# Low-discrepancy sequences for volume properties in solid modelling

T.J.G. Davies and R.R. Martin

**Abstract**

This paper investigates the use of low-discrepancy sequences for computing volume integrals in geometric modelling. An introduction to low-discrepancy point sequences is presented which explains how they can be used to replace random points in Monte Carlo methods. The relative advantages of using low-discrepancy methods compared to random point sequences are discussed theoretically, and then practical results are given for a series of test objects which clearly demonstrate the superiority of the low-discrepancy method.

## Low-Discrepancy Sequences

*Monte Carlo* methods of integration are used widely for calculating volume integrals in solid modelling. The Monte Carlo method uses *randomly* generated points inside a box enclosing an object of interest to calculate volume integrals. For example, the volume of the object can be estimated as the ratio of number of points that are contained within the object to the total number of points generated, multiplied by the volume of the box. Naturally, such a method is subject to errors because of the random nature of the sampling, and in particular we cannot guarantee that all parts of space will be sampled equally well. Quasi-Monte Carlo methods [4] use *pseudo-random* sequences of numbers, called *low-discrepancy* sequences, for computing multi-dimensional integrals, where here pseudo-random indicates that the sampling is to be done in a rather more structured manner.

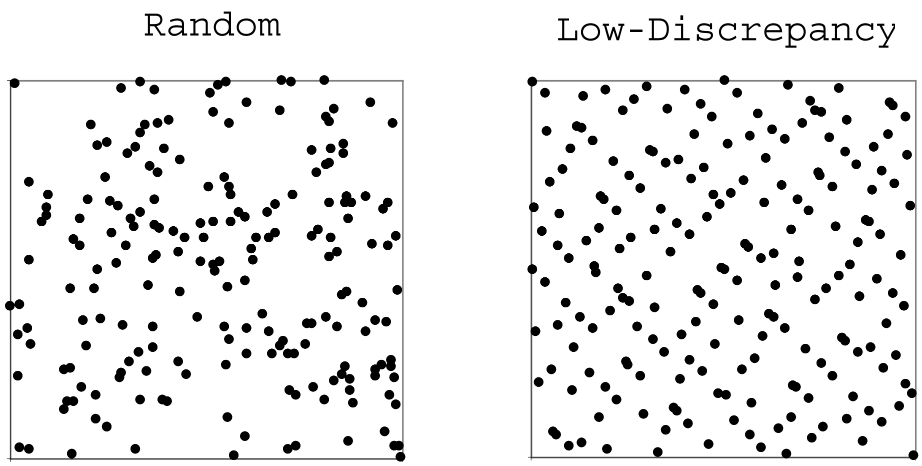


Figure 1. Comparison of points generated randomly and using low-discrepancy sequences.

The key idea is the one of *discrepancy*, which is a measure of how uniformly the points sample the space. (A simple introduction to low-discrepancy methods, in the context of applications to financial problems, can be found in [2].) Two sets of 200 points in two dimensions are shown in Figure 1. Those on the left were generated using a random number generator, while those on the right were generated using a low-discrepancy sequence. Clearly, there are some large 'holes' in the random sampling, while the holes in the low-discrepancy sampling are less pronounced. Note also, however, that the low-discrepancy samples do not form a regular grid. Such a grid can give large errors when used for volume integral computation, in cases where the object is just a little larger or a little smaller than the grid spacing, for example. This problem does not arise for the low-discrepancy point sequences.

To understand discrepancy, let us first consider one dimension. Take the interval [0,1] and let $E$ be any subset of this interval, defined by the characteristic function

$$f_E(x) = \begin{cases} 0 & \text{if } x \notin E \\ 1 & \text{if } x \in E. \end{cases} \tag{1}$$

Now define

$$A(E, N) = \sum_{n=1}^{N} f_E(x_n), \tag{2}$$

where $x_1$,$x_2$,...,$x_N$ are $N$ numbers in [0,1]. Thus $A(E,N)$ is the number of the $x_N$ which are in $E$. The *discrepancy* $D_N$ of the $N$ numbers $x_1$,$x_2$,...,$x_N$ is

$$D_N = \sup_J \left| \ \tfrac{A(J,N)}{N} - |J| \ \right|, \tag{3}$$

where $J$ now runs through all subintervals of [0,1], and $|J|$ is the length of $J$. Thus $D_N$ is the biggest possible error when estimating the length of any interval $J$ by sampling using the given set of $x_N$ and using $A(J,N)/N$ as the estimate of its length.

More generally, if $f$ is any function with bounded variation $V(f)$ on $I$ it can be shown that

$$\left| \ \tfrac{1}{N} \textstyle\sum_{n=1}^{N} f(x_n) - \int_0^1 f(t)dt \ \right| \leq V(f)D_N^*, \tag{4}$$

where $D_N^*$ is a slightly different definition of discrepancy based only on those intervals whose left hand ends start at 0.

Similar definitions and results apply in $m$ dimensions where the intervals are replaced by rectangular parallelepipeds. It can be shown that the two different definitions of discrepancy are of the same order for fixed $m$:

$$D_N^* \leq D_N \leq 2^m D_N^*. \tag{5}$$

Making use of this concept of discrepancy relies on the fact that there are known algorithms (see later) for generating sequences of points in $m$ dimensions which have low-discrepancy. In particular, the discrepancies of such sequences are smaller than the expected discrepancies for a random set of points.

In light of these remarks, we fully expect the use of such sequences to have an advantage in calculating volume integrals in solid modelling (even where the volumes to be integrated over are not smooth mathematical objects, but are perhaps mechanical components). These expectations are borne out in experimental tests which we present in the rest of this paper. Thus, the main purpose of this paper is to draw the attention of the geometric modelling community to the advantages of using low-discrepancy sequences for volume integration.

## Theoretical Advantage

Following an observation made by Woodwark [7] we may note the following, in the case of randomly generated points. If $N$ trials are made of a random event whose probability of success is $p$, then the expected number of successes is $Np$, and the standard deviation in that number is $\sqrt{Np(1-p)}$. Thus, when using points generated randomly in a Monte Carlo method to estimate volumes in this way, we would expect a relative error in the volume of a size comparable to

$$\frac{\sqrt{Np(1-p)}}{Np} = \sqrt{\frac{(1-p)}{Np}}, \tag{6}$$

which is

$$O(N^{-\frac{1}{2}}) \tag{7}$$

in the number of sample points.

However, when using low-discrepancy sequences, it is possible in $m$ dimensions to generate sequences of points whose discrepancy is $O(\log^m N)$, so giving an expected relative error in volume (see Equation 4) of

$$O(N^{-1}\log^m N). \tag{8}$$

Clearly, asymptotically, this means that the expected error for low-discrepancy sequences is lower than that for random points.

In practice, there are two additional considerations. Firstly, for small $N$, what are the relative slopes of these functions? As can be seen in Figure 2 for the case of three dimensions (the main case of interest for geometric modelling), while $O(N^{-\frac{1}{2}})$ may decrease slightly quicker with $N$ for $N$ between 100 and 1000 points, clearly by the time $N$ is above 10000 points, $O(N^{-1}\log^3 N)$ is decreasing more rapidly (Figure 2 uses logs to base 10).

Secondly, there is the question of the constants of proportionality in these different functions. (This corresponds to a relative vertical shift in the two curves in Figure 2—whereupon we get the question of at what value of $N$ the $O(N^{-1}\log^3 N)$ graph overtakes the $O(N^{-\frac{1}{2}})$ graph.) This depends on the particular low-discrepancy sequence used, and for example it is well known that Sobol's point generation method [6] has a worse constant of proportionality than

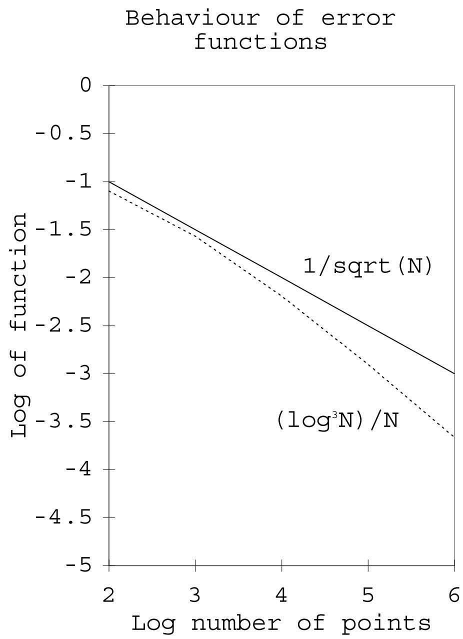


Figure 2. Comparison of error for Monte Carlo and low-discrepancy methods.

Niederreiter's [5]. We offer no further theoretical analysis on this point, but as the results show later, the constants of proportionality are such that low-discrepancy sequences have an advantage even for small $N$.

# Point sequences and test data used

We performed experiments to compute the volumes of objects using random points and two different low-discrepancy point sequences in Monte Carlo and quasi-Monte Carlo methods. For random points, the built-in UNIX random number generator was used. Although in principle pseudo-random number generators of this type can exhibit undesirable lattice structures in higher dimensions [4], we did not observe such effects here. The two low-discrepancy sequences used were Sobol's (for theory see [6]) and Niederreiter's (for theory see [5]). In both cases, implementations from *Collected Algorithms of the ACM* were used: for Sobol's method, see [1], and for Niederreiter's method, [3].

Various forms of Niederreiter's method exist. We used the base 2 method, which can be implemented more efficiently.

A small collection of test objects was compiled, comprising three simple shapes, and three more complex mechanical components. Objects 4 and 5 were supplied by J. Corney of Heriot-Watt University; the objects are available on the Web in the NIST Repository: `http://www.parts.nist.gov/`. Object 6 was supplied by A. Safa of Intergraph Italia. These objects are described below, as are the bounding boxes used for the volume calculations (note that these are not always as tight as possible).

- Object 1: Sphere, radius 1.0. Bounding box used: 2 x 2 x 2.
- Object 2: L-shaped block, width 2, height 2, length 3 with a block of width 1, height 1 and length 3 removed from the top right corner. Bounding box used: 4.5 x 6.5 x 4.5.
- Object 3: Block with cylindrical hole, width 2, height 2 length 3, with vertical cylindrical hole of diameter 1.0 through the centre. Bounding box used: 4.5 x 6.5 x 4.5.
- Object 4: HW1: A mechanical object—see Figure 3. Bounding box used: 318 x 148 x 30.
- Object 5: HW2: Another mechanical object—see Figure 4. Bounding box used: 123.709 x 117.919 x 475.
- Object 6: A valve—see Figure 5. Bounding box used: 0.237 x 0.165 x 0.1675.

# Experimental methodology

The volumes of the objects were computed in each case in three distinct ways: using random points, and using low-discrepancy point sequences generated by Sobol's method and Niederreiter's method. Each volume was calculated by generating points lying inside a rectangular box enclosing the object, using point-membership classification to decide if each point was in the box, and then using the formula:

$$V_{obj} = V_{box}(\frac{N_{in}}{N}). \tag{9}$$

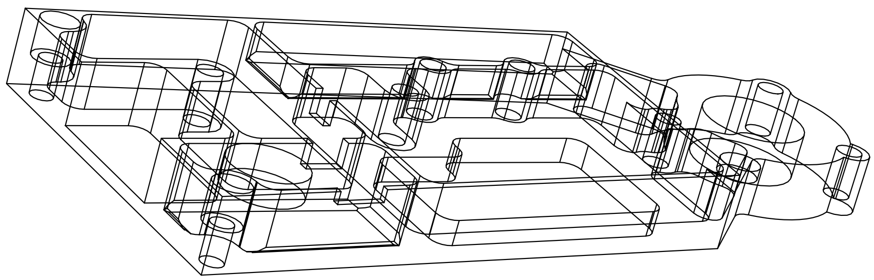

Figure 3. HW1 object.

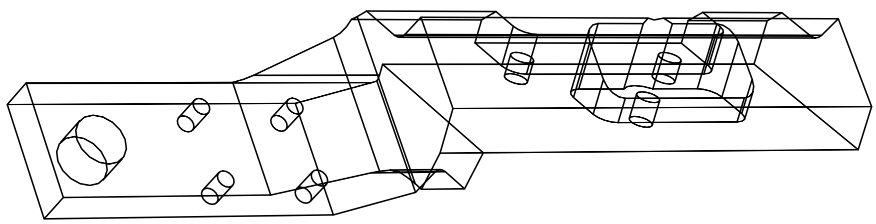

Figure 4. HW2 object.

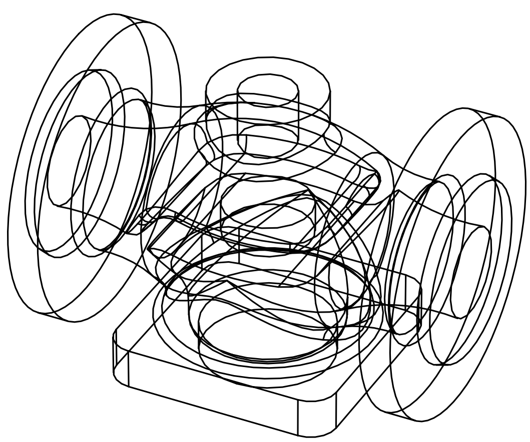

Figure 5. Valve object.

$V_{obj}$ is the estimated volume of the test object, $V_{box}$ is the volume of the box, $N_{in}$ is the number of points found in the object and $N$ is the total number of points generated.

The experiment carried out on each object, for each method, was to compute the volume of the object for an increasing number of points, and in each case to observe the fractional error in the computed volume relative to the true value. For Objects 1–3 the error was calculated at $10^2$, $10^3$, $10^4$, $10^5$ and $10^6$ points. For Objects 4–6 the error was calculated at $10^2$, $10^3$, $10^4$ and $10^5$ points.

For Object 2, the L-shaped block, errors were also calculated every 100 points up to $10^5$ points to investigate the behaviour of the low-discrepancy sequences in more detail.

Values used for the true volume of the object were computed theoretically for Objects 1–3, and found accurately using a commercial solid modeller for Objects 4–6.

# Results

## Timing observations

Using UNIX timing functions, it was found that when using all three methods, the point-classification step was much slower than generating the points, and in practice there was no observable time disadvantage in using any of the three methods to generate an equal number of sample points.

## Sobol's method

Experiments with the Sobol point generator proved disappointing, and gave results which were not much better than those achieved using random points. We thus do not present these results here. On the other hand, the Niederreiter point generator achieved impressive improvements over random point generation, and we give these results in detail below. As mentioned earlier, it is already known that Sobol's sequences do not have such good properties as Niederreiter's, which was borne out by our own experimental observations.

## Errors from random points

Note that each run of a Monte Carlo method with differing random points will give differing results, with differing errors. Using any

| *No. of points* | *Experimental* | *Theoretical* |
|---:|---:|---:|
| 10 | 1.395 | 1.263 |
| 100 | 0.427 | 0.4 |
| 1000 | 0.082 | 0.1263 |
| 10000 | 0.052 | 0.04 |
| 100000 | 0.0125 | 0.0126 |

Table 1. Errors in Monte Carlo method versus number of points.

one run of the Monte Carlo method as an indication of the errors obtained may thus be misleading. Instead, we have a theoretical estimate (see Equation 6) of how big that error is. We thus did a preliminary investigation to see if using the UNIX random number generator did produce relative errors in computed volumes of this size. For various numbers of points, we computed the volume ten times, and found the standard deviation in the volume computed. These are compared in Table 1 to the standard deviations predicted by Equation 6. As can be seen, the errors in practice match well to those predicted theoretically.

Thus, in the following section, we use the standard deviations predicted by theory for random points as 'typical' errors for comparison with errors from the low-discrepancy methods, to avoid statistical fluctuations in the Monte Carlo method affecting the comparison.

In contrast, note that only one result is possible for a given number of sample points using a given low-discrepancy sequence, as it is a well defined sequence of points.

## Results for each test object

Figures 6–11 show the results obtained in our tests. In each graph, the relative error in computed volume is plotted versus the number of point samples used to calculate the volume (the graphs are plotted on a logarithmic scale using logs to base 10). In each case, theoretical errors for a Monte Carlo method based on random points are compared to the actual relative errors obtained using Niederreiter's method.

From the results obtained for each object, it is clear that in general the Niederreiter method gives a distinct advantage over the

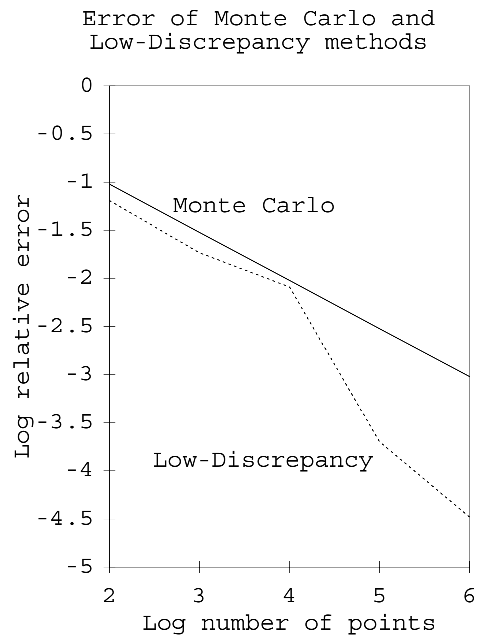


Figure 6. Accuracy versus number of points for the sphere.

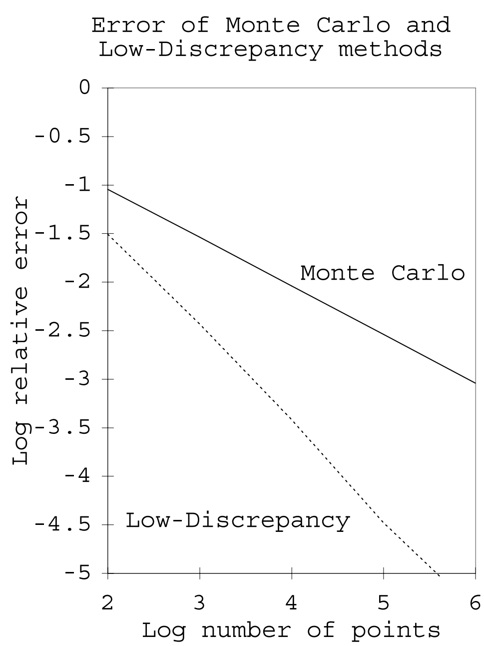


Figure 7. Accuracy versus number of points for the L-shaped block.

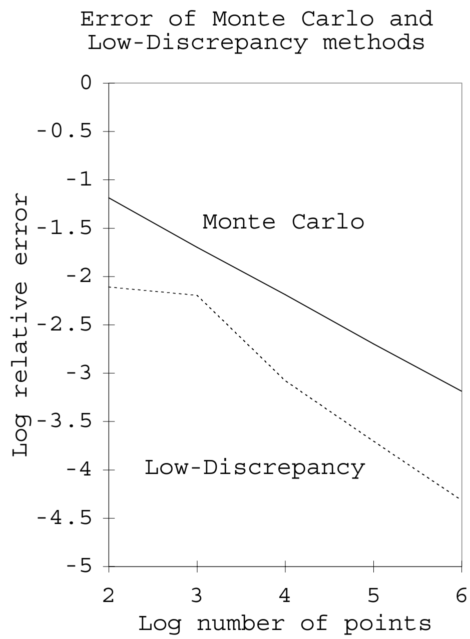


Figure 8. Accuracy versus number of points for the block with a cylindrical hole.

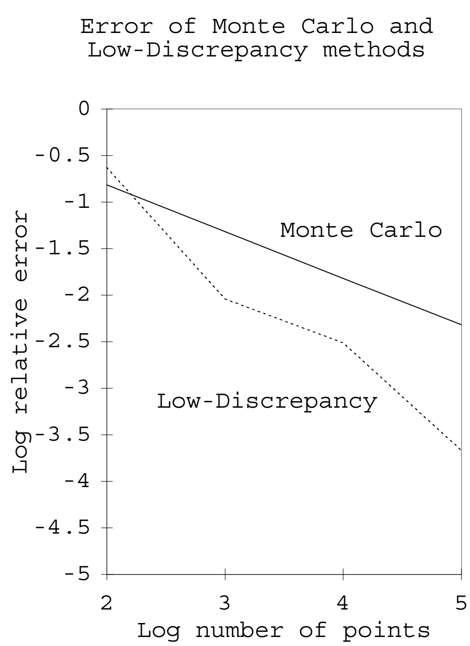


Figure 9. Accuracy versus number of points for the HW1 Object.

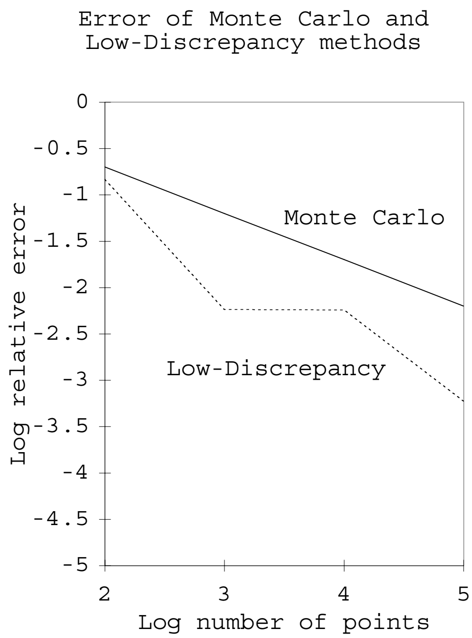


Figure 10. Accuracy versus number of points for the HW2 Object.

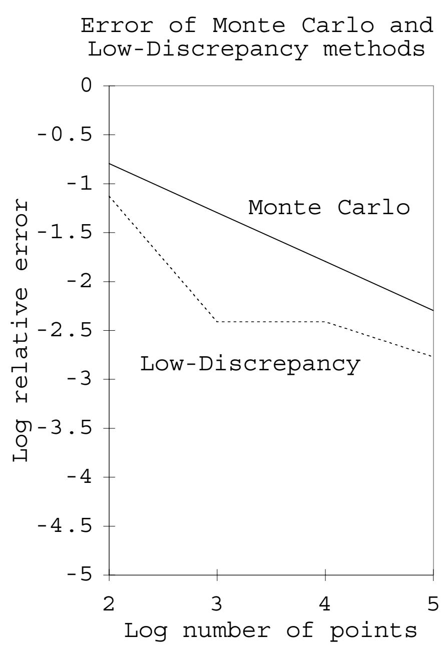


Figure 11. Accuracy versus number of points for the Valve object.

| *Object* | *Low-discrepancy slope* |
|---:|---:|
| Sphere | 0.85 |
| L-shaped block | 0.98 |
| Block with hole | 0.59 |
| HW1 | 0.96 |
| HW2 | 0.72 |
| Valve | 0.49 |

Table 2. Gradients of low-discrepancy graphs.

Monte Carlo method, in that many fewer points are needed to achieve a given accuracy, even for quite small numbers of points (more than a few hundred). We can also see that in most cases, at least, the graph for Niederreiter's points has a steeper slope than that for the Monte Carlo method on average: in each case we have found the best-fit straight line through this graph, and presented its slope in Table 2. The corresponding slope of the Monte Carlo graph is 0.5 in each case. This means that, as more points are chosen, the relative advantage of the low-discrepancy method increases relative to the Monte Carlo method. (For the Valve object, the gradient is in fact less than that of the Monte Carlo graph. Nevertheless, the low-discrepancy method is more accurate for this object for any given number of points in the experimental range than the Monte Carlo method. These graphs have been drawn from a small number of samples, which probably explains the low slope found in this particular case.)

## Detailed results

Figure 12 is the more detailed graph for the L-shaped block showing errors every 100 points. A best-fit line was drawn; the gradient of this line is 0.79. While there are considerable fluctuations in the errors as the number of sample points varies, nowhere is the actual error more than 10 times greater than the trend, shown by the best-fit line, and often, the method does much better than the trend. Because the error does not vary smoothly with the number of sample points, it would in general be difficult to give *guarantees* of the error obtained in computing volume integrals using low-discrepancy sequences (note that Equation 4 is only directly relevant for rect-

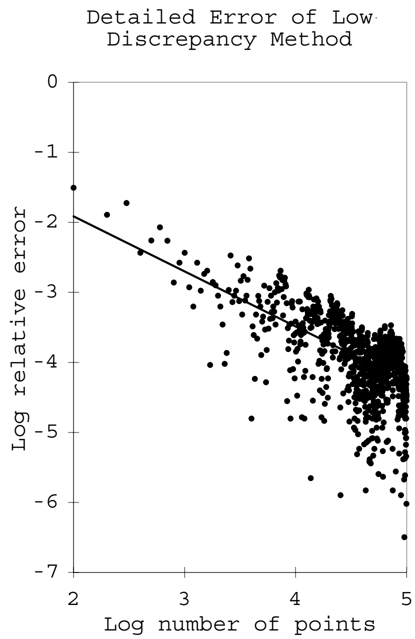

Figure 12. Detailed graph of errors for the low-discrepancy method for the L-shaped block.

angular parallelepipeds), although clearly reasonable estimates of likely errors can be given.

## Relative efficiency

Comparing the Monte Carlo method with Niederreiter's method, the actual results achieved are very impressive. In the case of Object 4 (the HW1 object), more than 20000 points are needed in the Monte Carlo method to achieve an accuracy of 1%, while fewer than 1000 points from the Niederreiter sequence are needed, for example. Table 3 shows for each object the relative efficiency of each method for each test object. In each case, the number of test points approximately needed to achieve 1% accuracy is shown, and the relative advantage of the low-discrepancy method computed. As can be seen from Figure 6, the results for the sphere are somewhat unlucky due to an upturn in the low-discrepancy curve just around 1% error, and the other results are probably more representative.

| *Object* | *Monte Carlo* | *Niederreiter* | *Ratio* |
|---:|---:|---:|---:|
| Sphere | 8110 | 5136 | 1.6 |
| L-shaped block | 8000 | 355 | 22.6 |
| Block with hole | 4782 | 72 | 66.4 |
| HW1 | 24547 | 893 | 27.5 |
| HW2 | 39355 | 662 | 59.4 |
| Valve | 25148 | 437 | 57.5 |

Table 3. Sample points approximately needed for 1% accuracy of volume for each object.

# Conclusions

It is clear from the results and graphs that using Niederreiter low-discrepancy point sequences in a quasi-Monte Carlo method are much better than using random points for computing volumes for all the test objects, even for a small number of points. Furthermore such low-discrepancy point sequences can be generated at negligible extra cost compared to random point sequences of the same number of points, when taking the overall computational time into account.

Of course, in a real geometric modeller, Monte Carlo methods are likely to be combined with a recursive subdivision scheme [8], so that detailed evaluation is only done near the boundaries of objects and not at places well inside or well outside the object. Nevertheless, the results presented here give good reason to believe that low-discrepancy sequences will give just as great a benefit when used in a local context as when used globally—the simplicity of Objects 1–3 perhaps being more representative of the local geometry of objects.

We fully expect low-discrepancy sequences to be adopted in the future for computing volume integrals in solid modelling.

# Acknowledgements

We would like to thank the Nuffield Foundation for funding T. Davies in this work with a bursary under program NUF-URB97.

We would also like to thank J. Corney of Heriot Watt University for supplying Objects 4 and 5 for this experiment, and A. Safa of Intergraph Italia for supplying Object 6.

# References

[1] P. Bratley, B.L. Fox, ALGORITHM 659 Implementing Sobol's quasi-random sequence generator, *ACM Transactions on Mathematical Software* **14**,1 (88–100), 1988.

[2] B. Cipra, In Math We Trust, *What's Happening in the Mathematical Sciences 1995–1996* (100–111), American Mathematical Society 1996.

[3] B.L. Fox, H. Niederreiter, ALGORITHM 738 Programs to generate Niederreiter's low-discrepancy sequences, *ACM Transactions on Mathematical Software* **20**, 4 (494–495), 1994.

[4] H. Niederreiter, Quasi-Monte Carlo methods and pseudo-random numbers, *Bulletin of the American Mathematical Society* **84**, (957–1041), 1978.

[5] H. Niederreiter, Low-discrepancy and low-dispersion sequences, *Journal of Number Theory* **30**, (51–70), 1988.

[6] I.M. Sobol, On the distribution of points in a cube and the approximate evaluation of integrals, *USSR Computers and Mathematical Physics* **7**, (51–70), 1988.

[7] J.R. Woodwark, Exercise in *Starting Work on Solid Models*, Geometric Modelling Society Course, Keble College, Oxford, January 1992.

[8] J.R. Woodwark, K.M. Quinlan, Reducing the effect of complexity on volume model evaluation, *Computer Aided Design* **14**,2 (89–95), 1982.