TABLES OF MAXIMALLY-EQUIDISTRIBUTED COMBINED LFSR GENERATORS

PIERRE L'ECUYER

ABSTRACT. We give the results of a computer search for maximally-equidistributed combined linear feedback shift register (or Tausworthe) random number generators, whose components are trinomials of degrees slightly less than 32 or 64. These generators are fast and have good statistical properties.

1. INTRODUCTION

Linear Feedback Shift Register (LFSR) random number generators, also called Tausworthe generators, are based on linear recurrences modulo 2 with primitive characteristic polynomials. Efficient implementations are available for the case where the characteristic polynomial is a trinomial and satisfies some additional conditions. Trinomial-based generators have important statistical defects, but combining them can yield generators that are relatively fast and robust. Such combinations have been proposed and analyzed in [4, 9, 10]. In [4], it was explained how to find combined generators with the best possible equidistribution properties in some sense, within specified classes of combined LFSR generators. Three specific combined generators, each with three components and period length near 2^{88}, were also given. In the present paper, we provide the results of more extensive computer searches, for combined generators with larger periods. The need for large periods is supported by several arguments given, e.g., in [2, 3, 5]. The generators given in [4] are for 32-bit computers. Since 64-bit computers are becoming increasingly common, it is important to have good generators designed to fully use the 64-bit words. Some of the generators proposed here do it.

The next section explains how we combine LFSR generators and recalls definitions and properties. Section 3 gives specific combined generators of different sizes. Section 4 provides computer implementations in C.

2. COMBINED LFSR GENERATORS AND EQUIDISTRIBUTION

Consider the LFSR recurrence

$$x_n = (a_1 x_{n-1} + \cdots + a_k x_{n-k}) \bmod 2, \tag{1}$$

whose characteristic polynomial is $P(z) = z^k - a_1 z^{k-1} - \cdots - a_k$. This is a linear recurrence in the finite field $\mathbf{F}_2$ with two elements, 0 and 1. The recurrence has

1991 *Mathematics Subject Classification.* 65C10.

Key words and phrases. Random number generation, equidistribution, combined generators, finite fields.

This work has been supported by NSERC-Canada grants # ODGP0110050 and SMF0169893, and FCAR-Québec grant # 93ER1654. I wish to thank Luc De Bellefeuille and Armand Nganou who helped performing the computer searches.

period length $\rho = 2^k - 1$ if and only if P is a primitive polynomial, which we now assume. Let

$$u_n = \sum_{i=1}^{L} x_{ns+i-1} 2^{-i}, \tag{2}$$

where the *step size* s and the *word length* L are positive integers. If $(x_0, \ldots, x_{k-1}) \neq 0$, and s is coprime to ρ, then the sequence (2) is also purely periodic with period ρ. An LFSR (or Tausworthe) random number generator is one that outputs a sequence $\{u_n,\ n \geq 0\}$ defined by (2).

Suppose now that we have J LFSR recurrences, the jth one having a primitive characteristic polynomial $P_j(z)$ of degree k_j, and step size s_j relatively prime with $\rho_j = 2^{k_j} - 1$. Assume that the $P_j(z)$ are pairwise relatively prime, that the ρ_j are also relatively prime, and that these LFSRs use a common L. Let $\{x_{j,n},\ n \geq 0\}$ be the jth LFSR sequence, and define $x_n = (x_{1,n} + \cdots + x_{J,n}) \bmod 2$ and u_n as in (2). Equivalently, if $\{u_{j,n},\ n \geq 0\}$ is the output sequence from the jth LFSR, then $u_n = u_{1,n} \oplus \cdots \oplus u_{J,n}$ where $\oplus$ denotes the bitwise exclusive-or in the binary expansion. The sequence $\{x_n\}$ is called the combined LFSR sequence and a generator that produces this $\{u_n\}$ is called a *combined LFSR* generator. In fact, $\{x_n\}$ follows a recurrence with reducible characteristic polynomial $P(z) = P_1(z) \cdots P_J(z)$ [9]. Under our assumptions, the sequences $\{x_n\}$ and $\{u_n\}$ have period length $\rho = (2^{k_1} - 1) \times \cdots \times (2^{k_J} - 1)$. This type of combination is interesting because it permits one to conciliate efficient implementation with statistical robustness, by choosing the P_j as trinomials for which the recurrence is easy to implement and runs fast, while making sure that $P(z)$ has many non-zero coefficients and that the combined generator has good equidistribution properties [1, 7, 10]. Of course, this is not the only way of constructing generators with good equidistribution; for other approaches, see, e.g., [5, 6, 8] and other references given there.

Let T_t be the set (in the sense of a *multiset*) of t-dimensional vectors of successive output values, from all possible initial states:

$$T_t = \left\{ \boldsymbol{u}_n = (u_n, \ldots, u_{n+t-1}) \mid n \geq 0,\ (x_0, \ldots, x_{k-1}) \in \{0,1\}^k \right\}.$$

Dividing the interval $[0,1)$ into 2^ℓ equal segments determines a partition of the unit hypercube $[0,1)^t$ into $2^{t\ell}$ cubic cells of equal size, called a (t,ℓ)*-equidissection* in base 2, and the set T_t is said to be (t,ℓ)*-equidistributed* if each cell contains the same number of points of T_t. The latter is possible only if $\ell \leq L$ and $t\ell \leq k$. If T_t is (t,ℓ_t^*)*-equidistributed* for $0 \leq t \leq k$, where $\ell_t^* = \min(L,\ \lfloor k/t \rfloor)$, then the (output) sequence is called *maximally-equidistributed* (ME). An ME sequence for which all non-empty cells contain exactly one point, for $t \geq 1$ and $\ell_t^* < \ell \leq L$ (i.e., when there are more cells than points), is called *collision-free* (CF). ME-CF sequences enjoy nice equidistribution properties; their point sets are very evenly distributed in all dimensions, in terms of equidissections. Verifying whether a sequence is ME or ME-CF amounts to computing the rank of a binary matrix that expresses the relevant bits of $\boldsymbol{u}_n$ in terms of $(x_{1,0}, \ldots, x_{1,k_1-1}), \ldots, (x_{J,0}, \ldots, x_{J,k_J-1})$, for different values of t, as explained in [4].

The above definitions of ME and ME-CF are based on the ℓ most significant bits of each u_n, so when t is large, we look only at a few most significant bits. What about the least significant bits? For the LFSR generators considered here, it turns out that any successive ℓ bits in each u_n have the same equidistribution properties as the most significant ones. More specifically, let r be an integer such that $0 \le r \le L - \ell$ and define

$$v_n = 2^r u_n \bmod 1 = \sum_{i=1}^{L-r} x_{r+ns+i-1} 2^{-i}.$$

Then, for any box C in the (t, ℓ)-equidissection,

$$\begin{aligned} & \left\{\boldsymbol{v}_n = (v_n, \ldots, v_{n+t-1}) \in C \mid n \ge 0,\, (x_0, \ldots, x_{k-1}) \in \{0,1\}^k\right\} \\ = \quad & \left\{\boldsymbol{u}_n = (u_n, \ldots, u_{n+t-1}) \in C \mid n \ge 0,\, (x_r, \ldots, x_{r+k-1}) \in \{0,1\}^k\right\}. \end{aligned}$$

Therefore, the sequence $\{v_n\}$ has exactly the same (t, ℓ)-equidistribution properties as $\{u_n\}$.

3. Some Maximally-Equidistributed Collision-Free Generators

We now give ME-CF combined LFSR generators with word-lengths $L = 32$ and 64, whose components have recurrences with primitive trinomials of the form $P_j(z) = z^{k_j} - z^{q_j} - 1$ with $0 < 2q_j < k_j$, and with step size s_j satisfying $0 < s_j \le k_j - q_j < k_j \le L$ and $\gcd(s_j, 2^{k_j} - 1) = 1$. Components that satisfy these conditions are implemented easily using the algorithm described in [4]. When they satisfy the additional condition that

$$L - k_j \le r_j - s_j \tag{3}$$

for all j, then the initialization procedure in [4, p. 205] is not necessary. All the parameter sets given in the forthcoming tables satisfy this additional condition.

For $L = 32$, three specific ME-CF generators with $J = 3$ were given in [4], and it was reported that there are 4744 ME-CF generators with $J = 4$, $k_1 = 31$, $k_2 = 29$, $k_3 = 28$, and $k_4 = 25$, among the 3.28 million that satisfy all our conditions except for (3). Since this paper was published, several people asked the author for specific instances of such generators. Table 1 gives a partial list. These combined generators have period lengths $(2^{31}-1)(2^{29}-1)(2^{28}-1)(2^{25}-1) \approx 2^{113}$ and their characteristic polynomials have degree 113. The 62 generators in Table 1 satisfy (3). They all have $(q_1, q_2, q_3, q_4) = (6, 2, 13, 3)$, so they have the same characteristic polynomial $P(z)$, which has 58 coefficients equal to zero and 55 coefficients equal to 1.

The following tables give selected results of random searches for ME-CF generators with $L = 64$, and with $J = 3$, 4, and 5 components. Here, $k = k_1 + \cdots + k_J$ is the degree of the product polynomial associated with the combination, N_1 is the number of coefficients that are 1 in that polynomial, and $\lg \rho = \text{lcm}\ (k_1, \ldots, k_J)$ is (approximately) the logarithm in base 2 of the period length of the generator.

In Table 2, the first 4 generators have full period length $\rho = (2^{k_1} - 1)(2^{k_2} - 1)(2^{k_3} - 1) \approx 2^k$. The remaining 6 do not have full period, because the k_j are not co-prime. Note that for all generators in this table, N_1 is rather small in comparison with k; that is, the characteristic polynomials have much more zeros than ones.

Table 3 gives 8 full-period ME-CF generators with $L = 64$, $J = 4$, $(k_1, k_2, k_3, k_4) = (63, 58, 55, 47)$, and $(q_1, q_2, q_3, q_4) = (31, 19, 24, 21)$. Their period length is approximately 2^{223} and their characteristic polynomial $P(z)$ (they all have the same) has 49 coefficients (out of 223) equal to 1. Table 4 gives a partial list of ME-CF generators with $(k_1, k_2, k_3, k_4) = (63, 58, 57, 55)$ and $(q_1, q_2, q_3, q_4) = (1, 19, 7, 24)$, so $k = 233$ and $\lg \rho = 230$, whereas Table 5 gives ME-CF generators with $(k_1, k_2, k_3, k_4) = (63, 60, 58, 57)$, which gives $k = 238$ and $\lg \rho = 220$. In all cases, the number of ones

Table 1. ME-CF generators with $L = 32$ and $J = 4$.

	s_1	s_2	s_3	s_4		s_1	s_2	s_3	s_4
1	18	2	7	13	32	4	16	8	3
2	13	3	4	9	33	22	17	4	6
3	24	3	11	12	34	21	17	4	13
4	10	4	2	6	35	20	17	7	8
5	16	4	2	12	36	19	17	11	6
6	11	5	4	3	37	4	17	11	7
7	17	5	4	6	38	12	17	11	15
8	12	5	11	9	39	15	18	4	9
9	23	5	11	12	40	17	18	4	15
10	23	6	7	8	41	12	18	7	4
11	14	8	2	9	42	15	18	8	11
12	22	8	7	4	43	6	18	11	13
13	21	8	11	4	44	8	19	2	9
14	10	9	8	2	45	13	19	4	2
15	22	9	11	9	46	5	19	8	3
16	3	10	4	15	47	6	19	8	11
17	24	10	7	8	48	24	19	11	5
18	21	10	8	4	49	6	20	2	10
19	12	10	8	15	50	13	20	4	10
20	17	10	11	6	51	24	21	2	7
21	3	11	4	12	52	14	21	8	13
22	9	11	4	13	53	10	22	8	13
23	9	11	7	4	54	7	22	8	14
24	11	12	4	10	55	15	23	8	5
25	20	12	7	15	56	9	23	11	4
26	17	12	11	11	57	20	24	4	8
27	21	13	4	14	58	16	24	4	14
28	11	14	8	7	59	20	24	4	14
29	6	14	8	13	60	23	24	7	3
30	20	15	7	13	61	14	24	8	10
31	12	16	2	10	62	16	24	11	12

Table 2. ME-CF generators with $L = 64$ and $J = 3$.

	k_1	k_2	k_3	q_1	q_2	q_3	s_1	s_2	s_3	k	$\lg \rho$	N_1
1	63	58	55	5	19	24	24	13	7	176	176	17
2	63	55	52	1	24	3	27	22	14	170	170	27
3	63	55	47	5	24	5	22	18	21	165	165	21
4	63	55	47	31	24	21	17	21	5	165	165	21
5	63	58	57	31	19	22	20	26	13	178	175	27
6	63	58	57	31	19	22	26	14	15	178	175	27
7	63	58	57	31	19	22	20	11	16	178	175	27
8	63	58	57	31	19	22	29	26	20	178	175	27
9	63	58	57	31	19	22	11	25	27	178	175	27
10	63	57	55	5	22	24	51	18	19	175	172	27

in the characteristic polynomial of the combined generator is significantly less than $k/2$, but still reasonably high.

Table 3. Full-period ME-CF generators with $L = 64$, $J = 4$, $k = 223$, and $N_1 = 49$.

	s_1	s_2	s_3	s_4		s_1	s_2	s_3	s_4
1	18	28	7	8	5	18	22	16	6
2	26	20	11	7	6	30	28	17	9
3	19	25	12	9	7	17	28	18	6
4	18	31	13	6	8	12	8	22	9

Table 6 lists 24 full-period ME-CF generators with $L = 64$, $J = 5$, $(k_1, k_2, k_3, k_4, k_5) = (63, 55, 52, 47, 41)$, $(q_1, q_2, q_3, q_4, q_5) = (1, 24, 3, 5, 3)$, $k = 258$, $\rho \approx 2^{258}$, and $N_1 = 103$. ME-CF generators with $L = 64$, $J = 5$, $(k_1, k_2, k_3, k_4, k_5) = (63, 57, 55, 52, 47)$, $(q_1, q_2, q_3, q_4, q_5) = (1, 7, 24, 3, 5)$, $k = 274$, $\rho \approx 2^{271}$, and $N_1 = 119$, are given in Table 7. As J increases, N_1 tends to approach $k/2$. With $J = 6$ or 7, one can probably obtain $N_1 \approx k/2$. However, as more components are added while making sure that $\lg \rho$ is close to k, one eventually comes up using polynomials P_j of relatively small degree k_j. Increasing J further then becomes less profitable.

One could also use polynomials P_j of larger degrees; e.g., use values of k_j near 128, having in mind (hypothetical) computers with 128-bit words. Still larger values of J would then be required in order to obtain N_1 near $k/2$.

4. Implementations

The procedure `lfsr113` in Figure 1 gives an implementation, in the language C, of the first ME-CF generator in Table 1, with $\rho \approx 2^{113}$. It uses the algorithm QuickTaus in Section 2.2 of [4], for each component of the combination. Before calling `lfsr113` for the first time, the variables `z1`, `z2`, `z3`, and `z4` must be initialized to any (random) integers larger than 1, 7, 15, and 127, respectively. In other words, the k_j most significant bits of z_j must be nonzero, for each j. (Note: this restriction also applies to the computer code given in [4], but was mistakenly not mentioned in that paper.) Ideally, the vector of initial seeds $(z_1, \ldots, z_j)$ would be drawn from a uniform distribution over the set of admissible values.

Figure 2 implements the first ME-CF generator in Table 6, whose period length is $\rho \approx 2^{258}$. The type "`unsigned long long`" refers to a 64-bit unsigned integer, available on 64-bit computers.

On a SUN UltraSparc 1, to generate 10 million (10^7) random numbers and add them up to print the sum, it took approximately 2.5 seconds with `lfsr113`, 3.1 seconds with `lfsr258`, and 0.2 seconds with the procedure `dummy` in Figure 1. For these speed comparisons, we used the `cc` compiler with the `-fast` option. We added the numbers and printed the sum to make sure that the optimizing compiler was not outsmarting us by skipping instructions after observing that the result was not used.

Table 4. ME-CF generators with $L = 64$, $J = 4$, $k = 233$, $\lg \rho = 230$, and $N_1 = 59$.

	s_1	s_2	s_3	s_4		s_1	s_2	s_3	s_4
1	18	10	23	11	47	43	16	31	18
2	26	10	13	11	48	38	23	37	18
3	48	17	30	11	49	46	25	39	18
4	27	20	9	11	50	47	4	26	19
5	46	22	9	11	51	33	7	27	19
6	23	29	24	11	52	18	11	17	19
7	25	29	13	11	53	43	11	37	19
8	34	29	9	11	54	5	14	13	19
9	50	7	38	12	55	53	20	27	19
10	15	8	19	12	56	24	25	25	19
11	44	22	16	12	57	30	25	27	19
12	6	23	29	12	58	34	29	41	19
13	16	5	22	13	59	18	5	36	20
14	11	10	25	13	60	15	11	18	20
15	18	11	40	13	61	52	11	34	20
16	19	16	30	13	62	5	22	10	20
17	45	23	24	13	63	9	22	10	20
18	17	7	9	14	64	16	23	38	20
19	52	11	20	14	65	17	23	26	20
20	52	22	30	14	66	40	23	37	20
21	25	23	26	14	67	46	23	5	20
22	27	7	19	15	68	6	28	27	20
23	25	11	13	15	69	25	28	33	20
24	6	26	31	15	70	5	32	26	20
25	19	28	25	15	71	13	7	37	21
26	38	28	37	15	72	26	8	41	21
27	53	28	18	15	73	37	10	43	21
28	50	29	32	15	74	38	10	11	21
29	17	32	41	15	75	30	13	39	21
30	39	8	12	16	76	38	16	43	21
31	53	13	33	16	77	9	17	32	21
32	12	5	13	17	78	34	25	17	21
33	16	5	11	17	79	38	26	41	21
34	25	7	32	17	80	8	28	31	21
35	54	10	36	17	81	19	29	12	21
36	45	11	29	17	82	37	32	27	21
37	30	20	18	17	83	27	8	5	22
38	39	20	43	17	84	8	10	29	22
39	19	22	22	17	85	41	10	25	22
40	50	23	25	17	86	50	13	4	22
41	11	26	19	17	87	55	13	37	22
42	19	26	11	17	88	50	17	36	22
43	13	29	40	17	89	39	26	29	22
44	46	32	29	17	90	55	26	23	22
45	20	4	31	18	91	13	28	16	22
46	5	10	33	18	92	51	32	10	22

TABLE 5. ME-CF generators with $L = 64$, $J = 4$, $k = 238$, $\lg \rho = 220$, and $N_1 = 71$.

	q_1	q_2	q_3	q_4	s_1	s_2	s_3	s_4
1	31	1	19	22	30	23	17	18
2	31	1	19	22	13	23	26	5
3	31	1	19	22	17	38	23	24
4	31	1	19	22	26	47	17	19
5	31	11	19	22	26	34	20	17
6	31	11	19	22	29	38	28	18

TABLE 6. ME-CF generators with $L = 64$, $J = 5$, $k = 258$, $\lg \rho = 258$, and $N_1 = 103$.

	s_1	s_2	s_3	s_4	s_5		s_1	s_2	s_3	s_4	s_5
1	10	5	29	23	8	13	26	5	31	14	13
2	12	5	11	16	15	14	36	5	32	16	8
3	17	5	16	6	7	15	36	5	32	21	8
4	17	5	19	16	14	16	39	5	19	6	8
5	18	5	37	7	3	17	43	5	14	20	15
6	19	5	31	15	13	18	44	5	14	15	15
7	20	5	11	13	6	19	44	5	29	6	13
8	22	5	17	10	11	20	44	5	34	25	9
9	23	5	37	13	7	21	45	5	16	21	8
10	24	5	7	16	8	22	51	5	28	3	12
11	26	5	22	4	9	23	53	5	26	16	8
12	26	5	26	13	12	24	54	5	28	13	3

TABLE 7. ME-CF generators with $L = 64$, $J = 5$, $k = 274$, $\lg \rho = 271$, and $N_1 = 119$.

	s_1	s_2	s_3	s_4	s_5		s_1	s_2	s_3	s_4	s_5
1	9	34	5	26	18	11	22	40	5	4	18
2	9	32	5	31	6	12	22	19	5	14	19
3	9	25	5	37	22	13	22	41	5	16	6
4	10	24	5	7	12	14	22	16	5	32	4
5	12	17	5	14	8	15	26	9	5	11	14
6	12	40	5	16	22	16	26	19	5	29	3
7	12	26	5	34	23	17	44	20	5	8	6
8	17	27	5	13	9	18	44	31	5	22	14
9	17	8	5	37	19	19	53	8	5	23	17
10	20	41	5	14	6	20	53	12	5	31	18

```
unsigned long z1, z2, z3, z4;

double lfsr113 ()
   {                    /* Generates numbers between 0 and 1. */
   unsigned long b;
   b  = (((z1 <<  6) ^ z1) >> 13);
   z1 = (((z1 & 4294967294) << 18) ^ b);
   b  = (((z2 <<  2) ^ z2) >> 27);
   z2 = (((z2 & 4294967288) <<  2) ^ b);
   b  = (((z3 << 13) ^ z3) >> 21);
   z3 = (((z3 & 4294967280) <<  7) ^ b);
   b  = (((z4 <<  3) ^ z4) >> 12);
   z4 = (((z4 & 4294967168) << 13) ^ b);
   return ((z1 ^ z2 ^ z3 ^ z4) * 2.3283064365387e-10);
   }

double dummy ()
   {
   return 0.5
   }
```

Figure 1. A 32-bit combined LFSR generator with 4 components.

```
unsigned long long z1, z2, z3, z4, z5;

double lfsr258 ()
   {                    /* Generates numbers between 0 and 1. */
   unsigned long long b;
   b  = (((z1 <<  1) ^ z1) >> 53);
   z1 = (((z1 & 18446744073709551614) << 10) ^ b);
   b  = (((z2 << 24) ^ z2) >> 50);
   z2 = (((z2 & 18446744073709551104) <<  5) ^ b);
   b  = (((z3 <<  3) ^ z3) >> 23);
   z3 = (((z3 & 18446744073709547520) << 29) ^ b);
   b  = (((z4 <<  5) ^ z4) >> 24);
   z4 = (((z4 & 18446744073709420544) << 23) ^ b);
   b  = (((z5 <<  3) ^ z5) >> 33);
   z5 = (((z5 & 18446744073701163008) <<  8) ^ b);
   return ((z1 ^ z2 ^ z3 ^ z4 ^ z5) * 5.4210108624275221e-20);
   }
```

Figure 2. A 64-bit combined LFSR generator with 5 components.

References

[1] A. Compagner, *The hierarchy of correlations in random binary sequences*, Journal of Statistical Physics **63** (1991), 883–896.

[2] ———, *Operational conditions for random number generation*, Physical Review E **52** (1995), no. 5-B, 5634–5645.

[3] P. L'Ecuyer, *Uniform random number generation*, Annals of Operations Research **53** (1994), 77–120.

[4] ———, *Maximally equidistributed combined Tausworthe generators*, Mathematics of Computation **65** (1996), no. 213, 203–213.

[5] ———, *Random number generation*, Handbook on Simulation (Jerry Banks, ed.), Wiley, 1998, To appear.

[6] M. Matsumoto and Y. Kurita, *Twisted GFSR generators II*, ACM Transactions on Modeling and Computer Simulation **4** (1994), no. 3, 254–266.

[7] ———, *Strong deviations from randomness in m-sequences based on trinomials*, ACM Transactions on Modeling and Computer Simulation **6** (1996), no. 2, 99–106.

[8] H. Niederreiter, *Random number generation and quasi-Monte Carlo methods*, SIAM CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 63, SIAM, Philadelphia, 1992.

[9] S. Tezuka and P. L'Ecuyer, *Efficient and portable combined Tausworthe random number generators*, ACM Transactions on Modeling and Computer Simulation **1** (1991), no. 2, 99–112.

[10] D. Wang and A. Compagner, *On the use of reducible polynomials as random number generators*, Mathematics of Computation **60** (1993), 363–374.

Département d'Informatique et de Recherche Opérationnelle, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, H3C 3J7, Canada

E-mail address: `lecuyer@iro.umontreal.ca`