

On the Lattice Structure of MIXMAX Random Number Generators

Pierre L'Ecuyer



Joint work with **Paul Wambergue**, **Erwan Bourceret**, and **Marc-Antoine Savard**

MCQMC, Rennes, July 2018

MIXMAX Generators [Akopov, Savvidy, et al. (1991)]

Output vector $\mathbf{u}_i \in \mathbb{R}^k$ follow the recurrence

$$\mathbf{u}_i = \mathbf{A}\mathbf{u}_{i-1} \bmod 1$$

where $\mathbf{u}_0 \in [0, 1)^k$ and \mathbf{A} is a $k \times k$ matrix such that:

- (1) $\det(\mathbf{A}) = 1$, so the linear transformation preserves volume;
- (2) the eigenvalues $\lambda_1, \dots, \lambda_k$ of \mathbf{A} are away from the unit circle.

Then, for appropriate \mathbf{A} , distance between trajectories that start from very close will diverge exponentially fast, as $\Theta(e^{hi})$ in i steps, where h is the entropy:

$$h = \sum_{j=1}^k \mathbb{I}[|\lambda_j| > 1] \cdot \log |\lambda_j|.$$

For almost all (irrational) \mathbf{u}_0 , the trajectory is aperiodic and space-filling in $[0, 1)^k$.

Original MIXMAX proposal:

$$\mathbf{A} = \mathbf{A}(k, d) = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 3 + d & 2 & 1 & \cdots & 1 & 1 \\ 1 & 4 & 3 & 2 & \cdots & 1 & 1 \\ \vdots & & & & \ddots & & \\ 1 & k & k - 1 & k - 2 & \cdots & 3 & 2 \end{pmatrix}.$$

Savvidy (2015) provides lower bounds on h that depend only on k for $\mathbf{A} = \mathbf{A}(k, d)$, and shows that h is large, e.g., much larger than for the AWC and SWB generators of Marsaglia and Zaman (1991).

However, all of this is for irrational state vectors \mathbf{u}_i , which cannot be implemented exactly.

The authors proposed to approximate the irrational recurrence by a rational one as follows. Take a large prime m and define

$$\mathbf{x}_i = \mathbf{A}\mathbf{x}_{i-1} \bmod m$$

in which $\mathbf{x}_i = (x_{i,0}, \dots, x_{i,k-1})^t \in \mathbb{Z}_m^k$ and

$$\mathbf{u}_i = (u_{i,0}, \dots, u_{i,k-1}) = \mathbf{x}_i/m \in [0, 1)^k.$$

This is a [matrix LCG](#) (Niederreiter 1986); gives a [periodic](#) recurrence! If $\det(\mathbf{A}) = 1$, the maximal period is $(m^k - 1)/(m - 1)$, i.e., $m - 1$ times shorter than general matrix LCGs.

Savvidy provides parameters (k, d) giving period $\rho = (m^k - 1)/(m - 1)$ for $m = 2^{61} - 1$ with k from 8 to 3150. Even for $k = 8$, this already gives $\rho \approx 2^{427}$. [Long enough period!](#)

He also provides an [efficient implementation](#) that uses only $2k$ additions and one multiplication by d to compute the next vector \mathbf{x}_i at each step. [Fast!](#)

MIXMAX became popular recently. Part of the ROOT library [used at CERN](#), Geneva.

In July 2016, I was at CERN, in a workshop devoted to the MIXMAX. Fred James was telling us: “We have a proof that this RNG passes all statistical tests; this is fantastic.” I was not fully convinced.

Other variants, with more parameters and flexibility

MIXMAX- (m, k, d, c) and MIXMAX- (m, k, d, c, b) :

$$\mathbf{A} = \mathbf{A}(k, d, c) = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & 1 & 1 & \cdots & 1 & 1 \\ 1 & c+2+d & 2 & 1 & \cdots & 1 & 1 \\ 1 & 2c+2 & c+2 & 2 & \cdots & 1 & 1 \\ 1 & 3c+2 & 2c+2 & c+2 & \cdots & 1 & 1 \\ & & & \cdots & & & \\ 1 & (k-2)c+2 & (k-3)c+2 & (k-4)c+2 & \cdots & c+2 & 2 \end{pmatrix}$$

$$\mathbf{A} = \mathbf{A}(k, d, c, b) = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 3c+d+b & 2 & 1 & \cdots & 1 & 1 \\ 1 & 4c+b & 3c+b & 2 & \cdots & 1 & 1 \\ 1 & 5c+b & 4c+b & 3c+b & \cdots & 1 & 1 \\ & & & \cdots & & & \\ 1 & kc+b & (k-1)c+b & (k-2)c+b & \cdots & 3c+b & 2 \end{pmatrix}$$

Lattice structure of matrix LCGs

Define successive output values of matrix LCGs as follows:

$$\begin{aligned} \mathbf{u}_0 &= (u_0, u_1, \dots, u_{k-1}) \\ \mathbf{u}_1 &= (u_k, u_{k+1}, \dots, u_{2k-1}) \\ \mathbf{u}_2 &= (u_{2k}, u_{2k+1}, \dots, u_{3k-1}), \quad \text{etc.} \end{aligned}$$

For any fixed $s > 0$, let

$$\Psi_s = \{(u_0, u_1, \dots, u_{s-1}) \in [0, 1)^s \mid \mathbf{x}_0 \in \mathbb{Z}_m^k\},$$

the set of vectors of s successive values obtained from all possible initial states. We want this set to cover $[0, 1)^s$ very evenly.

More general: for any finite set $I = \{i_1, \dots, i_s\}$ where $0 \leq i_1 < \dots < i_s$, let

$$\Psi_s(I) = \{(u_{i_1}, \dots, u_{i_s}) \in [0, 1)^s \mid \mathbf{x}_0 \in \mathbb{Z}_m^k\}.$$

Want this set to cover $[0, 1)^s$ very evenly for all I in some large family, e.g., $i_s \leq i_*$ and $s \leq s^*$.

Lattice structure

It is known that for each I ,

$$\Psi_s(I) = L_s(I) \cap [0, 1)^s$$

where $L_s(I)$ is a lattice in \mathbb{R}^s .

It means that all the points are in equidistant hyperplanes at distance $d_s(I) = 1/\ell_s(I)$ apart, where $\ell_s(I) =$ length of shortest vector in dual lattice.

Let $d_s^*(n)$ be the best (smallest) distance for a known lattice with density $n = \min(m^k, m^s)$ in s dimensions. Normalized measure: $S_s(I) = d_s^*(n)/d_s(I) \in (0, 1]$. Want it close to 1.

Example of worst-case figure of merit: For $s' > 0$, let

$$M_{s'} = \min_{I \subseteq \{1, \dots, s'\}} S_s(I).$$

If s' is very large, there are so many subsets that some of them must be bad. But for s' not too large, **most** MRGs and matrix LCGs have a good $M_{s'}$.

Example: MRG of order k

Multiple recursive generators (MRG):

$$x_i = (a_1 x_{i-1} + \cdots + a_k x_{i-k}) \bmod m, \quad u_i = x_i / m.$$

Equivalent to matrix LCG with

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ a_k & a_{k-1} & \cdots & a_1 \end{pmatrix}^k$$

For $k = 7$ and $m = 2^{63} - 52425$ (a prime), we searched random parameters $\mathbf{a} = (a_1, \dots, a_k)$ that gave full period $\rho = m^k - 1 \approx 2^{441}$, and computed M_{10} for each. Within our limited time budget, we obtained 7041 such vectors and the values of M_{10} ranged from 0.53858 (the best) to 0.14970 (the worst). Thus, values of M_{10} smaller than say 0.1 are very rare.

How to compute $\ell_s(I)$?

$$\mathbf{V} = \begin{pmatrix} \mathbf{I}/m & \mathbf{A}^t/m & \cdots & (\mathbf{A}^{\nu-1})^t/m & ([\mathbf{A}^\nu]_r)^t/m \\ \mathbf{0} & \mathbf{I} & & & \mathbf{0} \\ & & \ddots & & \\ & & & \mathbf{I} & \\ \mathbf{0} & & & & \mathbf{I} \end{pmatrix}$$

and

$$\mathbf{W} = \begin{pmatrix} m \cdot \mathbf{I} & \mathbf{0} & & \mathbf{0} \\ -\mathbf{A} & \mathbf{I} & & \\ \vdots & & \ddots & \\ -\mathbf{A}^{\nu-1} & & & \mathbf{I} \\ -[\mathbf{A}^\nu]_r & \mathbf{0} & & \mathbf{I} \end{pmatrix}$$

The rows of \mathbf{V} form a basis of L_s and the rows of \mathbf{W} form a basis of the dual lattice L_s^* . For a basis of $L_s(I)$, select the appropriate columns of \mathbf{V} , and invert this basis modulo 1 to get a basis for $L_s^*(I)$. Then compute the length $\ell_s(I)$ of a shortest nonzero vector in $L_s^*(I)$, using BB.

Lattice structure of MIXMAX (m, k, d)

$$\mathbf{W} = \begin{pmatrix} & & \mathbf{m} \cdot \mathbf{I} & & & & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ -1 & -1 & -1 & \dots & -1 & & & & & \\ -1 & -2 & -1 & \dots & -1 & & & & & \\ -1 & -3-d & -2 & \dots & -1 & & \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ -1 & -4 & -3 & \dots & -1 & & & & & \\ & & & \dots & & & & & & \\ -1 & -k & -(k-1) & \dots & -2 & & & & & \\ & & -\mathbf{A}^2 & & & & \mathbf{0} & \mathbf{I} & & \\ & & \vdots & & & & \vdots & & \ddots & \\ & & -[\mathbf{A}^r]_r & & & & \mathbf{0} & & & \mathbf{I} \end{pmatrix}.$$

The other MIXMAX are similar, with a slightly different $-\mathbf{A}$.

Any integer linear combination of rows of \mathbf{W} belongs to the dual lattice L_s^* .

In particular, if $k \geq 2$, for all the MIXMAX variants, we have

$$\mathbf{w}_{k+1} = (-1, -1, -1, \dots, -1, 1, 0, \dots, 0)$$

$$\mathbf{w}_{k+2} = (-1, -2, -1, \dots, -1, 0, 1, \dots, 0)$$

↑ coordinate $k + 1$

so

$$\mathbf{w} = \mathbf{w}_{k+1} - \mathbf{w}_{k+2} = (0, 1, 0, \dots, 0, 1, -1, 0, \dots, 0) \in L_s^* \text{ for } s \geq k + 2.$$

↑ coordinate $k + 1$

This vector has Euclidean length $\sqrt{3}$.

Its presence implies that the successive output values satisfy $(u_1 + u_k - u_{k+1}) \bmod 1 = 0$.

That is, $u_1 + u_k - u_{k+1} = 0$ or 1 , because $0 \leq u_i < 1$ for all i .

So when $I := \{1, k, k + 1\} \subseteq I'$, $\Psi_3(I')$ is covered by only two planes, $1/\sqrt{3}$ apart.

An example with $k = 8$

MIXMAX- (m, k, d, c) from Savvidy (2017), with $m = 2^{61} - 1$, $k = 8$, $d = 0$, and $c = 2^{53} + 1$.

Then, whenever $\{1, 8, 9\} = I \subseteq I'$, the points of $\Psi_3(I')$ are in only two parallel planes.

This gives $S_s(I) = 6.69 \times 10^{-19}$ for this I , and then $M_{s'} \leq 6.69 \times 10^{-19}$ for all $s' \geq k + 2$.

We applied the following empirical statistical tests to this generator.

Collision test: Generate n points in s dimensions, using all coordinates of output vectors.

Partition the cube $[0, 1)^s$ in d^s cubic cells.

Count the number C of collisions (a point falling in an occupied cell).

C should be approx. Poisson with mean $\lambda = n^2/(2d^s)$.

Repeat N times and count the total number of collisions, then the p -value.

With $s = 16$, $d = 8$, $n = 4 \times 10^7$, $N = 10$, we get $p \approx 6 \times 10^{-206}$.

We expect about 28 collisions and we get 314.

If we take only the first three values of each block and $s = 6$, with $d = 128$, same n and N , we get $p < 10^{-300}$. We expect about 1,818 collisions and we get 116,218.

Birthday spacings test: With $s = 16$, $d = 16$, $n = 3 \times 10^7$, $N = 10$, we get $p < 10^{-300}$.

Skipping coordinates

To get rid of the bad structure, one can skip some coordinates of each vector \mathbf{u}_i . Lüscher (1994) already proposed this for the AWC/SWB generators.

For example, if we skip the second coordinate of each vector \mathbf{u}_i , we no longer have the relationship $u_1 + u_k - u_{k+1} = 0$ or 1 .

But other relationships can be found between the remaining coordinates!

For example, for the MIXMAX- (m, k, d) , if $k \geq 6$, then

$$\mathbf{w} = \mathbf{w}_{k+4} - 2\mathbf{w}_{k+5} + \mathbf{w}_{k+6} = (0, \dots, 0, \underset{\uparrow \text{coord. } 6}{-1}, 0, \dots, 0, 1, -2, 1, 0, \dots)$$

from which we find $-u_5 + u_{k+3} - 2u_{k+4} + u_{k+5} = q$ for $q \in \{-2, -1, 0, 1\}$. It follows that whenever $\{5, k+3, k+4, k+5\} \subseteq I$, $\Psi_s(I)$ is contained in at most 4 equidistant parallel hyperplanes at distance $1/\sqrt{7}$ apart. Even if we skip the first three coordinates of each block \mathbf{u}_i , we still have this structure.

MIXMAX- (m, k, d, c)

$$\mathbf{A} = \mathbf{A}(k, d, c) = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & 1 & 1 & \cdots & 1 & 1 \\ 1 & c+2+d & 2 & 1 & \cdots & 1 & 1 \\ 1 & 2c+2 & c+2 & 2 & \cdots & 1 & 1 \\ 1 & 3c+2 & 2c+2 & c+2 & \cdots & 1 & 1 \\ & & & \cdots & & & \\ 1 & (k-2)c+2 & (k-3)c+2 & (k-4)c+2 & \cdots & c+2 & 2 \end{pmatrix}$$

Skipping coordinates for MIXMAX- (m, k, d, c) with $c \geq 1$.

Suppose $m = qc + r$ where $q > 0$ and $|r|$ are small integers. Then

$$\begin{aligned} \mathbf{w} &= q(\mathbf{w}_{k+4} - 2\mathbf{w}_{k+5} + \mathbf{w}_{k+6}) \bmod m \\ &= q(0, \dots, 0, 1 - c, -1, 0, \dots, 0, 1, -2, 1, 0, \dots) \bmod m \\ &= (0, \dots, 0, q + r, -q, 0, \dots, 0, q, -2q, q, 0, \dots), \end{aligned}$$

in which $q + r$ is at position 5 and $-2q$ is at position $k + 5$.

We thus have $((q + r)u_4 - qu_5 + qu_{k+3} - 2qu_{k+4} + qu_{k+5}) \bmod 1 = 0$.

So if $\{4, 5, k + 3, k + 4, k + 5\} \subseteq I$, $\Psi_s(I)$ is covered by at most $5q + |q + r| - 1$ equidistant parallel hyperplanes, at distance $1/\ell$ apart, where $\ell^2 = 7q^2 + (q + r)^2$.

Example with $k = 8$ again

MIXMAX- (m, k, d, c) from Savvidy (2017), with $m = 2^{61} - 1$, $k = 8$, $d = 0$, and $c = 2^{53} + 1$.

Here, $m = 256c - 257$, so $q = 256$ and $r = -257$.

Therefore, if $\{4, 5, 11, 12, 13\} \subseteq I$, then $\Psi_s(I)$ is covered by at most $5q + |q + r| - 1 = 1280$ equidistant parallel hyperplanes at distance $1/\ell = 1/\sqrt{7q^2 + (q + r)^2} \approx 1/677.3$, for $s = 5$.

This ℓ is in fact the exact $\ell_s(I)$ and it gives $S_s(I) \approx 2.3859 \times 10^{-16}$.

So even by skipping the first three coordinates of each output vector, we still have a very bad structure.

By searching at random for bad MRGs with a similar m and k , it is extremely rare to find such a bad one!

Birthday spacings test: Skip first 3 values of each block and keep the next 5. Apply the test with $t = 10$, $d = 64$, $n = 10^7$, $N = 10$. We get $p < 10^{-300}$. We have $\mathbb{E}[C] = 2168$ and get $C = 4220$.

Conclusion

- ▶ The MIXMAX is a fast generator with a very large period.
It may look very good and safe at first sight.
It passed the Crush batteries of TestU01.
But we showed here that it has significant weaknesses that show up in statistical tests.
- ▶ We have seen this type of story several times before, with other very fast generators with large periods.
- ▶ Bottom line: It is important to analyze and understand (theoretically) the structure of the points produced by RNGs.

Some references

- Afflerbach, L., and H. Grothe. 1988. "The Lattice Structure of Pseudo-Random Vectors Generated by Matrix Generators". *Journal of Computational and Applied Mathematics* 23:127–131.
- Couture, R., and P. L'Ecuyer. 1994. "On the Lattice Structure of Certain Linear Congruential Sequences Related to AWC/SWB Generators". *Mathematics of Computation* 62 (206): 798–808.
- L'Ecuyer, P. 1997. "Bad Lattice Structures for Vectors of Non-Successive Values Produced by Some Linear Recurrences". *INFORMS Journal on Computing* 9 (1): 57–60.
- L'Ecuyer, P. 1999. "Good Parameters and Implementations for Combined Multiple Recursive Random Number Generators". *Operations Research* 47 (1): 159–164.
- L'Ecuyer, P. 2017. "History of Uniform Random Number Generation". In *Proceedings of the 2017 Winter Simulation Conference*, 202–230: IEEE Press.
- L'Ecuyer, P., and R. Couture. 1997. "An Implementation of the Lattice and Spectral Tests for Multiple Recursive Linear Random Number Generators". *INFORMS Journal on Computing* 9 (2): 206–217.

- L'Ecuyer, P., and R. Simard. 2007, August. "TestU01: A C Library for Empirical Testing of 17 Random Number Generators". *ACM Transactions on Mathematical Software* 33 (4): Article 22.
- L'Ecuyer, P., and R. Simard. 2014. "On the Lattice Structure of a Special Class of Multiple Recursive Random Number Generators". *INFORMS Journal on Computing* 26 (2): 449–460.
- Lüscher, M. 1994. "A Portable High-Quality Random Number Generator for Lattice Field Theory Simulations". *Computer Physics Communications* 79:100–110.
- Marsaglia, G., and A. Zaman. 1991. "A New Class of Random Number Generators". *The Annals of Applied Probability* 1:462–480.
- Savvidy, K. G. 2015. "The MIXMAX Random Number Generator". *Computer Physics Communications* 196:161–165.
- Savvidy, K. G. 2017. *MIXMAX manual*. see <https://www.hepforge.org/archive/mixmax/MANUAL.pdf>.
- Tezuka, S., P. L'Ecuyer, and R. Couture. 1993. "On the Add-with-Carry and Subtract-with-Borrow Random Number Generators". *ACM Transactions of Modeling and Computer Simulation* 3 (4): 315–331.