I. A whole bouquet of Sobol sequences

Sobol sequences are one of the oldest types of low discrepancy sequences (LSDs) [SOB67]. Their production is easy and fast, and they play an appreciable role in applications. In our recent papers [SCH03, SCH04] we mentioned we had used them to some extent to compute true discrepancies and run test integrations, but in those papers we were mainly concerned with Halton and Niederreiter sequences, and stated only that the behavior of the Sobol sequence is not too different from Niederreiter ones. However, a closer look shows that “the” Sobol sequence does not exist. Similar to Niederreiter sequences, which can be produced from different bases, Sobol sequences (even if we define them as always having base 2) can be produced from different sets of the so called direction numbers (DNs). For a given dimension $s$ the number of possible such sets is finite, but it increases more than exponentially with dimension $s$, and is already 2048 for $s = 6$. Different restrictions on the DNs have been introduced to obtain Sobol points with good uniformity properties, beginning with Sobol’s “property A” [SOB76]. The renewed interest in Sobol sequences after the introduction of Quasi-Monte Carlo integration into finance has led to several new proposals to produce “good” DNs, and tables have been published up to dimension 5000 [BRO03, JAK02, JOK03, SIB03, SIL03, WIN03]. Programs into which these sets can be incorporated are available in the public [BRF88, HOH03]. The quality of these proposals has sometimes just been postulated from their construction, and sometimes argued from computed $L_2$-discrepancies $T^*$.

A better quality parameter would, of course, be the $L_4$-discrepancy $D^*$ which, unfortunately, cannot be computed for values of $s$ and $N$ of practical interest. Nevertheless, we have calculated some $D^*$ for $s = 6$ and report them below. The somewhat easier computation of $L_2$-discrepancies $T^*$ (and $T$) [MOC94, JHK97] has been extended by us to $N > 10^6$ for $s = 6, 12,$ and 24. From our results it is, however, doubtful whether $T^*$ (or $T$) is a good measure for the practical quality of a sequence.

The main purpose of this supplement is, therefore, to present results of integration tests with several well known test functions. These functions and the details of the integration procedure are described in our paper [SCH04]. But, different from the paper, we use here other steps for $N$, viz. steps of $2^k$ and steps of $10^k/4$, $k = 1, 2, ...$. By comparing these we wanted to see whether the lattice structure of sequences terminating at $N = 2^k$ makes any difference to more irregular step sizes. (For our test functions and parameters the answer is no.) We remind the reader that the effective dimensions of test function 1 (TF1) is only about $s/4$, that of TF2 and TF3 is practically equal to $s$, and that of TF4 is very small and almost independent of $s$. In the figures we denote the different DN sets as follows: $H = \text{Ref. [HOH03]}$ stemming from Ref. [JOK03], and devised to improve the original Sobol set. This set was used in Ref. [BRF88], and is identical to set $H$ for dimensions $< 21$. Also, $K = \text{Ref. [BRO03]}$, $S = \text{Ref. [SIL03]}$, $W = \text{Ref. [WIN03]}$. For $s = 3$ only set $B$ differs from the others, while $W$ is identical to $H$ for $s < 7$. We also used several “wrong” sets, e.g. the direction numbers of Ref. [SIL03], which should be used with irreducible polynomials starting with degree 3, used here with the irreducible polynomials of the other sets starting with degree 1. These are valid Sobol sequences, however not those intended by the authors. Finally we run a few integrations with the “worst possible” direction.
numbers, i.e. all set to 1, which we called set E.

II. True $D^*$ discrepancies of some Sobol sequences

We did this calculation only for $s = 6$, because for $N = 3$ the considered DN sets are equal, and for $s = 12$ our computing resources would limit us to about $N < 40$, which we considered uninteresting in our context. Even for $s = 6$ the practical limit for $N$ is about 300 for us. As for other sequences [SCH04] after a short initial region, which assures that $D^*$ is $< 1.0$, and goes to about $N = 8$, $D^*$ decreases with a power law, whose exponent is approximately -0.6 here. Fig. 1 shows, how near the $D^*(N)$ of the sequences are to each other. For $N \geq 10$ the average ratio of any pair of data sets differs from 1.0 by less than 10%, whereas due to fluctuations the maximum observed ratio of any pair at any $N$ is 1.5. The original Sobol sequence (code “H”) has on the average a slightly higher $D^*$ than both others, while the averages of $S$ and $K$ are ordered $K < S$ for $10 < N < 100$, but $K > S$ for $100 < N < 350$. Note that $D^*$ for set E is comparable to the others, and, e.g., better than for set H for $N > 100$. Unfortunately we are unable to extend this plot to large enough $N$ to see whether $K$ remains permanently smaller than the others.

III. $L_2$-discrepancies $T^*$ of some Sobol sequences

In contrast to $L_4$-discrepancies $D^*$, the $L_2$-discrepancies $T$ (for arbitrary axis-parallel hyper-rectangles) and $T^*$ (for axis-parallel hyper-rectangles containing the origin as one corner) can be calculated to much higher $N$. We did that for $s = 6$, 12, and 24 up to more than $N = 10^6$. Here, we concentrate on $T^*$ which is always orders of magnitude larger than $T$, and for which the inequality $T^* < D^*$ holds. Moreover, it has been used in the literature to grade the quality of different sequences (e.g. Refs. JHK97, JAK02).

Fig. 2 shows the typical situation for small dimension (here $s=6$): With the exception of sequence $S$ and the Halton sequence (which we show for comparison), the values of $T^*$ stay below the expectation for a random sequence (straight line). But for $N \geq 100$ they are not far from random behavior, and only then turn into a steeper decay. The log-log slope is slowly increasing from about -0.6 to about -0.85 for $H$ and $K$ (somewhat less for $S$). Such increase is expected since $T^* < D^*$, and $D^*$ is asymptotically $O((\log N)^3/N)$ with a slope approaching -1. However, the crossover of the asymptotic $D^*$ and the random expectation of $T^*$ is very far out, e.g. for $s = 6$ at $N = 10^{23}$, and for $s = 12$ already at $10^{60}$. The differences between the various Sobol sequences are very small except for $N \leq 30$ where $S$ (and Halton) are above random. Especially, for $N \geq 10^3$ the points of the sequences H=W, and $K$ (and also Halton) fall practically upon each other.

The situation is similar for $s = 12$ (not shown), except that the kink, after which the Sobol sequences are better than random, is only at $N \approx 10 000$. This behavior is in concord with that shown for $s = 10$ in Ref. [BRO03].

Finally, for $s = 24$ (Fig. 3) the kink (which must occur somewhere) is not reached in our calculations, and one may only speculate how far it occurs (certainly much farther than $N = 10^6$). Sequences $H$, $K$, and $W$ are somewhat below random for small $N$ ($\approx 1000$), and $S$ and Halton are much above it! But for $N \geq 1000$ all Sobol $T^*$ are practically equal to the random case. Further remarks on the behavior of $T^*$ can be found below in Appendix A.

But we emphasize already here that the connection of $T^*$ with the quality of Quasi-Monte Carlo integration is very, very loose, i.e. $T^*$ is not a reliable measure to assess the quality of different low discrepancy sequences for Quasi-Monte Carlo integration!

IV. Test integrations with various Sobol sequences

In our paper [SCH04] we have already mentioned that we had done a few test integra-
tions also with Sobol sequences. This was done with the type of sequence which we denote “H” here. In this supplement we extend these calculations to the four different Sobol sequences mentioned in the introduction. The calculations were restricted to N <10⁹, in order to avoid rounding errors and not to be forced to use quadruple precision as in the main paper. Typical results are shown in the next few figures.

Our main conclusion is that the quality of integration with these four sets of different direction numbers is essentially the same. The fluctuations of the average error ε of ten integrations, whether correctly scrambled (Fig. 4) or only repeated with consecutive pieces of a fixed sequence (Fig. 6), are comparable to or even larger than the differences between sequences from DN sets (at least for those tested by us, which were selected by their authors to be particularly good). The rms errors σ are barely different (Fig. 5). This situation is very much the same for s = 12 (not shown) and s = 24 (Figs. 7 and 8) as for s = 6. In all cases, after eliminating some outliers for very small N, say N <30, the various fits to ε and to σ are nearly the same. I.e. the differences for ε for some N are within one order of magnitude, which is less than the fluctuations as a function of N, and those for σ within a factor of 3, again less than the fluctuations. Since the fluctuation maxima and minima for different DN sets occur at different values of N, one can, however, recommend the use of different DN sets as a consistency check, especially if one intends to do the integration with a fixed number of trials N and to omit other convergence checks.

Finally a few integrations with direction number set E (all direction numbers equal to 1) again did not show the essentially different results. (One could have anticipated that from remarks in the literature.) Because of the large fluctuations only fits to the average absolute errors ε or the rms errors σ produce meaningful results. Evaluated at N = 1000 or 10 000 set E is indeed worse than the average of the others, but typically only by 50%. Similar to the Halton case, while pictures of pair-wise projections show great gaps for a single point sequence of moderate length N, these gaps are apparently averaged out, if the integration is repeated either with pieces of the same sequence or with scrambled sequences.

V. Conclusions

What we show in this supplement is that quasi-Monte Carlo integration with Sobol sequences produced from different direction number sets give more or less equivalent results. This is in concord with the results of our main paper, where we compared several sequences of the Halton, Niederreiter and Niederreiter-Xing type, and also found not the big differences which one could anticipate from the literature. What is wrong? Let us examine on what the expectations are based.

1) Theoretical formulas about worst case errors, as discussed in the main paper. We showed there, I hope convincingly, that practical computations are neither in the asymptotic regime nor do they have worst case errors.

2) Computations of the L₂-discrepancy T*, the only one which can feasibly been computed. We showed above that T* is of doubtful value for the assessment of integration errors.

3) Computational tests of more or less practical problems, which show many conflicting results (as e.g. discussed for problems in financial mathematics in Ref. SIB03). In view of the large fluctuations of quasi-Monte Carlo integration as a function of the number of trials N - see Figs. 4,6,7 of this supplement and Figs. 2,4,6 of the main paper, which are already fluctuations of the average of 10 integrations - we suspect that many tests in the literature show just the result of a positive or negative fluctuation. Unfortunately in many tests only fits but not the
fluctuating points are shown, and often the full set of parameters (number of repetitions, range of the fits) are not given, which makes the assessment of such papers difficult.

On the other hand we will not conceal here the limitations of our own work: (a) We used a limited number of test functions, and only rather simple ones. (Note, however, that TF4 is not a simple product function, and that we know the effective dimension of each one.) (b) Our largest dimension was 24. In financial mathematics 240 might be a more typical number. However, it has also been shown at least in one case [OW03], Keister’s integral, that the effective dimension of these problems may also be very small. (c) We concentrated on large N, i.e. $10^4$ to $10^9$, which for not too exotic test functions give integration errors much smaller than 1%. If one is interested in the error range of 1% or even less, one should again examine the range $N < 1000$, and reduce the fluctuations by an increased number of repetitions, say 1000 or more. Our conclusion in the main paper about the uselessness of the Hlawka-Koksma formula for error anticipation holds, however, also here.

Appendix A. Some strange behavior of $T^*$

When reading Ref. [JHK97] we hit upon some strange results for the $L_2$-discrepancy $T^*$ which did not conform with common (or at least the author’s) expectations for low-discrepancy sequences. The $T^*$-discrepancy has been used in that paper and elsewhere to discuss the efficiency of Quasi-Monte Carlo integration by comparing $T^*(N)$ with the expectation value for pseudo-random sequences, $T^*_{\text{ran}}(N)$, which is known to be

$$T^* = \sqrt{\frac{\langle y^2 \rangle - \langle y \rangle^2}{N}}. \quad (1)$$

Two of the results (which were limited to $N \leq 150,000$ by the computers of 1997) looked so strange to us that we re-computed them: (1) the behavior of $T^*$ of Halton sequences for $s > 18$ (Fig. 15b of Ref. [JHK97]), where, in the covered range of $N$, $T^*$ is much larger than the expected $T^*$ of a random sequence, and (2) a hump of $T^*(N)$ of the Sobol sequence for $s = 11$ at $N$ about 65,000 (Fig. 14 of Ref. [JHK97]). The faster computers of today allowed us to extend $N$ to about $2 \times 10^6$. Note, however, that in contrast to test integrations the computing time for $T^*$ is quadratic in $N$, still limiting the range for which $T^*$ can be computed.

From Figs. 2 and 3 we saw already that Halton (and some Sobol) sequences have a $T^*$ larger than the random expectation for small $N$, where “how small is small” depends on the dimension. Figs. 9 and 10 demonstrate this again for the Halton and Silva’s Sobol sequence. The obviously forgotten and, therefore, deceiving feature is that the expectation value of $T^*$ for pseudo-random sequences (eq. 1) starts near $2^{-s/2}$, while $T^*$ and $D^*$ for all low discrepancy sequences start near 1. On the other hand, theory tells us that for every low discrepancy sequence $T^*(N)/T^*_{\text{ran}}(N)$ must vanish in the limit $N \to \infty$. I.e. for large enough $N$ all the points in Figs. 9 and 10 will lie below zero! Unfortunately we do not know where that happens, and to compute it even for $s = 24$ is probably out of reach today.

However, it must be stressed that integration errors show a behavior different from that of $T^*$. As Fig. 11 shows the errors for integration with sequences S and Halton for $N = 1000$ are nearly the same while Fig. 3 shows a logarithmic difference of more than 1.0 for the values of $T^*$. This (and other observations not shown here) means that $T^*$ is a somewhat doubtful measure for the quality assessment of different low discrepancy sequences for Quasi-Monte Carlo integration.

With respect to problem (2) found in Ref. [JHK97] our new calculations with different types of Sobol sequences show that the strange hump in $T^*(N)$ for “the” Sobol sequence at $N \approx 65000$ for $s = 11$ (and less so for $s = 10$ and 12) is apparently only a fluctuation, which is
specific for the DN set used. As Fig. 12 shows, Sobol sequences with direction number sets H and K (see Section I) show this effect, but those with sets S and W do not. One sees, moreover, how large the fluctuations of $T^*$ are!
References

BRF88  P. Bratley and B. L. Fox, Implementing Sobol’s Quasirandom sequence generator, Algorithm 659, ACM TOMS 14 (1988) 88-100

BRO03  http://www.broda.co.uk/software.htm


JAK02  P. Jäckel, Monte Carlo Methods in Finance, Weihem, Wiley 2002


Figures - All Plots are log-log plots

- Fig. 1

Star discrepancies $D^*$ of different Sobol sequences for dimension 6. Color codes (see text): green: E, red: H (identical to W), magenta: K, blue: S. Fits to the points N>6 are shown dashed.

- Fig. 2

$L_2$-Discrepancies $T^*$ of different Sobol sequences and the Halton sequence for dimension 6. Color codes (see text): red: H (=W), magenta: K, blue: S, black: Halton. The dots for H are thicker to see them below those for K. The pseudo-random expectation for $T^*$ is shown as a green straight line.
**Fig. 3**

$L_2$-Discrepancies $T^*$ of the Halton and different Sobol sequences for dimension 24. Color codes (see text): red: H, magenta: K, blue: S, cyan: W, black: Halton. The pseudo-random expectation is shown as green straight line. Note that all low-discrepancy sequences must eventually decrease systematically below the random line like those in Fig. 2, since $T^* \leq D^*$ and $D^* = O((\ln N)/N)$.

**Fig. 4**

Average error of ten Hickernell-Owen scrambled quasi-Monte Carlo integrations of test function TF3 with four different Sobol sequences in dimension 6. The integrals are evaluated at steps $10^{13}$. Color codes (see text): red: H, magenta: K, blue: S, cyan: W. Fits are shown dashed.
Fig. 5

RMS error (standard deviation) of ten Hickernell-Owen scrambled quasi-Monte Carlo integrations of test function TF3 with four different Sobol sequences in dimension 6. The integrals are evaluated at steps $10^k$. Color codes (see text): red: H, magenta: K, blue: S, cyan: W. Fits are shown dashed.

Fig. 6

Average error of ten repeated quasi-Monte Carlo integrations of test function TF3 with four different Sobol sequences in dimension 6. The integrals are evaluated at steps $10^k$. Color codes (see text): red: H, magenta: K, blue: S, cyan: W. For $s = 6$ the direction number set W is the same as H, therefore without random scrambling the red curve is overlaid by the cyan one. Note that the fluctuations are slightly larger than those for Owen-Hickernell scrambling (cf. Fig. 4). Fits are shown dashed.
Fig. 7

Average error of ten Hickernell-Owen scrambled quasi-Monte Carlo integrations of test function TF1 with four different Sobol sequences in dimension 24. The integrals are evaluated at steps $10^{4k}$. Color codes (see text): red: H, magenta: K, blue: S, cyan: W. Fits are shown dashed.

Fig. 8

RMS error (standard deviation) of ten Hickernell-Owen scrambled quasi-Monte Carlo integrations of test function TF1 with four different Sobol sequences in dimension 24. The integrals are evaluated at steps $10^{4k}$. Color codes (see text): red: H, magenta: K, blue: S, cyan: W. Fits are shown dashed.
**Fig. 9**

$L_2$-discrepancy $T^*$ for the Halton sequence (with primes 2, 3, 5, ...) relative to that expected for a random sequence as function of dimension (logarithmic scale, $s = 12$ through $25$) at some values of $N$: 50k (red), 100k (magenta), 250k (cyan), 500k (grey), 1M (black). See the Appendix.

**Fig. 10**

$L_2$-discrepancy $T^*$ relative to that expected for a random sequence for Sobol sequence $S$ as function of $N$ for dimensions 12 (black), 18 (blue), and 24 (red). The Sobol sequence has a "better" $T^*$ than the random one (negative abscissa in the plot) only above some value of $N$, which increases heavily with dimension.
Average integration error (Owen-Hickenell scrambling) for the integration of TF4 in 12 dimensions with the original Sobol (code "H"), the Halton, and a pseudo-random sequence. Color codes: red: H, magenta: Halton, black: pseudo-random. For comparison an arbitrary straight line with power -1/2 is also plotted. Note that TF4 has a small effective dimension, so quasi-MC is much faster than pseudo-MC even in 12 dimensions.

$L_2$-discrepancy $T^*$ for Sobol sequences H (red), K(magenta), S (blue), W (cyan). The hump near $N \sim 65000$ ($\log_{10} N \sim 4.8$), here for dimension 11 (visible also for $s = 12$ and 13, but not for $s = 9$ or 10) is apparently a somewhat larger fluctuation. It appears only for Sobol sequences with direction number sets H and K, not for S and W. The gray band is the expectation for a random sequence.