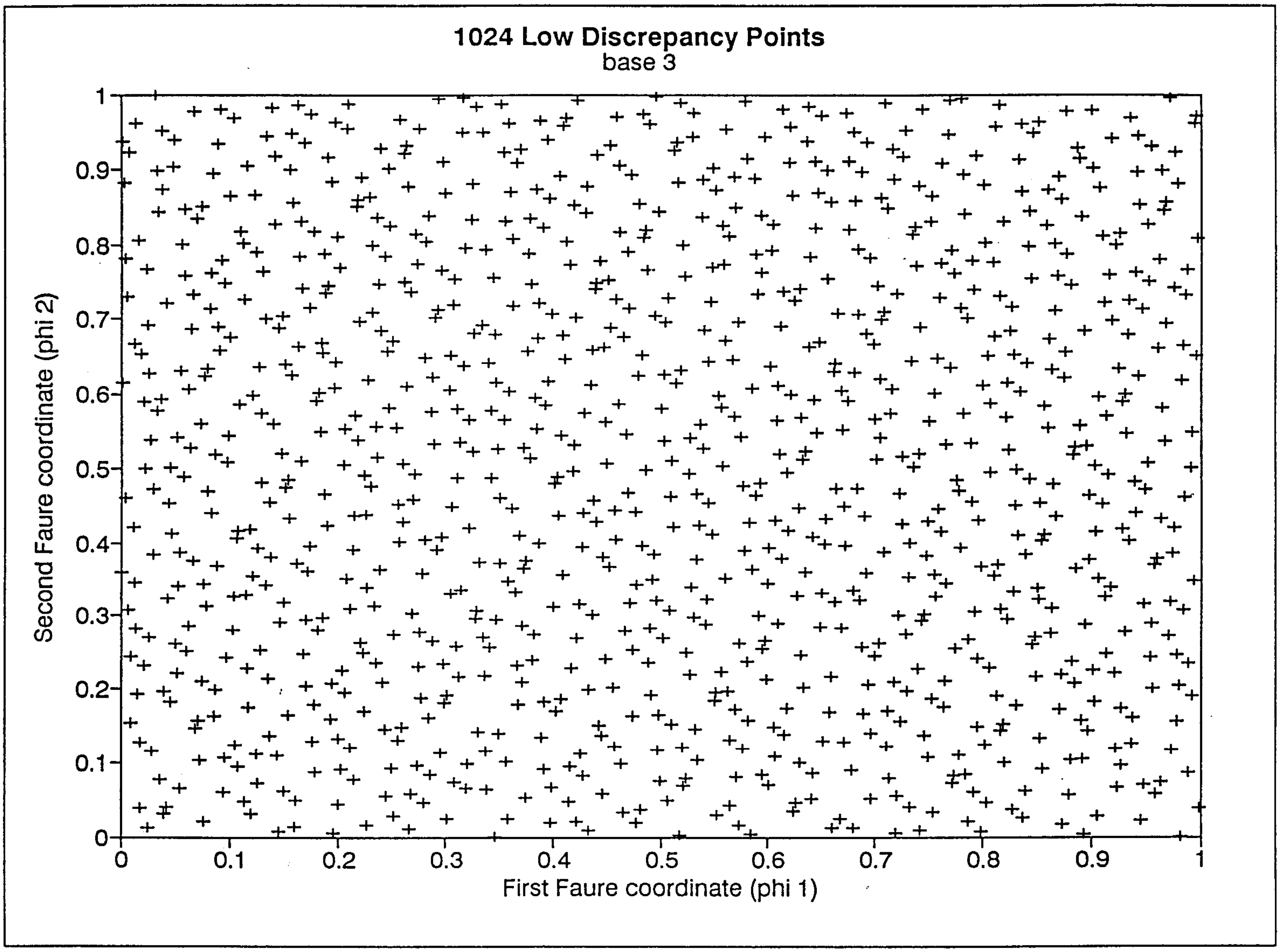

1024 Low Discrepancy Points
base 3
Second Faure coordinate (phi 2)
First Faure coordinate (phi 1)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Fig 3

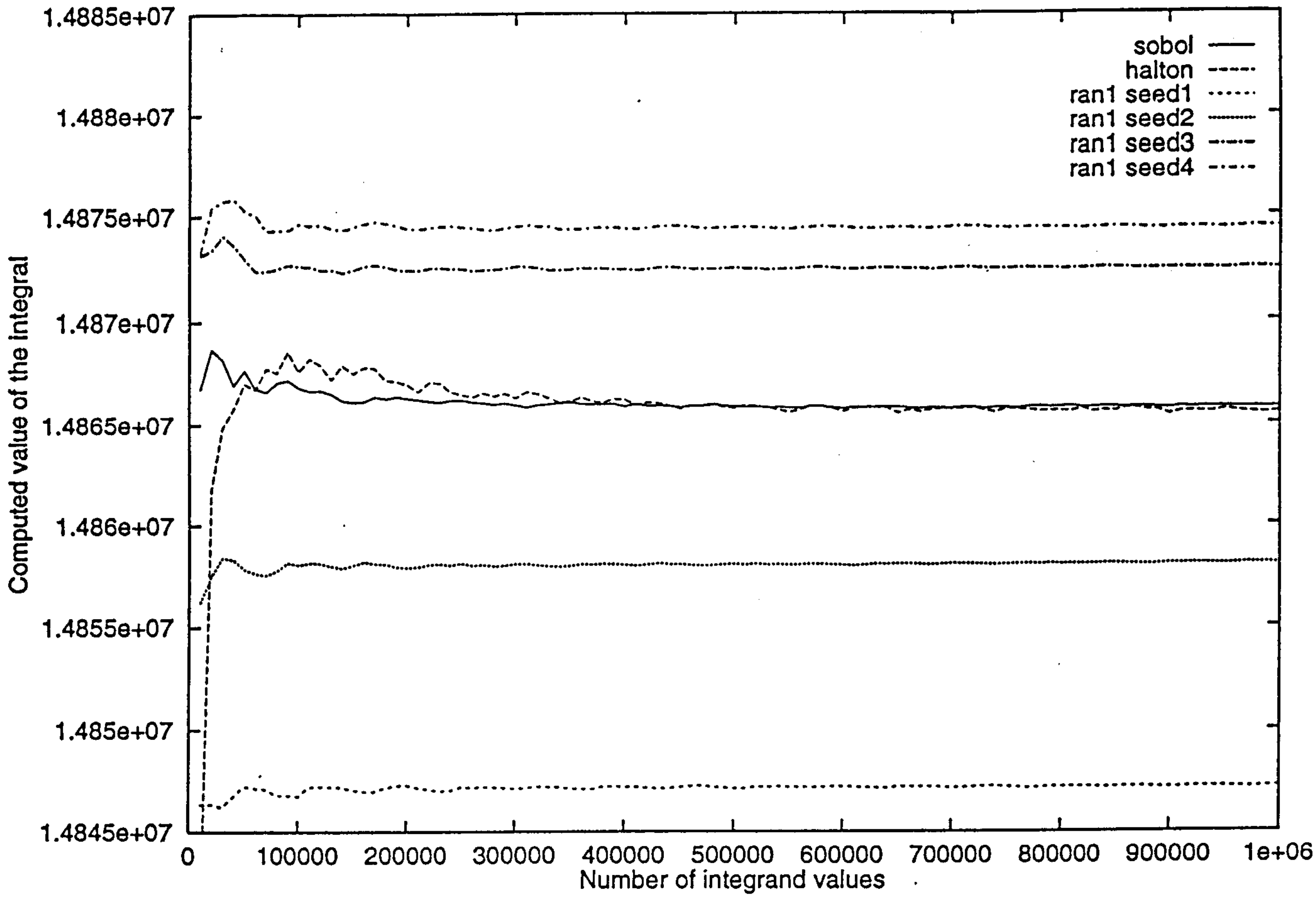


Figure 4: Sobol and Halton runs for tranche A and four Monte Carlo runs using ran1

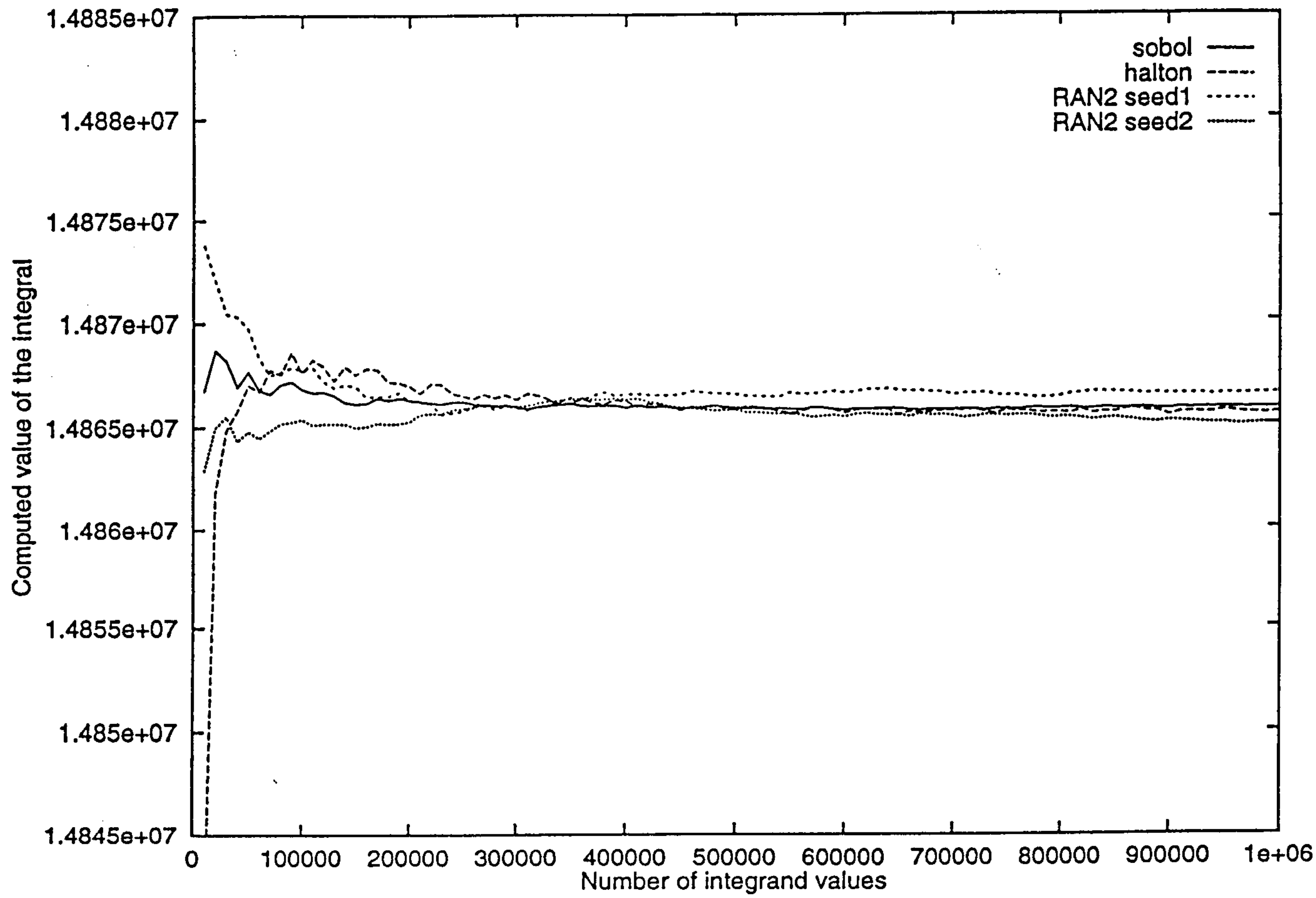


Figure 5: Sobol and Halton runs for tranche A and two antithetic variables runs using RAN2

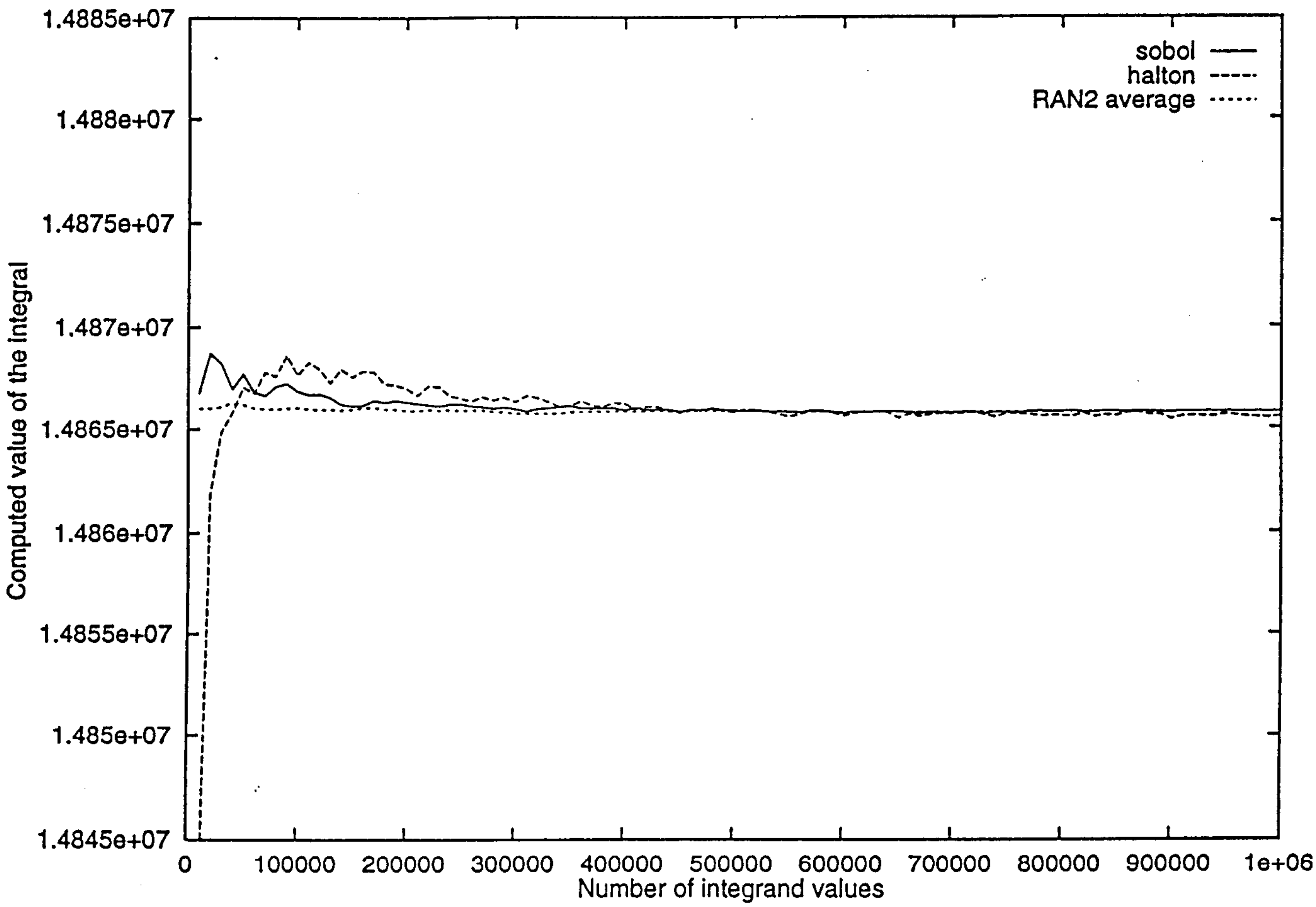


Figure 6: Sobol and Halton runs for tranche A and an average of twenty antithetic variables runs using RAN2

(15)