On the Distribution of Counter-Dependent Nonlinear Congruential Pseudorandom Number Generators in Residue Rings

Edwin D. El-Mahassni
Department of Computing, Macquarie University
North Ryde, NSW, 2109
edwinelm@ics.mq.edu.au

Domingo Gomez
Johann Radon Institute for
Computational and Applied Mathematics
Altenberger Straße 69, A-4040 Linz, Austria.
Domingo.Gomez@ricam.oeaw.ac.at

Abstract

Nonlinear congruential pseudorandom number generators can have unexpectedly short periods. Shamir and Tsaban introduced the class of counter-dependent generators which admit much longer periods. In this paper, using a technique developed by Niederreiter and Shparlinski, we present discrepancy bounds for sequences of s-tuples of successive pseudorandom numbers generated by counter-dependent generators modulo a composite M.

1 Introduction

In this paper we study some distribution properties of *counter-dependent nonlinear congruential pseudorandom number generators* introduced by [17] and defined by a recurrence congruence modulo an integer M of the form

$$u_{n+1} = f(u_n, n) \pmod M, \quad 0 \le u_n \le M-1, \qquad n = 0, 1, \dots , \tag{1}$$

with some *initial value* u_0, where $f(X,Y) \in \mathbb{Z}_M[X,Y]$ is a polynomial over the residue ring $\mathbb{Z}_M = \mathbb{Z}/M\mathbb{Z}$.

It is obvious that the sequence (1) eventually becomes periodic with some period $t \le M^2$. Throughout this paper we assume that this sequence is *purely periodic*, that is, $u_n = u_{n+t}$ beginning with $n = 0$, otherwise we consider a shift of the original sequence.

In the case that $f(X,Y) = h(X) \in \mathbb{Z}_M[X]$ does not depend on the second variable we get the well-studied *nonlinear congruential pseudorandom number generators*, see [4, 6, 8, 13] for the distribution of the elements and for the distribution of powers in prime fields see [15]. However, in this case the period t is at most M and it is possible that the generated sequences have unexpectedly short period as it is noted in [17]. In the case that $f(X,Y) = g(X) + Y \in \mathbb{Z}_M[X,Y]$ we get the *counter-assisted nonlinear congruential pseudorandom number generators* defined in [17]. These generators are special *nonlinear congruential pseudorandom number generators of order* 2 defined by

$$u_{n+1} = f(u_n, u_{n-1}) \pmod M, \quad 0 \le u_n \le M-1, \qquad n = 1, 2, \dots$$

where $f(X,Y) = g(X) - g(Y) + X + 1$ with some special initial values u_0 and u_1 satisfying $u_1 = g(u_0) + 1$. The case where the order is non trivial and $M = p$ is a prime, has been analyzed in [7, 9, 18].

Distribution and structural properties of general counter-dependent nonlinear congruential generators over finite fields have first been analyzed in [5]. Here, we establish results about the distribution about residue rings using a technique introduced in [13].

The first Section is devoted to introduce some notations and stating known theorems. In Section 3 we prove results about the distribution of the points

$$\left(\frac{u_n}{M}, \dots, \frac{u_{n+s-1}}{M}\right) \tag{2}$$

in the s-dimensional unit cube $[0,1)^s$ in terms of a discrepancy bound, where n runs through a part of the period, $n = 0, \dots, N-1$, $1 \le N \le t$.

A uniform distribution of these points, i.e., a low discrepancy, is a desirable feature for pseudorandom numbers in quasi-Monte Carlo methods, see e.g. [11, 12, 14, 19].

Finally, in Section 4, we show how for some M, we obtain improvements on these distribution results.

2 Definitions and Auxiliary Results

Given an integer M, we define $\omega(M)$ to be the number of distinct prime divisors of M and $\tau(M)$ as the number of divisors of M. The first lemma follows directly from Theorem 317 in [10].

Lemma 1. *For every sufficiently large M, the bound*

$$\tau(M) = M^{O(1/\log\log M)}$$

holds.

This bound holds for suffiently large M, but for most values of M we can obtain improvements due to Hardy and Ramanujan (see [10]).

Lemma 2. *The bound*

$$\tau(M) \le (\log M)^2$$

holds for all, except $o(X)$ numbers when $1 \le M \le X$.

For a sequence of N points

$$\Gamma = (\gamma_{1,n}, \ldots, \gamma_{s,n})_{n=1}^{N} \tag{3}$$

of the half-open interval $[0,1)^s$, denote by Δ_Γ its *discrepancy*, that is,

$$\Delta_\Gamma = \sup_{B\subseteq[0,1)^s} \left| \frac{T_\Gamma(B)}{N} - |B| \right|,$$

where $T_\Gamma(B)$ is the number of points of the sequence Γ which hit the box

$$B = [\alpha_1, \beta_1) \times \ldots \times [\alpha_s, \beta_s) \subseteq [0,1)^s$$

and the supremum is taken over all such boxes. For an integer vector $\mathbf{a} = (a_1, \ldots, a_s) \in \mathbb{Z}^s$ we put

$$|\mathbf{a}| = \max_{i=1,\ldots,s} |a_i|, \qquad r(\mathbf{a}) = \prod_{i=1}^{s} \max\{|a_i|, 1\}. \tag{4}$$

Also, denote by $\gcd(\alpha_0, \ldots, \alpha_{N-1})$ the greatest common divisor of the integers $\alpha_0, \ldots, \alpha_{N-1}$. We need the *Erdös–Turán–Koksma inequality* (see Theorem 1.21 of [3]) for the discrepancy of a sequence of points of the s-dimensional unit cube, which we present in the following form.

Lemma 3. *There exists a constant $C_s > 0$ depending only on the dimension s such that, for any integer $L \geq 1$, for the discrepancy of a sequence of points* (3) *the bound*

$$\Delta_\Gamma < C_s \left(\frac{1}{L} + \frac{1}{N} \sum_{0<|\mathbf{a}|\leq L} \frac{1}{r(\mathbf{a})} \left| \sum_{n=1}^{N} \exp\left(2\pi i \sum_{j=1}^{s} a_j \gamma_{j,n} \right) \right| \right)$$

holds, where $|\mathbf{a}|$, $r(\mathbf{a})$ are defined by (4) *and the sum is taken over all integer vectors*

$$\mathbf{a} = (a_1, \ldots, a_s) \in \mathbb{Z}^s$$

with $0 < |\mathbf{a}| \leq L$.

The currently best value of C_s is given in [2]. We put

$$\mathbf{e}_M(z) = \exp(2\pi i z/M).$$

For a polynomial $f(X,Y) \in \mathbb{Z}_M[X,Y]$ of total degree d we define the sequence of polynomials $f_k(X,Y) \in \mathbb{Z}_M[X,Y]$ by the recurrence relation

$$f_{k+1}(X,Y) = f\left(f_k(X,Y), Y+k\right), \qquad k = 0, 1, \ldots, \tag{5}$$

where $f_0(X,Y) = X$. It is clear that $\deg f_k \leq d^k$ and that

$$u_{n+k} = f_k\left(u_n, n\right).$$

This allows us to state the following Lemma:

Lemma 4. *Let $f(X,Y) \in \mathbb{Z}_M[X,Y]$ be a polynomial of local degree in X of value $d_p \geq 2$ modulo every prime divisor p of M and $f_k(X,Y)$ is defined as in* (5). *Then the local degree in X of $f_k^{(p)}(X,Y) = f_k(X,Y) \pmod{p}$ equals d_p^k, $k = 0, 1, \ldots$.*

Proof. It is trivial to see that

$$f_k^{(p)}(X,Y) = f^{(p)}(f_{k-1}^{(p)}(X,Y), Y+k-1) \pmod{p}.$$

So, using Lemma 3 of [5], we arrive at the desired result □

The following Lemma is the 2-dimensional version of Theorem 2.6 in [1] in a slightly weaker form.

Lemma 5. *Let $f(X,Y)$ be a polynomial with integer coefficients with the gcd of all of them, except the constant term, is one and total degree d then the bound*

$$\left| \sum_{x,y=1}^{M} \mathbf{e}_M(f(x,y)) \right| \leq e^{14d} 3^{2\omega(M)} (\tau(M)) M^{2-1/d}$$

holds.

This now allows us to state and prove the following Lemma.

Lemma 6. *Let $f(X,Y)$ be a polynomial with integer coefficients and total degree d. Then the bound*

$$\left| \sum_{x,y=1}^{M} \mathbf{e}_M(f(x,y)) \right| \leq e^{14d} (\tau(M/G))^5 M^{2-1/d} G^{1/d}$$

holds, where G is the gcd of all the coefficients of f except the constant term.

Proof. Let $f_G(x,y) = (f(x,y) - f(0,0))/G$ and $m = M/G$. Then,

$$\left| \sum_{x,y=1}^{M} \mathbf{e}_M(f(x,y)) \right| = \left| \sum_{x,y=1}^{M} \mathbf{e}_M(f(x,y) - f(0,0)) \right| = G^2 \left| \sum_{x,y=1}^{m} \mathbf{e}_m(f_G(x,y)) \right|.$$

Now $f_G(X,Y)$ satisfies the conditions in Lemma 5, so:

$$G^2 \left| \sum_{x,y=1}^{m} \mathbf{e}_m(f_G(x,y)) \right| \leq G^2 e^{14d} 3^{2\omega(m)} \tau(m) (m)^{2-1/d}$$

and noting $2^{\omega(m)} \leq \tau(m)$, the result follows. □

Now, we are going to introduce some results about the sequence $f_k(X,Y)$ that we will have to use in the proofs.

Lemma 7. *Let $f(X,Y) \in \mathbb{Z}_M[X,Y]$ be a polynomial of local degree in X, $d_p \geq 2$ modulo every prime divisor p of M and let*

$$\sum_{j=0}^{s-1} a_j \left(f_{k+j}(X,Y) - f_{l+j}(X,Y)\right) = \sum_{i_1=0}^{D_1}\sum_{i_2=0}^{D_2} B_{i_1 i_2} X^i Y^j.$$

Then, for any $k \neq l$, the equality

$$\gcd(B_{10}, B_{01}, \ldots, B_{D_1 D_2}, M) = \gcd(a_0, \ldots, a_{s-1}, M).$$

holds.

Proof. The main ideas of the proof come from Lemma 5 in [4]. We put $A_j = a_j/G$, $j = 0, \ldots, s-1$ and $m = M/G$, where $G = \gcd(a_0, \ldots, a_{s-1}, M)$. In particular,

$$\gcd(A_0, \ldots, A_{s-1}, m) = 1. \tag{6}$$

It is enough to show that

$$H(X,Y) = \sum_{j=0}^{s-1} A_j \left(f_{k+j}(X,Y) - f_{l+j}(X,Y)\right)$$

is not a constant polynomial modulo any prime $p|m$. We take $f^{(p)}$ to be the reduction of f modulo p. By our assumption, the local degree of X in $f^{(p)}$ is $d_p \geq 2$. Denote by $f_k^{(p)}$ as in Lemma 4 and $H^{(p)}(X,Y)$ as $H(X,Y) \mod p$. Thus,

$$H^{(p)}(X,Y) = \sum_{j=0}^{s-1} A_j \left(f_{k+j}^{(p)}(X,Y) - f_{l+j}^{(p)}(X,Y)\right) \pmod{p}.$$

Let h be the largest $j = 1, \ldots, s$ with $\gcd(A_j, p) = 1$ (we see from (6) that such h exists). Then, by Lemma 4, for $k > l$ the polynomial $H^{(p)}(X,Y)$ has local degree in X exactly d_p^{k+h}, finishing the proof. □

3 Discrepancy Bound

Let the sequence (u_n) generated by (1) be purely periodic with an arbitrary period t. For an integer vector $\mathbf{a} = (a_0, \dots, a_{s-1}) \in \mathbb{Z}^s$ we introduce the exponential sum

$$S_{\mathbf{a}}(N) = \sum_{n=0}^{N-1} \mathbf{e}_M \left(\sum_{j=0}^{s-1} a_j u_{n+j} \right).$$

Theorem 8. *Let the sequence (u_n), given by (1) with a polynomial $f(X,Y) \in \mathbb{Z}_M[X,Y]$ with $f(X,Y)$ of total degree d and local degree in X, at least 2 modulo every prime divisor p of M, be purely periodic with period t, and $t \ge N \ge 1$, then the bound*

$$\max_{\gcd(a_0,\dots,a_{s-1},M)=G} |S_{\mathbf{a}}(N)| = O\left(N^{1/2} M (\log\log\log(M/G))^{-1/2}\right)$$

holds, where the implied constant depends only on s and d.

Proof. Select any $\mathbf{a} = (a_0, \dots, a_{s-1}) \in \mathbb{Z}^s$ with $\gcd(a_0, \dots, a_{s-1}, M) = G$. It is obvious that for any integer $k \ge 0$ we have

$$\left| S_{\mathbf{a}}(N) - \sum_{n=0}^{N-1} \mathbf{e}_M \left(\sum_{j=0}^{s-1} a_j u_{n+k+j} \right) \right| \le 2k.$$

Therefore, for any integer $K \ge 1$,

$$K |S_{\mathbf{a}}(N)| \le W + K^2,$$

where

$$W = \left| \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} \mathbf{e}_M \left(\sum_{j=0}^{s-1} a_j u_{n+k+j} \right) \right| \le \sum_{n=0}^{N-1} \left| \sum_{k=0}^{K-1} \mathbf{e}_M \left(\sum_{j=0}^{s-1} a_j u_{n+k+j} \right) \right|.$$

Accordingly, we obtain

$$\begin{aligned} W^2 &\le N \sum_{n=0}^{N-1} \left| \sum_{k=0}^{K-1} \mathbf{e}_M \left(\sum_{j=0}^{s-1} a_j f_{k+j}(u_n, n) \right) \right|^2 \\ &\le N \sum_{x,y=1}^{M} \left| \sum_{k=0}^{K-1} \mathbf{e}_M \left(\sum_{j=0}^{s-1} a_j f_{k+j}(x, y) \right) \right|^2 \\ &= N \sum_{k=0}^{K-1} \sum_{l=0}^{K-1} \sum_{x,y=1}^{M} \mathbf{e}_M \left(\sum_{j=0}^{s-1} a_j \left(f_{k+j}(x,y) - f_{l+j}(x,y) \right) \right). \end{aligned}$$

If $k = l$, then the inner sum is trivially equal to M^2. There are K such sums. Otherwise, using Lemma 4, the polynomial $\sum_{j=0}^{s-1} a_j \left(f_{k+j}\left(x,y\right) - f_{l+j}\left(x,y\right)\right)$ is nonconstant and has total degree at most d^{K+s-2}. Hence we can apply Lemmas 6 and 7 together with Lemma 1 to the inner sum, obtaining the upper bound

$$e^{c_0 d^{K+s-2}} M^{2-1/d^{K+s-2}+5c_1/\log\log(M/G)} G^{1/d^{K+s-2}}$$

for at most K^2 sums and positive constants c_0, c_1. Hence,

$$W^2 \leq KNM^2 + K^2 N e^{c_0 d^{K+s-2}} M^{2-1/d^{K+s-2}+5c_1/\log\log(M/G)} G^{1/d^{K+s-2}} \tag{7}$$

Now, without too much loss of generality we may assume $d^{K+s-2} \geq 2$. Next we put $K = \lceil c_2 \log\log\log(M/G) \rceil$, for some constant c_2 to guarantee that the first term dominates and the result follows. □

Next, let $D_s(N)$ denote the discrepancy of the points defined in 2 in the s-dimensional unit cube $[0,1)^s$. Using the last theorem, we proof the following:

Theorem 9. *If the sequence (u_n), given by (1) with a polynomial $f(X,Y) \in \mathbb{Z}_M[X,Y]$ with $f(X,Y)$ of total degree d and local degree in X at least 2 modulo every prime divisor of M, is purely periodic with period t and $t \geq N \geq 1$, then the bound*

$$D_s(N) = O\left(N^{-1/2} M (\log\log\log\log M)^s / (\log\log\log M)^{1/2}\right)$$

holds, where the implied constant depends only on s and d.

Proof. The statement follows from Lemma 3, taken with

$$L = \left\lceil N^{1/2} M^{-1} (\log\log\log M)^{1/2} \right\rceil$$

and the bound of Theorem 8, where all occurring $G = \gcd(a_1, \ldots, a_s, M)$ are at most L. □

4 Improvements on bounds for some M

In this section we will show that for some values of M, we can improve our bounds. Let $S_{\mathbf{a}}(N)$ and $D_s(N)$ be defined as before.

Theorem 10. *Let the sequence* (u_n)*, given by* (1) *with a polynomial* $f(X,Y) \in \mathbb{Z}_M[X,Y]$ *with* $f(X,Y)$ *of total degree* d *and local degree in* X*, at least 2 modulo every prime divisor of* M*, be purely periodic with period* t *and* $t \geq N \geq 1$*. Also suppose that*

$$\tau(M) \leq (\log M)^2.$$

Then the bound

$$\max_{\gcd(a_0,\ldots,a_{s-1},M)=G} |S_{\mathbf{a}}(N)| = O\left(N^{1/2}M(\log\log(M/G))^{-1/2}\right)$$

holds, where the implied constant depends only on s *and* d*.*

Proof. The proof is basically the same as for Theorem 8, except we use the smaller bound for $\tau(M)$ instead of Lemma 1. Hence (7), becomes:

$$W^2 \leq KNM^2 + K^2Ne^{c_0 d^{K+s-2}}(\log(M/G))^{10}M^{2-1/d^{K+s-2}}G^{1/d^{K+s-2}}$$

and putting $K = \lceil c_1 \log\log(M/G) \rceil$, for some constant c_1 to guarantee that the first term dominates, the result then follows. □

Recalling Lemma 2 we obtain:

Corollary 11. *Let* A *a positive integer number and the sequence* (u_n)*, given by* (1) *with a polynomial*
$f(X,Y) \in \mathbb{Z}_M[X,Y]$ *with* $f(X,Y)$ *of total degree* d *and local degree in* X *at least 2 modulo every prime divisor of* M*, be purely periodic with period* t *and* $t \geq N \geq 1$*, then for all* $M < A$*, except* $o(A)$ *of them, the bound*

$$\max_{\gcd(a_0,\ldots,a_{s-1},M)=G} |S_{\mathbf{a}}(N)| = O\left(N^{1/2}M(\log\log(M/G))^{-1/2}\right)$$

holds, where the implied constant depends only on s *and* d*.*

These last two theorems now allow us to prove stronger bounds on the discrepancy. Using Theorem 10 we get the following result:

Theorem 12. *Let the sequence* (u_n)*, given by* (1) *with a polynomial* $f(X,Y) \in \mathbb{Z}_M[X,Y]$ *with* $f(X,Y)$ *of total degree* d *and local degree in* X *at least 2 modulo every prime divisor of* M*, be purely periodic with period* t *and* $t \geq N \geq 1$*. Also suppose that* M *satisfies the inequality*

$$\tau(M) \leq (\log M)^2.$$

Then the bound

$$D_s(N) = O\left(N^{-1/2} M (\log\log\log M)^s / (\log\log M)^{1/2}\right)$$

holds, where the implied constant depends only on s *and* d.

Proof. The statement follows from Lemma 3, taken with

$$L = \left\lceil N^{1/2} M^{-1} (\log\log M)^{1/2} \right\rceil$$

and the bound of Theorem 10, where all occurring $G = \gcd(a_1, \ldots, a_s, M)$ are at most L. □

Combinating the last Theorem and Lemma 1:

Corollary 13. *Let* A *a positive integer number. If the sequence* (u_n)*, given by* (1) *with a polynomial* $f(X,Y) \in \mathbb{Z}_M[X,Y]$ *with* $f(X,Y)$ *of total degree* d *and local degree in* X *at least 2 modulo every prime divisor of* M*, be purely periodic with period* t *and* $t \geq N \geq 1$*, then for all* $M < A$ *but* $o(A)$ *choices of them, the bound*

$$D_s(N) = O\left(N^{-1/2} M (\log\log\log M)^s / (\log\log M)^{1/2}\right)$$

holds, where the implied constant depends only on s *and* d.

5 Open Questions

We remark that the technique used in [16] can not be employed here. It would be useful if an improvement using such or a similar method could be found.

Acknowledgments.

During the preparation of this paper, Domingo Gomez was supported by FWF grant S8313. The authors would especially like to thank Arne Winterhof for helpful advice and assistance.

References

[1] G. I. Arkhipov, V. N. Chubarikov, and A. A. Karatsuba, *Trigonometric Sums in Number Theory and Analysis*, de Grutyer Expositions in Mathematics **39**, W.de Grutyer, Berlin, 2004.

[2] T. Cochrane, 'Trigonometric approximation and uniform distribution modulo 1', *Proc. Amer. Math. Soc.*, **103** (1988), 695–702.

[3] M. Drmota and R. F. Tichy, *Sequences, discrepancies and applications*, Springer-Verlag, Berlin, 1997.

[4] E. D. El-Mahassni, I. E. Shparlinski, and A. Winterhof, 'Distribution of nonlinear congruential pseudorandom numbers for almost squarefree integers, *Monatsh. Math.*, **148** (2006), 297–307.

[5] E. El-Mahassni and A. Winterhof, 'On the distribution and linear complexity of counter-dependent nonlinear congruential pseudorandom number generators', JP Journal of Algebra, Number Theory and Applications (JANTA), Pushpa Publishing House, **6**II (2006), 411–423.

[6] E. D. El-Mahassni and A. Winterhof, 'On the distribution of nonlinear congruential pseudorandom numbers in residue rings', *Intern. J. Number Th.*, **2**(1) (2006), 163–168.

[7] J. Gutierrez and D. Gomez-Perez, 'Iterations of multivariate polynomials and discrepancy of pseudorandom numbers', Proc. 14th Symp. Appl. Algebra Algebraic Alg. Error-Correcting Codes. Lecture Notes in Comp. Sci., Springer, Berlin, **2227** (2001), 192–199.

[8] J. Gutierrez, I. Shparlinski and A. Winterhof, 'On the linear and nonlinear complexity profile of nonlinear pseudorandom number-generators', IEEE Trans. Inform. Theory **49**(1) (2003), 60–64.

[9] F. Griffin, H. Niederreiter and I. Shparlinski, 'On the distribution of nonlinear recursive congruential pseudorandom numbers of higher orders', Lecture Notes in Comp. Sci., Springer, Berlin, **1719** (1999), 87–93.

[10] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, Clarendon Press, Oxford, UK, 3rd ed., 1979.

[11] H. Niederreiter, *Random number generation and Quasi–Monte Carlo methods*, SIAM Press, 1992.

[12] H. Niederreiter, 'Design and analysis of nonlinear pseudorandom number generators', *Monte Carlo Simulation*, A.A. Balkema Publishers, Rotterdam, 2001, 3–9.

[13] H. Niederreiter and I. E. Shparlinski, 'On the distribution and lattice structure of nonlinear congruential pseudorandom numbers', *Finite Fields and Their Appl.*, **5** (1999), 246–253.

[14] H. Niederreiter and I. E. Shparlinski, 'Dynamical systems generated by rational functions', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **2643** (2003), 6–17.

[15] H. Niederreiter and A. Winterhof, 'Multiplicative character sums for nonlinear recurring sequences', Acta Arith. 111, (2004), 299-305 .

[16] H. Niederreiter and A. Winterhof, 'Exponential sums for nonlinear recurring sequences', *Finite Fields and their Applications*, to appear.

[17] A. Shamir and B. Tsaban, 'Guaranteeing the diversity of number generators', Inform. and Comp., **171** (2001), 350–363.

[18] A. Topuzŏglu and A. Winterhof, 'On the linear complexity profile of nonlinear congruential pseudorandom number generators of higher orders', *Applicable Algebra in Engineering, Communications and Computing*, **16** (2005), 219–228.

[19] A. Topuzŏglu and A. Winterhof, 'Pseudorandom Sequences', in Topics in Geometry, Cryptography and Coding Theory, Springer, Berlin, 2006, 135–166.