Computing 78, 173–193 (2006)
Digital Object Identifier (DOI) 10.1007/s00607-006-0173-3

Computing
Printed in Austria

Distribution of Lattice Points

F. Sezgin, Ankara

Received January 11, 2006; revised June 21, 2006
Published online: September 18, 2006

Abstract

We discuss the lattice structure of congruential random number generators and examine figures of merit. Distribution properties of lattice measures in various dimensions are demonstrated by using large numerical data. Systematic search methods are introduced to diagnose multiplier areas exhibiting *good, bad* and *worst* lattice structures. We present two formulae to express multipliers producing worst and bad laice points. The conventional criterion of normalised lattice rule is also questioned and it is shown that this measure used with a fixed threshold is not suitable for an effective discrimination of lattice structures. Usage of percentiles represents different dimensions in a fair fashion and provides consistency for different figures of merits.

AMS Subject Classifications: 65C10, 65Y05, 68Q22, 11A55.

Keywords: Bad lattice points, good lattice points, lattice rules, linear congruential generators, random number, spectral test.

1. Introduction

Random numbers are essential tools in many applications such as simulation, education, criptography, arts, numerical analysis, computer programming, VLSI testing, recreation and sampling. Because of their efficiency and ease of implementation, linear congruential generators attracted the attention of many researchers and became de facto standards. Random number generators must be subjected to several theoretical and empirical tests to detect certain kinds of weaknesses before their use for serious applications. The most popular theoretical measure for assessing the quality of random number generators is the distribution of t-tuples in t dimensional space. This performance is measured by various figures of merits.

It is well known that the t dimensional vectors of successive numbers in dimension $t \geq 2$ produced by a linear congruential generator have a lattice structure. Several authors have examined this property and discussed various measures for assessing it. In Sect. 2, we summarize and compare basic techniques for assessing the lattice structure. In Sect. 3, by examining some patterns in figure of merits, we address the problem of diagnosing good and bad multipliers. Conventional works on spectral test deal with the problem of finding best lattice structure. But in a recent paper, Entacher et al [8] studied some cases giving rise to bad lattice points. In Sect. 4, we

enhance the classification of bad lattice points and advert some systematic patterns to identify areas comprised of bad lattice points.

We also question the usage of figure of merit based on normalized spectral test with a fixed threshold value. For this purpose, in Sect. 5, distribution properties of normalized spectral test are investigated thoroughly on large amounts of data obtained from various generators. It is shown that the distribution curves of test values have completely distinct patterns in different dimensions. Therefore we show that a percentile-based measure may be more appropriate in order to avoid the deteriorating impact of fixed threshold in smaller dimensions and to provide similar rankings with respect to different methods of assessment.

2. Assessing the Lattice Structure

Several references address the methods of assessment for the lattice structure [1], [6], [11]–[13], [19], [29], and [30].

(a) The squared Euclidean distance ν_t^2,
(b) The distance between adjacent parallel hyperplanes d,
(c) Minimal number of parallel hyperplanes,
(d) The Euclidean distance between points ν_t,
(e) Number of bits of accuracy: $\log_2 \nu_t$,
(f) Standardized figure of merit μ_t,
(g) Beyer quotient q_t: Ratio of the shortest and the longest basis vector lengths,
(h) Normalized figures of merit $S_{1,k}(A, M)$, $S_{2,k}(A, M)$ and $S_{3,k}(A, M)$ with respect to criteria (b), (c) and (d) [11],
(i) Lattice packing constants [11],
(j) Discrepancy.

Since all these methods have their advantages and shortcomings some of them are widely used simultaneously in random number literature.

Given a congruential random number generator with multiplier a, relatively prime to modulus M, for $2 \leq t \leq T$, the spectral test uses integers $\{S_1, \ldots, S_t\} \neq (0, \ldots, 0)$ satisfying the relation

$$\sum_{i=1}^{t} S_i a^{i-1} \equiv 0 \,(\mathrm{mod}\, M). \tag{1}$$

Letting $0 < a < M$, it determines the values of

$$\nu_t^2 = \min\left\{\sum_{i=1}^{t} S_i^2\right\}. \tag{2}$$

The relation (1) may be written as an equation for a certain k satisfying

$$\sum_{i=1}^{t} S_i a^{i-1} = kM. \tag{3}$$

We must stress that contrary to the conventionally accepted definition $0 \leq S_i < M$, the S_i values can not be larger than a. Because for any dimension t, the expression (3) will be equal to zero by choosing only two nonzero coefficients: $S_t = -1$ and $S_{t-1} = a$, giving ${\nu_t}^2 = a^2 + 1$. In fact, this is an upper bound for the squared Euclidean distance.

The spectral test is a very reliable theoretical tool to distinguish bad and good congruential generators. This test is explained in detail by Knuth [19]. Although a plethora of papers address the question of rating various generators, there is no universally adopted criterion to tell whether or not a particular random number generator passes or fails the spectral test. This is partly because the success measure is case-dependent and several generators considered as adequate in common application, fail in specific cases [2], [9], and [10].

In order to make this criterion independent of M, Knuth suggests the standardized figure of merit

$$\mu_t = \frac{\pi^{t/2} \nu_t^t}{\Gamma(t/2+1)M}. \tag{4}$$

There are other criteria adopted by various authors. A very common measure is the normalized figure of merit

$$M_T = \mathrm{Min}\left\{d_t^*/d_t,\ 2 \leq t \leq T\right\}, \tag{5}$$

where, $d_t = 1/\nu_t$ represents the maximum distance between adjacent hyperplanes determined by the points of the lattice in t-dimensional space and d_t^* is the lower bound of this distance.

Although widely used by several authors, the figure of merit q_t, called Beyer Quotient has also received severe criticism. By referring to two works of S. S. Ryshkov on the Minkowski-reduced lattice bases, Leeb [27] reminded the users that Beyer quotient is not uniquely determined for dimensions $t > 6$. It is rather surprising to observe recent uses of q_t for dimensions up to 40. Leeb lists some of these incorrect uses. To mention a few more, we can present the following list with their maximum dimensions T: L'Ecuyer and Tezuka [26] $T = 12$, L'Ecuyer [21] $T = 20$, L'Ecuyer and Couture [24] $T = 30$, L'Ecuyer, Blouin and Couture [23] $T = 20$, and Kao and Tang [14] $T = 8$. Dyadkin and Hamilton [4] and [5] conducted extensive analyses to identify multipliers for 64 and 128-bit multipliers taking $T = 20$. L'Ecuyer and Couture [24] presented a package implementing lattice and spectral tests. Authors support the usage of $M_T = \min_{k \leq t \leq T} S_t$ but also deal with q_t for historical reasons. They also argue that: "One advantage of using Beyer quotients is that they are all normalized (between 0 and 1) and that Q_T is defined for all positive T, in contrast to M_T. One may then compare (and rank) generators of the same size using the figure of merit Q_T for large T."

The proposed thresholds are also subjective and arbitrary. For example, Knuth [19] considers $\nu_t \geq 2^{30/t}$ adequate for a good generator for most purposes but admits that this criterion was chosen partly because 30 is conveniently divisible by 2, 3, 5,

and 6. The same author proposed $\mu_t \geq 0.1$ as a threshold for passing the spectral test for $2 \leq t \leq 6$ and adhered to this rule in the last edition of his book. But this limit is not satisfactory. In application, generators having $\mu_t > 3$ abound even for single-precision floating-point arithmetic (Sezgin [31]).

M_t also has been used with different thresholds. One of the earliest applications belongs to Kurita [20], who by screening 7440 multipliers for $M = 2^{31} - 1$ obtained multipliers having values larger than 60% of the upper limits. Some other thresholds used are 80% [11] and [13], 70% [16], 75% [17], and 84% [18].

Using the density sphere packings formulas given by Conway and Sloane [3], L'Ecuyer [22] listed the upper bound of M_T up to $T = 48$. In this work, L'Ecuyer presented test values for 290 multipliers belonging to various size moduli between $2^8 - 5$ and $2^{128} - 159$. Investigation of these values show that increasing the dimension causes a fall in M_T in 95% of cases. The sizes of test values are distributed as follows:

$$M_8 > M_{16} = M_{32} \text{ in } 44.5\% \text{ of cases,}$$
$$M_8 > M_{16} > M_{32} \text{ in } 37.2\% \text{ of cases,}$$
$$M_8 = M_{16} > M_{32} \text{ in } 13.4\% \text{ of cases,}$$
$$M_8 = M_{16} = M_{32} \text{ in } 4.8\% \text{ of cases.}$$

$M_8 = M_{16} > M_{32}$ is more common in smaller moduli. Increasing modulus causes an increase in cases of $M_8 > M_{16} > M_{32}$. This points out the discrete character of the lattice structure in smaller moduli. For the same reason $M_8 = M_{16} = M_{32}$ cases are also very common for smaller moduli. After this study, the usage of fixed threshold criterion with large T values gained popularity in the literature. Some extensions of M_T are proposed in recent studies: For example, Lemieux and L'Ecuyer [28] proposed $M_{t,k}$ criterion taking into account the projections of the lattice over subspaces of small or successive dimensions. Entacher et al [7] consider M' as a measure of the minimum test value of the multiplier a itself and additionally the subsequence generators with multipliers a_k for a set of different k values. Kao and Tang [15] derived the upper bounds of spectral test for multiple recursive random number generators and conducted several searches.

3. Some Patterns in Spectral Test Figure of Merits

Entacher et al [8] classify the causes of bad lattices under four headings:

(a) If the parameter a is small such as 2, 3, 4, ..., worst lattice points occur.
(b) If i and a are small, a^i is also small and results in bad lattices.
(c) If $a = 2^\alpha + 1$, for high dimensions short dual vectors occur. Since M can be expressed as $\Sigma c_j a^j$ for some integers c_j, the number of hyperplanes containing all lattice points will be $n_s \leq \Sigma|c_j|$. Therefore multipliers near a power of 2 induce bad lattice points.

Table 1. The values of μ_t near $M/2 \approx 1073741823$ for dimensions $2 \leq t \leq 6$

Multiplier	Dimensions (t)				
	2	3	4	5	6
1073741777	0.000	0.002	0.172	2.859	1.447
1073741784	0.000	0.001	0.090	3.182	0.132
1073741796	0.000	0.000	0.021	1.238	0.582
1073741799	0.000	0.000	0.013	0.695	0.843
1073741800	0.000	0.000	0.008	0.363	0.199
1073741805	0.000	0.000	0.004	0.171	6.228
1073741807	0.000	0.000	0.002	0.051	1.452
1073741812	0.000	0.000	0.001	0.016	0.350
1073741814	0.000	0.000	0.000	0.006	0.117
1073741815	0.000	0.000	0.000	0.004	0.061
1073741816	0.000	0.000	0.000	0.000	0.000
1073741817	0.000	0.000	0.000	0.000	0.000
1073741827	0.000	0.000	0.000	0.000	0.000
1073741829	0.000	0.000	0.000	0.000	0.000
1073741839	0.000	0.000	0.002	0.071	2.162
1073741843	0.000	0.000	0.003	0.092	2.948
1073741846	0.000	0.000	0.009	0.455	0.528
1073741849	0.000	0.000	0.013	0.647	1.731
1073741850	0.000	0.000	0.018	1.029	1.317
1073741852	0.000	0.000	0.024	1.479	0.730
1073741860	0.000	0.001	0.065	5.091	0.602
1073741861	0.000	0.001	0.073	5.827	0.362

(d) Apart from these cases causing bad lattice structure for a multiplier, good multipliers will generate bad sub-lattices if the lags $l = (m-1)/i$ are employed with small i values.

In the above classification, cases a and b correspond to the same situation, because if a and the power are small, the resulting parameter is also small. We would like to summarize the causes of the worst and the bad lattice structures under two headings:

(1) Cases where a can be expressed as $(K(M + n_1) + n_2)/N$ with small N and $n = Kn_1 + n_2$ values. This case will produce the worst lattice points.
(2) Cases where a is $(k_1 M/k_2 + n)^{1/t}$. If k_2 and n are small, this form will produce bad lattice points in dimension $t + 1$. These headings are examined below in detail.

3.1. Cases where $a = (K(M + n_1) + n_2)/N$

Sezgin [32] studied the behavior of the Euclidean distance and showed that v_T^2 becomes very small if a takes values close to KM/N where integers K and N are $0 \leq K < N$, and N is small. For example the μ_t values in Table 1 are obtained for $M = 2^{31} - 1$ by taking $K = 1$ and $N = 2$.

It is interesting to note that when μ_t is very small for two dimensional space, those of higher dimensions are also small. This fact is demonstrated by Figs. 1–3. It must also be noted that the area of bad multipliers gets narrower in higher dimensions.

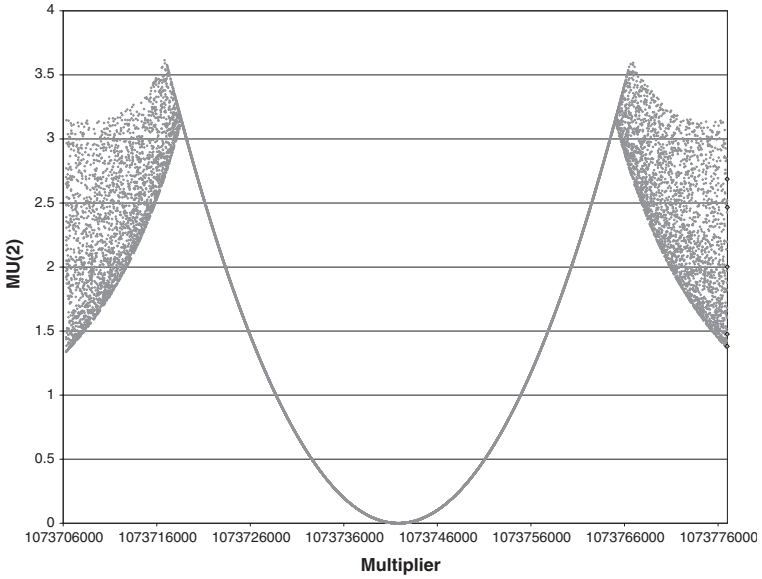

Fig. 1. Symmetrical distribution of μ_2 around kM/n as seen for $M/2 = 1073741824$

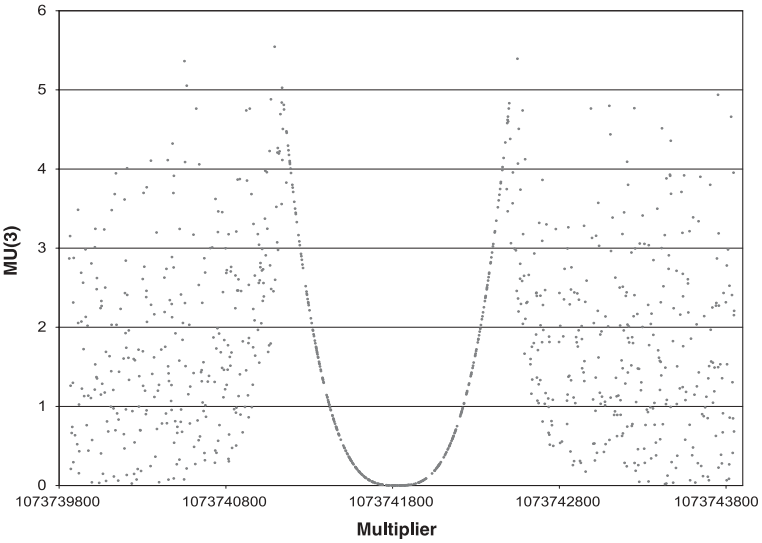

Fig. 2. Distribution of μ_3 around $M/2$ for $M = 2147483647$

In order to rate the lattice structure in a more general framework, we can use the fact that any multiplier a can be expressed as

$$a = \frac{K(M + n_1) + n_2}{N}, \tag{6}$$

where K, N, n_1 and n_2 are integers $0 < N < M$, $0 \leq K < N$, and $-N < n_1 < N$.

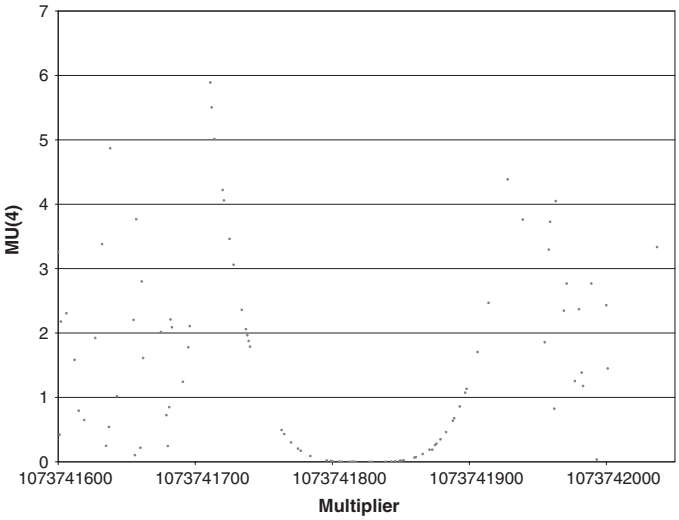

Fig. 3. Distribution of μ_4 around $M/2$ for $M = 2147483647$

Therefore, (1) can be written as

$$\sum_{i=1}^{t} \left(\frac{K(M+n_1)+n_2}{N} \right)^{i-1} S_i = kM \equiv 0 \pmod{M}. \tag{7}$$

Multiplying both sides by N^{t-1}, (7) can be written as

$$\sum_{i=1}^{t} N^{t-i} \left(KM + Kn_1 + n_2\right)^{i-1} S_i \equiv 0 \pmod{M}. \tag{8}$$

Letting $n = Kn_1 + n_2$ and expanding the binomial expression we get

$$\sum_{i=1}^{t} N^{t-i} S_i \sum_{j=0}^{i-1} \binom{i-1}{j} (KM)^j n^{i-j-1} = N^{t-1} kM.$$

If we group terms containing nonzero powers of M separately we get

$$\sum_{i=1}^{t} N^{t-i} S_i n^{i-1} = N^{t-1} kM - \sum_{i=1}^{t} N^{t-i} S_i \sum_{j=1}^{i-1} \binom{i-1}{j} (KM)^j n^{i-j-1}.$$

Noting that terms containing factor M are $\equiv 0 (\text{mod } M)$ in the right-hand side of the equation, we get the minimum Euclidean distance being subject to condition

$$\sum_{i=1}^{t} N^{t-i} n^{i-1} S_i^* \equiv 0 \pmod{M}. \tag{9}$$

Therefore the minimization problem reduces to solving the above relation and choosing the set $S_1^*, \ldots, S_t^*$. This expression is equivalent to

$$\sum_{i=1}^{t} N^{t-i} n^{i-1} S_i^* = k^* M \tag{10}$$

for an integer k^*. This equation does not depend on a and suggests a general identification method for the worst multipliers mentioned by Entacher et al [8]. Especially when N and n values are very small, the worst multipliers will be obtained. For small dimensions and moderate N values it is possible to get very small S_i^* values. For example, in $M = 2^{31} - 1$ when $N = 3, K = 2, n_1 = 1$ and $n_2 = -1$, we can get very small spectral test values for $t = 2$. Setting $k = 0$ it immediately follows that $3S_1^* + S_2^* = 0$. This equality has the minimal solution $S_1^* = 1$ and $S_2^* = -3$. Therefore $\nu_2^2 = 1^2 + (-3)^2 = 10$, a very small value. In a similar manner $\nu_3^2 = N^2 S_1^* + NnS_2^* + n^2 S_3^* = 0$ will have a solution satisfying $9S_1^* + 3S_2^* + S_3^* = 0$. With $S_1^* = 0$, $S_2^* = 1$, and, $S_3^* = -3$, we get $\nu_3^2 = 10$ again. The same value can be obtained for all other dimensions, because, for $t = T$ we will have $S_1^* = S_2^* = \ldots = S_{T-2}^* = 0$ but, $S_{T-1}^* = -3$, $S_T^* = 1$. It must be noted that for these constants, all possible multipliers are not acceptable for practice and a must be primitive root of M. Therefore n_2 must be –2. These examples explain why the worst lattice points are accumulated in certain areas.

By taking $S_{T-1}^* = -n$, $S_T^* = N$, in the worst lattice points, all dimensions will attain the same value. For this reason, in worst areas we will have $\nu_1^2 = \nu_2^2 = \ldots = \nu_T^2$ with a common solution since no smaller ν_t^2 value can be reached by including new nonzero lower degree terms. Worst lattice points will accumulate in certain areas. By changing n_1 and n_2 slightly, we observe a family of bad multipliers around the worst point. For example for $N = 2$, $K = 1$, and $n_1 = -1$, very small ν_2^2 will be observed. In $M = 2^{31} - 1$, by taking $n_2 = 8$, we will get the nearest primitive element $a = 1073741827$ with $N = 2, n = 7$ and $\nu_2^2 = 7^2 + 2^2 = 53$. In this multiplier, for all dimensions we will have the same test value. With increasing n, starting from the higher dimensions, there will be different ν_t^2 values. For example, when $n = 31$, in dimensions 7 and 8 we will have different results since the multiplier 1073741839 have the following ν_t^2 values for $t = 2, \ldots, 8$: 965, 965, 965, 965, 965, 324, and 165. Dimension 2 and 3 test results will remain identical until $a = 1073742497$, where $\nu_2^2 = \nu_3^2 = 1814413$. It is remarkable that when $k = 0$, bad points do not depend on M.

This approach also explains the bad multipliers presented by L'Ecuyer and Hellekalek [25]. They tabulate some good and bad (baby) LCGs for 25 moduli between $2^{12} - 3$ and $2^{36} - 5$. We present in Table 2 the explanation for 15 of these cases.

Table 2. Representation of lattice structure of some bad multipliers presented by L'Ecuyer and Hellekalek having largest prime moduli less than 2^e, for $12 \le e \le 26$

M	A	N	K	n_1	n_2	$n = Kn_1 + n_2$	${\nu_2}^2$
$2^{12} - 3$	5	1	0	0	5	5	26
$2^{13} - 1$	2341	7	2	2	1	5	74
$2^{14} - 3$	2731	6	1	5	0	5	61
$2^{15} - 19$	10	1	0	0	10	10	101
$2^{16} - 15$	17	1	0	0	17	17	290
$2^{17} - 1$	68985	19	10	1	−5	5	386
$2^{18} - 5$	203883	9	7	−4	2	−26	757
$2^{19} - 1$	458756	8	7	5	4	39	1585
$2^{20} - 3$	213598	54	11	−1	0	−11	3037
$2^{21} - 9$	202947	31	3	−24	0	−72	6145
$2^{22} - 3$	4079911	110	107	0	3	3	12109
$2^{23} - 15$	2696339	28	9	17	2	155	24809
$2^{24} - 3$	486293	69	2	−68	−73	−209	48442
$2^{25} - 39$	5431467	278	45	3	6	141	97165
$2^{26} - 5$	42038579	439	275	0	−44	−44	194657

3.2. *Cases where* $a = (k_1 M/k_2 + n)^{1/t}$

If a^t is $kM+n$, the $(t+1)$th dimension will have bad lattice for small n values. Because in Eq. (3), $\Sigma S_i a^{i-1}$ will be expressed as $-n + (kM + n) = kM \equiv 0 \,(\text{mod } M)$ and coefficients can be chosen as $S_1 = -n$ and $S_t = 1$. Since n is small and $\nu_t^2 = n^2 + 1$ does not depend on k, very small test results will be obtained. This case may be extended to situations $(k_1 M/k_2 + n)$ with small k_2 and n values. Because this will give

$$-nk_2 + (k_1 M + k_2 n) = k_2 kM \equiv 0 \,(\text{mod } M).$$

Letting $S_1 = -nk_2$ and $S_t = 1$ we get $\nu_t^2 = (nk_2)^2 + 1$. Table 3 represents a few examples of these cases obtained for $M = 2^{31} - 1$.

Examination of Table 3 shows that there is a fall in spectral test results for dimension $t + 1$ compared to the neighboring dimensions. These values are highlighted with bold numbers. The falls are very drastic when k_2 and n are very small. When the t-th root is very small, in other words, t is very large in operation $(.)^{1/t}$, we get very small multipliers as in the cases of cube root and fourth root in Table 3. As a result, very small test values will be observed in small dimensions too. This case is evident in multipliers 1285, 1625, 283, and, 952.

These two main patterns have also their counterparts as additive and multiplicative inverses. These are two relations facilitating the search for good or bad lattices. Once a multiplier is determined, we can say that its inverses with respect to modular addition and multiplication have the same lattice character. For example, if we have a multiplier a with spectral test value ν_t^2, we will have the same value for the additive inverse $\grave{a} = M - a$, because the expression (1) will lead to

Table 3. The spectral test values for various multipliers in the form $a = (k_1 M/k_2 + n)^{1/t}$

$\sqrt[t]{k_1 M/k_2 + n}$	a	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$	$t=8$
$\sqrt{M+282689}$	46344	0.931	**0.078**	0.439	0.532	0.669	0.697	0.537
$\sqrt{2M+524306}$	65540	0.658	**0.011**	0.063	0.178	0.349	0.557	0.436
$\sqrt[3]{M-25659522}$	1285	0.026	0.887	**0.417**	0.642	0.620	0.618	0.556
$\sqrt[3]{2M-3951669}$	1625	0.033	0.770	**0.331**	0.696	0.497	0.535	0.716
$\sqrt[4]{3M-28203020}$	283	0.006	0.192	0.534	**0.367**	0.690	0.692	0.562
$\sqrt{2M/5+141262}$	29311	0.589	**0.572**	0.687	0.682	0.655	0.673	0.648
$\sqrt[3]{2M/5+3807949}$	952	0.019	0.657	**0.437**	0.567	0.646	0.731	0.560

$$\sum_{i=1}^{t} S_i (M-a)^{i-1} \equiv 0 \,(\mathrm{mod}\ M).$$

In the binomial expansion all terms containing M will vanish and we will be left only with terms containing powers of a which will lead to the same test value. Sezgin [32] and Kao and Wong [18] considered additive inverses and gave examples. On the other hand, multiplicative inverse a^* has the property $aa^* = 1$ (Mod M) and produces the same set of random numbers but in reverse order. Several authors have presented multiplicative inverses. Fishman and Moore [13], Fishman [11], L'Ecuyer [22], and Tang [33] can be mentioned as examples.

4. Bad Regions

Table 1 and Figs. 1–3 show that very bad lattice values are obtained systematically around certain points. This fact was also observed in examples of Sect. 3.1. Now let us take the set of coefficients $S_1, \ldots, S_t$ and investigate the behavior of v_t in the neighborhood of a. Sezgin [32] showed that when y is small, the multiplier $a + y$ must reach the same integer multiple of M as in the Eq. (3) defined for multiplier a. Therefore $a + y$ will have coefficients $S_{y1}, \ldots, S_{yt}$ satisfying

$$\sum_{i=1}^{t} (a+y)^{i-1} S_{yi} \equiv 0 \ (\mathrm{mod}\ M). \tag{11}$$

By expanding binomial terms, and arranging with respect to powers of a, we will get

$$S_{t-k} = \sum_{i=0}^{k} \binom{t-k-1+i}{i} S_{y(t-k+i)} y^i. \tag{12}$$

These relations and pattern of constants are shown below explicitly in matrix form. Since most authors use dimensions up to 8, we contend to give the matrix up to this value.

$$\begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \\ S_7 \\ S_8 \end{pmatrix} = \begin{pmatrix} 1 & y & y^2 & y^3 & y^4 & y^5 & y^6 & y^7 \\ & 1 & 2y & 3y^2 & 4y^3 & 5y^4 & 6y^5 & 7y^6 \\ & & 1 & 3y & 6y^2 & 10y^3 & 15y^4 & 21y^5 \\ & & & 1 & 4y & 10y^2 & 20y^3 & 35y^4 \\ & & & & 1 & 5y & 15y^2 & 35y^3 \\ & & & & & 1 & 6y & 21y^2 \\ & & & & & & 1 & 7y \\ & & & & & & & 1 \end{pmatrix} \begin{pmatrix} S_{y1} \\ S_{y2} \\ S_{y3} \\ S_{y4} \\ S_{y5} \\ S_{y6} \\ S_{y7} \\ S_{y8} \end{pmatrix}. \tag{13}$$

The extension of this matrix representation is obvious. The system has a regular pattern and S_{yi} can be obtained by using the inverse of the matrix.

$$\begin{pmatrix} S_{y1} \\ S_{y2} \\ S_{y3} \\ S_{y4} \\ S_{y5} \\ S_{y6} \\ S_{y7} \\ S_{y8} \end{pmatrix} = \begin{pmatrix} 1 & -y & y^2 & -y^3 & y^4 & -y^5 & y^6 & -y^7 \\ & 1 & -2y & 3y^2 & -4y^3 & 5y^4 & -6y^5 & 7y^6 \\ & & 1 & -3y & 6y^2 & -10y^3 & 15y^4 & -21y^5 \\ & & & 1 & -4y & 10y^2 & -20y^3 & 35y^4 \\ & & & & 1 & -5y & 15y^2 & -35y^3 \\ & & & & & 1 & -6y & 21y^2 \\ & & & & & & 1 & -7y \\ & & & & & & & 1 \end{pmatrix} \begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \\ S_7 \\ S_8 \end{pmatrix}. \tag{14}$$

Therefore the solution for S_{yi} is

$$S_{y(t-k)} = \sum_{i=0}^{k} \binom{t-k-1+i}{i} S_{t-k+i}(-y)^i. \tag{15}$$

Now we can express the Euclidean distances for multipliers $a + y$, around the value a, as a function of $S_1, \ldots, S_t$. Since the definition of v_2^2 is $v_2^2 = S_1^2 + S_2^2$, it is possible to write the figure of merit v_{y2}^2 for multiplier $a + y$ as

$$\begin{aligned} v_{y2}^2 &= S_{y1}^2 + S_{y2}^2 \\ &= S_1^2 - 2S_1S_2y + S_2^2y^2 + S_2^2 \\ &= S_2^2y^2 - 2S_1S_2y + v_2^2. \end{aligned} \tag{16}$$

By similar calculations figure of merits for higher dimensions can be obtained. We present them here only up to $t = 5$:

$$v_{y3}^2 = S_3^2y^4 - 2S_2S_3y^3 + (4S_3^2 + S_2^2 + 2S_1S_3)y^2 - 2S_2(S_1 + 2S_3)y + v_3^2 \tag{17}$$

$$\begin{aligned} v_{y4}^2 = {} & S_4^2y^6 - 2S_3S_4y^5 + (S_3^2 + 2S_2S_4 + 9S_4^2)y^4 - 2(S_1S_4 + S_2S_3 + 6S_3S_4)y^3 \\ & + (S_2^2 + 2S_1S_3 + 4S_3^2 + 6S_2S_4 + 9S_4^2)y^2 - 2(S_1S_2 + 2S_2S_3 + 3S_3S_4)y + v_4^2 \end{aligned}$$

$$
\begin{aligned}
\nu_{y5}^2 = {} & S_5^2 y^8 - 2S_4S_5 y^7 + (S_4^2 + 2S_3S_5 + 16S_5^2) y^6 - 2(S_2S_5 + S_3S_4 + 12S_4S_5) y^5 \\
& + (S_3^2 + 2S_1S_5 + 2S_2S_4 + 9S_4^2 + 16S_3S_5 + 36S_5^2) y^4 \\
& - 2(S_1S_4 + S_2S_3 + 4S_2S_5 + 6S_3S_4 + 18S_4S_5) y^3 \\
& + (S_2^2 + 2S_1S_3 + 4S_3^2 + 6S_2S_4 + 9S_4^2 + 12S_3S_5 + 16S_5^2) y^2 \\
& - 2(S_1S_2 + 2S_2S_3 + 3S_3S_4 + 4S_4S_5) y + \nu_5^2.
\end{aligned}
$$

These relations clearly explain the behavior of lattice points. ν_{yt}^2 is a polynomial of degree $2(t-1)$ and this causes very chaotic behavior in higher dimensions.

A perusal of Figs. 1–3 obtained for $M = 2^{31} - 1$ reveals that μ_t values approach zero as the multiplier goes to $M/2$. This is a common phenomenon for multipliers of the form $(KM + n)/N$ for small n and N. At the right and left sides of minimum points, μ_t starts to increase. This increase continues until μ_t reaches a maximum. For example, μ_2 reaches its maxima at *1073716869* at left and *1073766925* at right. There is a distance of 50056 between these summits. Similar peaks are observed for higher dimensions. For example μ_3 has peaks at a = *1073741164* and *1073742500*. The distance between maxima gets shorter with increasing t. Therefore the search for bad lattice points areas must start from smaller dimensions. Since ν_{yt}^2 is a polynomial of degree $2(t-1)$, the figures of merit are very erratic for large dimensions when M is not extremely large. For a 32 bit modulus, the investigation of bad regions in two-dimensional space will be an efficient search strategy. For 64 and 128 bits however, it will be worthwhile to take into account bad regions in the third and forth dimensions.

The relation between the divisor N and starting point of the areas for good or bad lattice points exhibits a very regular and simple form. It is possible to express this relation more concisely. Referring to the formula of figure of merit μ_t, assume that we want to find the end point of the bad area where $\mu_2 < C$. By noting that $S_2 = N$, $S_1 = KM - Na$, $a \approx KM/N$, and using Eqs. (2)–(4), and (16), we get the lower and upper limits of bad multipliers region as $a - y$ and $a + y$ where

$$
y = \frac{1}{N}\sqrt{\frac{CM}{\pi}}. \tag{18}
$$

Example 1: For M = *2147483647* and $N = 2$ we get a = *1073741824* with a very small figure of merit $5\pi/M \approx 0$. The y value satisfying the above inequality is $y = 13073\sqrt{C}$. Therefore if we choose $C = 0.01$, it may be said that there is a bad lattice area having $\mu_2 < 0.01$, between 1073740516 and 1073743131. This agrees remarkably well with the actual calculations.

Example 2: Currently, the moduli having 256 bits are not common. Assume that in future we will need 256 bit muduli and during a search we will try to find a region of bad multipliers having $\mu_2 < 0.0001$ near an arbitrary KM/N value such as $a \approx 71M/78942$. Then we must look below $a + y$ and above $a - y$ where

$$y = \frac{2^{128}}{78942}\sqrt{\frac{0.0001}{\pi}} = \frac{2^{128}}{7894200\sqrt{\pi}} = 1.37209 \times 10^{31}.$$

In application the search for good or bad multipliers must start from the largest areas and go gradually to narrower ones. This implies starting from $N = 1$ and going to 2, 3, 4, ..., etc. Although the length of area is inversely proportional to N, growing K will compensate this loss and the cumulative number of multipliers having a certain quality increases with N.

5. Some Properties of Lattice Points Distribution

The distribution of lattice quality measures for different dimensions is very crucial for thoroughly understanding the behavior of multipliers. But these investigations are surprisingly neglected in the random number generation literature. Entacher et al [8] remarked that the normalized spectral test measure is not suitable for use with a fixed threshold value such as $M_T = 0.80$ adopted by many authors during their evaluation of random number generators. Figure 4 gives the cumulative frequency distribution of normalized measure for dimensions 2–6 obtained on 10,167,840 primitive roots out of 20 million possible candidates by a random search for $M = 2^{31} - 1$.

In Fig. 4, the cumulative curves of smaller dimensions are above the curves of higher dimensions for small S values. The situation is opposite for higher S values. For example in dimension $t = 2$, 9.9% of the multipliers have normalized test measure below 0.3. For the same threshold in dimensions 3, 4, 5, and 6, these proportions are 6.5, 3.6, 1.7, and 0.9, respectively. On the other hand, in the good quality side,

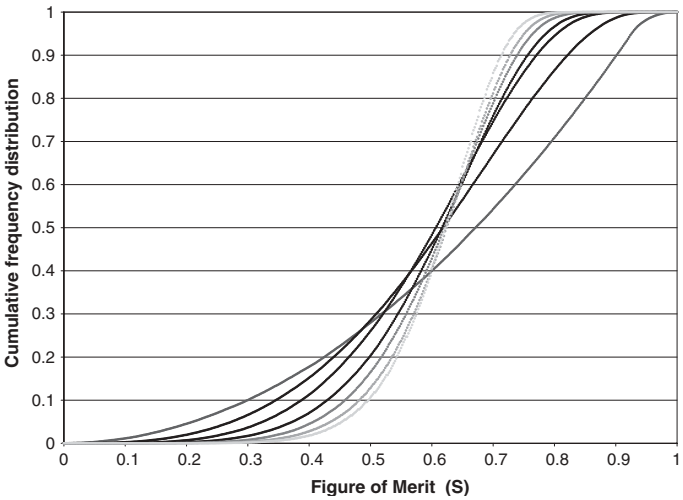

Fig. 4. Cumulative distributions of normalized spectral test value S

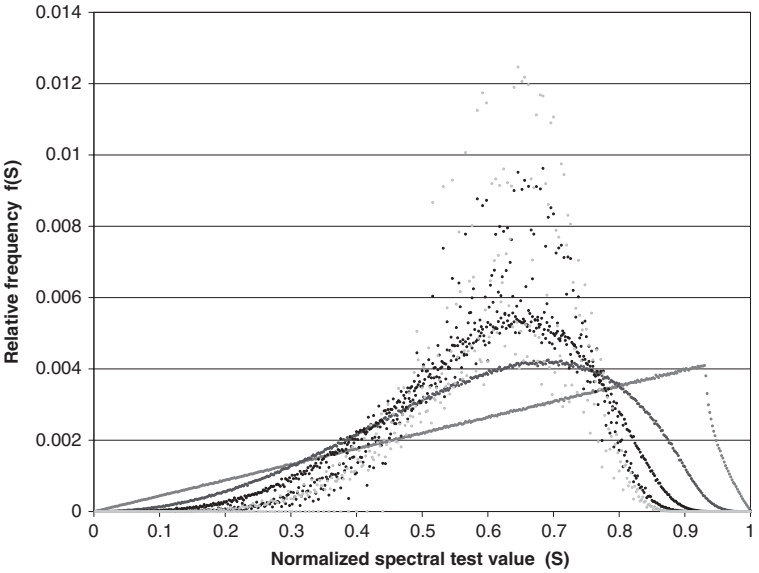

Fig. 5. Empirical frequency distributions of normalized spectral test value S

the multipliers having S values above the conventional 0.8 level is 29.4% for $t = 2$. In higher dimensions, this proportion falls to 13.8, 5.8, 3.3, and 1.2%, respectively. These data clearly show that fixed thresholds correspond to different percentiles in different dimensions. It is more selective in higher dimensions but allows a lower quality in smaller dimensions. This fact is not acceptable from a practical point of view because bad lattice structure will be more frequent for small t values.

Our study also reveals that discrete nature of spectral test results are displayed clearly for small moduli. For example, according to our data (not presented here), for $M = 32767$, $F(S)$ curves are step-functions. This is the natural result of filling the t-dimensional space with a limited number of points. On the other hand for large moduli, this discrete nature is still observed in higher dimensions. This is clearly seen in scattered points representing dimensions 5 and 6 for the empirical frequency distributions depicted in Fig. 5. In smaller dimensions, however, frequency points accumulate densely about regular curves.

The above properties of distribution curves are caused by remarkable patterns of the frequency curves of normalized test results in different dimensions. The frequency distribution obtained by normalizing ν_2 is particularly interesting. The curve is a straight line with slope 0.0044 until $S = 0.93$. From this point on, the frequency falls rapidly implying a sharp decrease in the proportion of exceptionally good quality multipliers. The curves are ordered from right to left in the plot area according to their increasing t values. The curves for smaller dimensions remain above others in the tail areas. The case is opposite near the medial region. It is interesting to see that tail areas get consistently smaller with increasing t, and observations accumulate densely about the mode values.

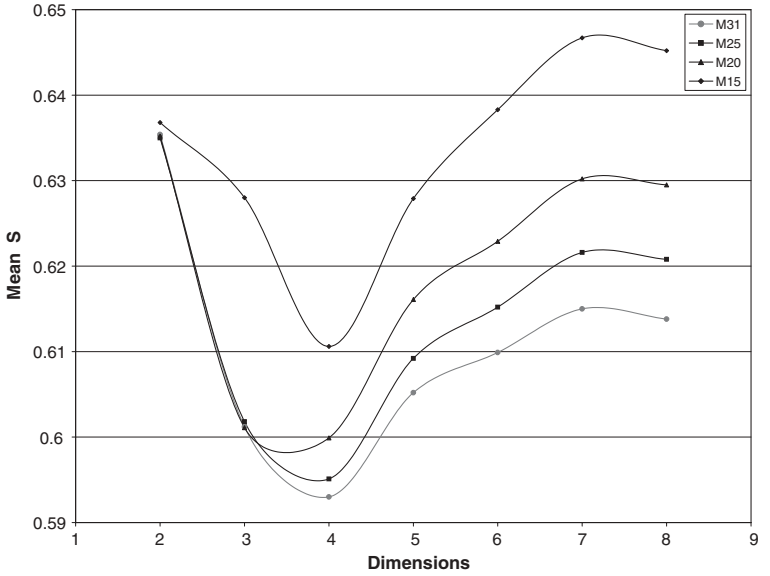

Fig. 6. Mean values of S for different modulus sizes and dimensions

We studied the normalized test values for different moduli and dimensions in detail. For this purpose, four multipliers are examined. Nearest prime modulus values to 2^{15}, 2^{20}, 2^{25} and 2^{31}, and number of prime roots tested are summarized below:

Modulus	**Number of prime roots**
M15 = $2^{15} - 19$ = 32,749	10,912
M20 = 2^{20}+ 7 = 1,048,583	133,485
M25 = 2^{25}+35 = 33,554,467	250,000
M31 = $2^{31} - 1$ = 2,147,483,647	124,473

Important descriptive characteristics of these curves are presented in Figs. 6–11. We investigate below these figures in detail.

(1) Means: Means of S values are depicted in Fig. 6. In dimension $t = 2$, means of different modulus values do not have great differences. *M31* has a mean value of 0.635 whereas *M15* has a mean value of 0.637. The difference grows with increasing dimensions. Since smaller moduli give coarser lattice distributions in higher dimension, they have higher means. For example, *M31* has a mean of 0.614 in $t = 8$, whereas the mean of *M15* is 0.645. There are interesting patterns in Fig. 6. After a general fall in dimensions 3 and 4, all generators start to give progressively higher means with increasing dimensions. The increasing divergence between the generators with increasing dimensions is remarkable. This is the result of enhanced filling capacity of generators with larger periods.

(2) Variances: According to Fig. 7, unlike means, the distribution of variances is quite stable between different modulus values. For $t = 2$, *M31* has a variance 0.0507.

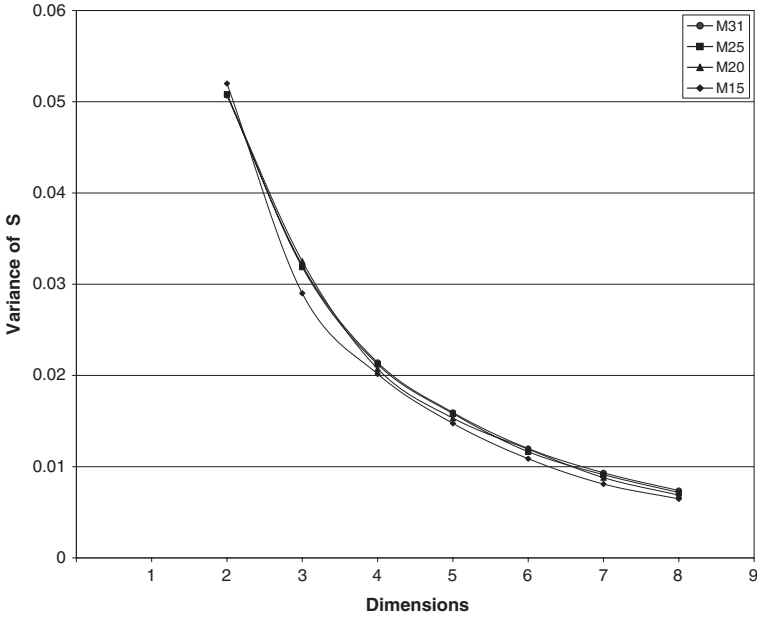

Fig. 7. Distributions of variances of *S* for different modulus sizes and dimensions

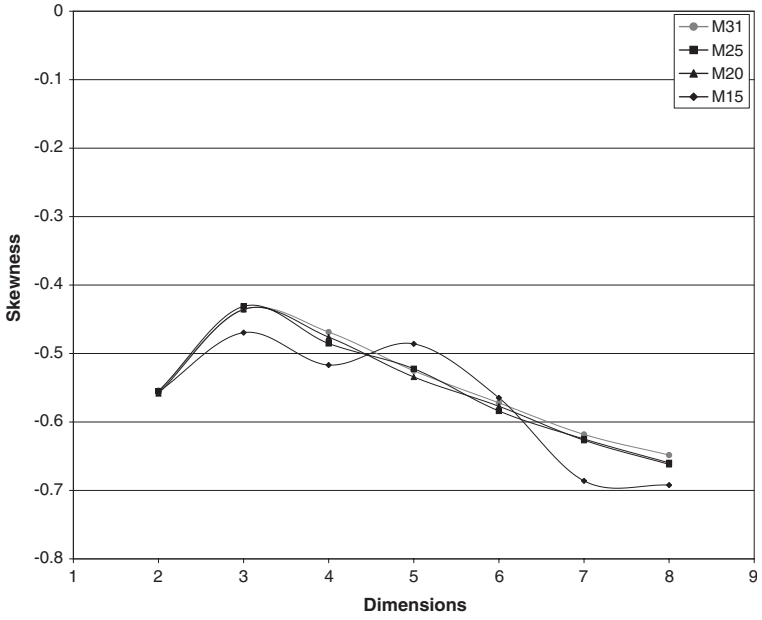

Fig. 8. Skewness coefficients of *S* for different modulus sizes and dimensions

The variance of *M15* is 0.0520, not a very different value. Other two variances are both 0.0508. The variances of generators *M31*, *M25*, *M20*, and *M15*, for $t = 8$, are 0.0074, 0.0072, 0.0064 and 0.0065, respectively. Although they do not change

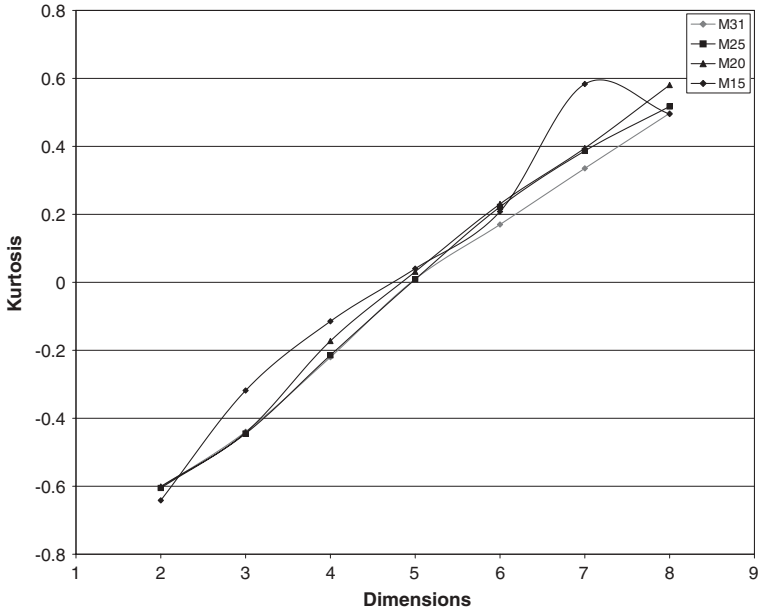

Fig. 9. Kurtosis coefficients of S for different modulus sizes and dimensions

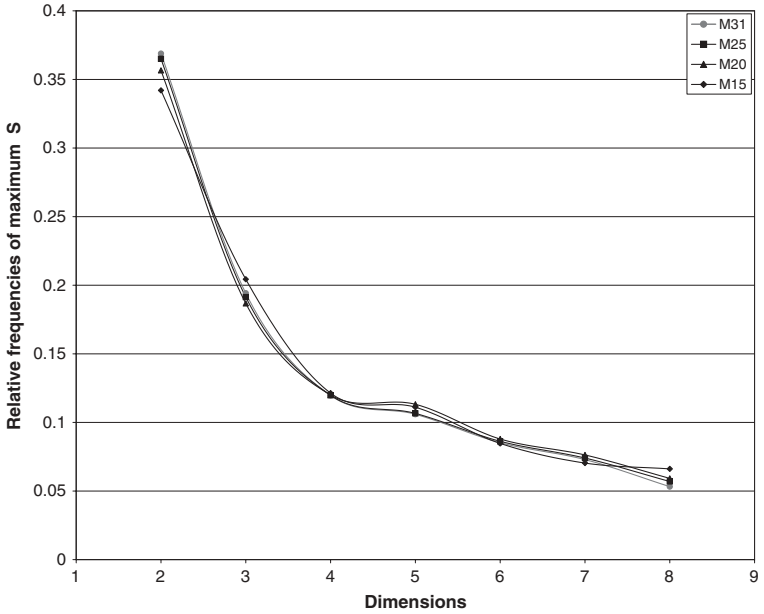

Fig. 10. Relative frequencies of maximum normalized spectral test values S

very much between different modulus values, there is an obvious tendency of having smaller variances with increasing dimensions. This is the result of diminishing tail regions with increasing dimensions. As seen from Fig. 5, by increasing dimensions,

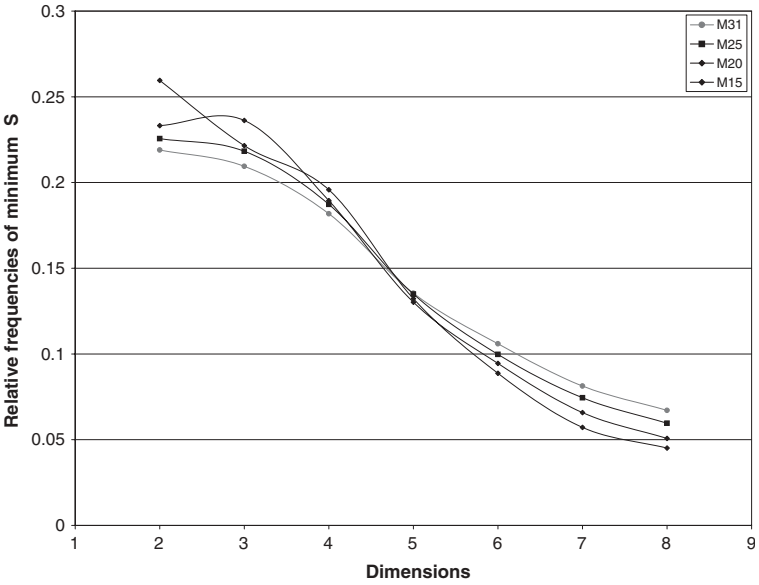

Fig. 11. Relative frequencies of minimum normalized spectral test values S

the curves tend to have smaller tail areas and values accumulate densely about the mode.

(3) Skewness: Skewness is an important property of distributions representing the asymmetry with respect to mean. Figure 8 shows that the distribution of normalized test value S is skew to left for all modulus sizes and skewness grows with increasing dimension. This property implies that values larger than the mean are more frequent and the right tails get heavier with increasing dimension. One can deduce the same information from Fig. 12 by examining distances between various percentiles. In this figure, distances between lower quantiles are always larger than distances between higher quantiles. These differences get smaller with increasing dimension.

(4) Kurtosis: Kurtosis gives useful information by expressing the excess of observations near the mean and far from it. Figure 9 shows that curves are platykurtic for $t = 2, 3,$ and 4. After $t = 5$ they become increasingly leptokurtic. This means that the distribution is flat for small dimensions therefore contain less observation near mean and more observations in tails. In larger dimension, however, curves are peaked. These facts can also be deduced from the Fig. 12 of percentiles. For smaller dimensions percentiles are situated in a larger region, but for larger dimensions, frequencies tend to accumulate densely near the central areas.

(5) Minimum and maximum S values: Fixed threshold uses a single value across dimensions and therefore the failure of a single dimension is enough to reject the multiplier. Our detailed examination showed that both large and small values are more frequent in smaller dimensions. This fact can be seen in Figs. 10 and 11. Two-dimensional test shows the maximum S value with approximately 35% probability.

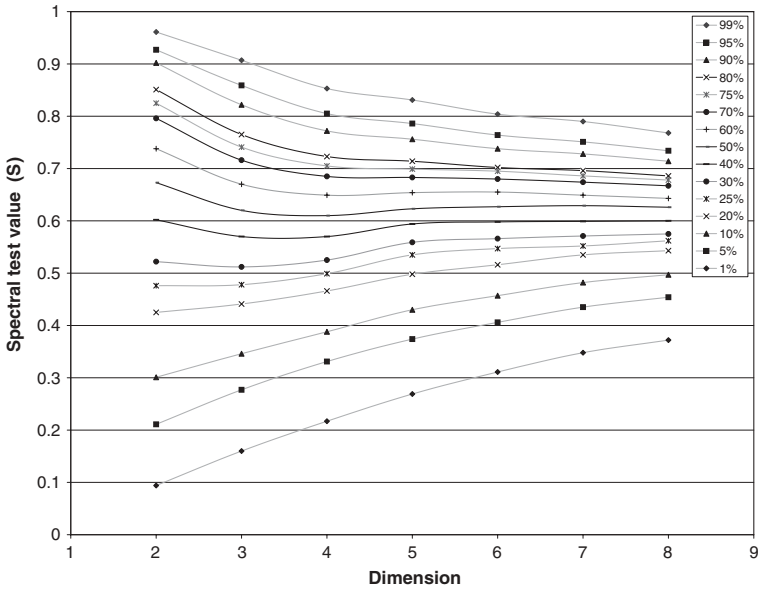

Fig. 12. Percentiles of normalized spectral test value (S) for $M = 2^{31} - 1$

Table 4. Some useful percentiles of normalized spectral test value S

Percentile	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$	$t = 7$	$t = 8$
0.01	0.095	0.164	0.223	0.272	0.322	0.348	0.372
0.05	0.213	0.277	0.331	0.378	0.419	0.435	0454
0.10	0.302	0.349	0.393	0.443	0.468	0.482	0.497
0.20	0.426	0.443	0.470	0.506	0.526	0.535	0.544
0.80	0.851	0.767	0.726	0.714	0.709	0.696	0.686
0.90	0.903	0.824	0.774	0.752	0.744	0.728	0.714
0.95	0.928	0.861	0.808	0.781	0.770	0.751	0.734
0.99	0.964	0.909	0.856	0.825	0.809	0.790	0.768

The probability decreases steadily and $t = 8$ has only about 7% probability of representing the maximum S.

Although the former probability gets slightly higher and the second slightly smaller with increasing modulus size, this pattern remains unchanged across modulus values. Similar comments are valid for minimum S. More than 20% of minimums are recorded in $t = 2$. For $t = 8$, this is slightly above 5%. As for the maximums, the pattern remains the same across modulus sizes, but with a slightly higher percentage of small modulus in smaller dimensions and higher percentage of larger modulus in higher dimensions. Entacher et al [8] noted that: "We observe a steady decrease of the number of multipliers rated as bad with increasing dimensions t and no more such multipliers for dimensions $t \geq 7$. The phenomenon may be explained in two ways: either there are no poor quality lattice rules in higher dimensions or our criterion is not suited to identify them." According to the percentages of Figs. 4, 5 and 12, it is now clear that the first explanation is valid for this fact.

As a guide for applied test studies we present some percentiles of normalized lattice test in Table 4. Readers are reminded that these are calculated from 20 million

candidate multipliers of 31 bit generators described in Sect. 5 and presented in Fig. 12. Usage of percentiles is desirable for two reasons. As can be seen from the Fig. 12, the first reason is the lack of a common fixed threshold for different dimensions. The second reason is the validity of percentiles for different figure of merit criteria. As pointed out in Sect. 2, lattice structure can be assessed by different measures. Since the relation between these measures preserves the order of magnitudes, percentiles of different figures of merits will have a correspondence, whereas there is no correspondence between values obtained as percentages of the maximum attainable values of different measures. For example if in t dimensional space ν_t exceeds *80*% of the maximum attainable value ν_t, the standardized value μ_t will exceed only $100*0.8^t$% of the maximum attainable value μ_t. ν_t having 0.8 spectral threshold value of M_T, will produce $\mu_t = 6.26$ in dimension $t = 6$ that corresponds only to 26% of the upper bound 23.87.

6. Conclusions

In the present article it is shown that multipliers producing worst lattice points form certain clusters. The same is true for best multipliers. Using these facts it is possible to diagnose systematically certain regions as fertile and unproductive before detailed numerical searches.

Some distributional properties of lattice test are also investigated theoretically and results are supported by a large body of empirical data. It is shown that the usage of conventional fixed threshold technique applied on normalized spectral test figure of merit is not appropriate as the measure of quality. This method gives greater emphasis on higher dimensions and neglects the most important part, the smaller dimensions. Therefore a threshold vector based on percentiles of the distribution is more appropriate.

Acknowledgements

I would like to thank the editor Prof. Craig C. Douglas and two anonymous referees for their helpful guidance during the revision process. I also would like to thank my son Dr. Tevfik Metin Sezgin for his useful discussion concerning certain equations and contribution for the LaTeX version of the manuscript.

References

[1] Afflerbach, L.: Criteria for the assessment off random number generators. J. Comp. Appl. Math. *31*, 3–10 (1990).
[2] Coddington, P. D., Ko, S. H.: Techniques for empirical testing of parallel random number generators. Technical Report DHPC-025. New York: Syracuse University 1998.
[3] Conway, J. H., Sloan, N. J. A.: Sphere packings, lattices and groups. Grundlehren der Mathematischen Wissenschaften 290. New York: Springer 1988.
[4] Dyadkin, I. G., Hamilton, K. G.: A study of 64-bit multipliers for Lehmer pseudorandom number generators. Comput. Phys. Commun. *103*, 103–130 (1997).
[5] Dyadkin, I. G., Hamilton, K. G.: A study of 128-bit multipliers for congruential pseudorandom number generators. Comput. Phys. Commun. *125*, 239–258 (2000).
[6] Entacher, K.: A collection of classical pseudorandom number generators with linear structures – advanced version. http://crypto.mat.sbg.ac.at/results/karl/server/ June 16, 2000.
[7] Entacher, K., Schell, T., Uhl, A.: Efficient lattice assessment for LCG and GLP parameter searches. Math. Comp. *71*, 1231–1242 (2002).

[8] Entacher, K., Schell, T., Uhl, A.: Bad lattice points. Computing *75*, 281–295 (2005).
[9] Ferrenberg, A. M., Landau, D. P., Wong, Y. J.: Monte Carlo simulations: Hidden errors from "good" random number generators. Phys. Rev. Lett. *69*, 3382–3374 (1992).
[10] Filk, T., Marcu, M., Fredenhagen, K.: Long range correlations in random number generators and their influence on Monte Carlo simulations. Phys. Lett. *165B*(1–3), 125–130 (1985).
[11] Fishman, G. S.: Multiplicative congruential random number generators with modulus 2^{β}: An exhaustive analysis for $\beta = 32$ and a partial analysis for $\beta = 48$. Math. Comp. *54*, 331–344 (1990).
[12] Fishman, G. S.: Monte Carlo concepts, algorithms, and applications. New York: Springer 1996.
[13] Fishman, G., Moore, L. R.: An exhaustive analysis of multiplicative congruential random number generators with modulus $2^{31} - 1$. SIAM J. Sci. Stat. Comput. *7*, 24–45 (1986).
[14] Kao, C., Tang, H. C.: Several extensively tested multiple recursive random number generators. Comput. Math. Appl. *36*, 129–136 (1998).
[15] Kao, C., Tang, H. C.: Upper bounds in spectral test for multiple recursive random number generators with missing terms. Comput. Math. Appl. *33*, 119–125 (1997).
[16] Kao, C., Wong, J. Y.: Several extensively tested random number generators. Comput. Ops. Res. *21*, 1035–1039 (1994).
[17] Kao, C., Wong, J. Y.: An exhaustive analysis of prime modulus multiplicative congruential random number generators with modulus smaller than 2^{15}. J. Statist. Comp. Simul. *54*, 29–35 (1996).
[18] Kao, C., Wong, J.Y.: Random number generators with long period and sound statistical properties. Comput. Math. Appl. *36*, 113–121 (1998).
[19] Knuth, D. E.: The art of computer programming, vol. 2: Seminumerical algorithms, 3rd ed. Reading, MA: Addison-Wesley 1998.
[20] Kurita, Y.: Choosing parameters for congruential random number generators. In: Recent developments in statistics (Barra, J. R. et al., eds.) Amsterdam: North Holland 1997, pp. 697–704.
[21] L'Ecuyer, P.: Random numbers for simulation. Commun. ACM. *33*, 86–97 (1990).
[22] L'Ecuyer, P.: Tables of linear congruential generators of different sizes and good lattice structure. Math. Comp. *68*, 249–260 (1999).
[23] L'Ecuyer, P., Blouin, F., Couture, R.: A search for good multiple recursive random number generators. ACM Trans. Model. Comput. Simulation *3*, 87–98 (1993).
[24] L'Ecuyer, P., Couture, R.: An implementation of the lattice and spectral tests for multiple recursive linear random number generators. INFORMS J. Comput. *9*, 206–217 (1997).
[25] L'Ecuyer, P., Hellekalek, P.: Random number generators: Selection criteria and testing. In: Random and Quasi-random Point Sets (Hellekalek, P. and Larcher G., eds.). Lectures Notes in Statistics, vol. *138*. Springer 1998, pp. 223–266.
[26] L'Ecuyer, P., Tezuka, S.: Structural properties for two classes of combined random number generator. Math. Comp. *57*, 735–746 (1991).
[27] Leeb, H.: Random numbers for computer simulation. Master's thesis, Institute für Mathematik, Universität Salzburg, Austria, 1995. Available from: http://random.mat.sbg.ac.at/.
[28] Lemieux, C., L'Ecuyer, P.: On selection criteria for lattice rules and other quasi–Monte Carlo point sets. Math. Comput. Simulation *55*, 139–148 (2001).
[29] Niederreiter, H.: Random number generation and quasi–Monte Carlo methods. Society for Industrial and Applied Mathematics, Philadelphia, PA 1992.
[30] Ripley, B.D.: Stochastic simulation. New York: Wiley 1987.
[31] Sezgin, F.: Some improvements for a random number generator with single-precision floating-point arithmetic. Comput. Geosci. *22*(4), 453–455 (1996).
[32] Sezgin, F.: A method of systematic search for optimal multipliers in congruential random number generators. BIT Numer. Math. *44*, 135–149 (2004).
[33] Tang, H. C.: Modulus of linear congruential random number generator. Quality Quantity *39*, 413–422 (2005).

F. Sezgin
Bilkent University
06580 Ankara
Turkey
e-mail: fatin@bilkent.edu.tr

Verleger: Springer-Verlag GmbH, Sachsenplatz 4–6, 1201 Wien, Austria. – Herausgeber: Prof. Dr. Wolfgang Hackbusch, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Inselstraße 22–26, 04103 Leipzig, Germany. – Satz und Umbruch: Scientific Publishing Services (P) Ltd., Chennai, India. – Offsetdruck: Manz Crossmedia, 1051 Wien, Austria. – Verlagsort: Wien. – Herstellungsort: Wien. – Printed in Austria.