Designs, Codes and Cryptography (2023) 91:2473–2479
https://doi.org/10.1007/s10623-023-01209-5

Fixed points of the subset sum pseudorandom number generators

Igor E. Shparlinski[1]

Received: 17 August 2022 / Revised: 24 February 2023 / Accepted: 5 March 2023 /
Published online: 23 March 2023

Abstract
We give upper bounds on the power moments of the number of fixed points of a family of subset sum pseudorandom number generators, introduced by Rueppel (Analysis and design of stream ciphers, Springer-Verlag, Berlin, 1986).

Keywords Pseudorandom number generators · Subset sum · Fixed points

Mathematics Subject Classification 11K45 · 11T71 · 94A60

1 Introduction

For a positive integer t, we use $\mathbb{Z}_t$ to denote the residue ring modulo t, which we always assume to be represented by the set $\{0, \ldots, t-1\}$.

We fix an r-dimensional integer vector

$$\mathbf{z} = (z_1, \ldots, z_r) \in \mathbb{Z}_t^r \tag{1.1}$$

and define the function $S_{r,t,\mathbf{z}} : \mathbb{Z}_t \to \mathbb{Z}_t$ as follows. Given $w \in \mathbb{Z}_t$ (which following our convention we interpret as an integer from the set $\{0, \ldots, t-1\}$) we expand w in binary $w = \overline{u_s \ldots u_1}$, where u_i represents the i-th least significant bit of w, that is, the i-th bit from the right (if $r > s$ we pad w with $r - s$ leading zeroes) and then set

$$S_{r,t,\mathbf{z}}(w) = \sum_{i=1}^{r} u_i z_i \in \mathbb{Z}_t.$$

Furthermore, for a fixed vector $\mathbf{z}$ and a given initial value $w_0 \in \mathbb{Z}_t$ we define the sequence

$$v(0) = w_0, \quad v(n+1) = S_{r,t,\mathbf{z}}(v(n)), \qquad n = 0, 1, \ldots.$$

Communicated by L. Mérai.

✉ Igor E. Shparlinski
igor.shparlinski@unsw.edu.au

[1] School of Mathematics and Statistics, University of New South Wales, Sydney, NSW 2052, Australia

This construction has been introduced by Rueppel [23, Chapter 7] (see also [24, 25]), is known as *the subset sum pseudorandom number generator*. The efficiency of the generator and its cryptographic properties have been studied by Impagliazzo and Naor [18]. This generator is believed to be cryptographically secure since it relies on a combinatorial rather than an algebraic structure, which prevents mounting attacks similar to those designed in [2–4, 12–15, 19, 21], see also the references therein.

We note that one of the parameters characterising the pseudorandom properties of any map is the number of its fixed points since it reflects the mixing properties of this map. For example, the statistics of fixed points has been investigated for such classical cryptographic maps as the RSA encryption function [5] and the discrete logarithm [6, 7, 16, 17]. Several other examples of such results can be found in [1, 8, 9, 11, 20, 22]. A survey of such results, and of related results on short cycles in these maps, can be found in [27].

Here we consider this question for the map $w \mapsto S_{r,t,\mathbf{z}}(w)$. That is, we define and study

$$F_{r,t}(\mathbf{z}) = \#\{w \in \mathbb{Z}_t : \ w = S_{r,t,\mathbf{z}}(w)\}.$$

More precisely, we are interested in the power moments of this quantity over all t^r possible choices of the vectors (1.1):

$$M_\nu(r,t) = \frac{1}{t^r} \sum_{\mathbf{z} \in \mathbb{Z}_t^r} F_{r,t}(\mathbf{z})^\nu, \qquad \nu = 1, 2, \ldots.$$

In particular for the first moment, that is, for the average values of $F_{r,t}(\mathbf{z})$ we simplify the notation as

$$A(r,t) = M_1(r,t).$$

We recall that it has been shown in [26, Theorem 31.2] that for $t \ge 2^r$ the bound

$$A(r,t) \le (2t)^{1/2} + 2 \tag{1.2}$$

holds.

Here we improve this bound and also obtain a new bound for higher moments.

We note that the subset sum pseudorandom number generator is very fast as no modular multiplication is needed and no weaknesses has been discovered so far. However so far very few theoretical results have been known. Thus besides giving some concrete theoretic results, we also hope to attract more attention to this generator.

2 Evaluation of the average value of the number of fixed points

We start with a significant improvement of (1.2) and in fact we evaluate $A(r,t)$ explicitly.

Theorem 2.1 *For $t \ge 2^r$, we have*

$$A(r,t) = 2 - \left\lceil t/2^r \right\rceil / t.$$

Proof Let

$$m = \lceil \log t / \log 2 \rceil. \tag{2.1}$$

be the length of the binary expansion of t. Hence we write binary representations of $w \in \mathbb{Z}_t$ as binary strings of length exactly m (possible with some zeros on the left, that is, on the most significant positions).

Note that by our assumption $t \ge 2^r$ we have $m \ge r$. $m > r$.

Changing the order of summation we write

$$A(r,t) = \frac{1}{t^r} \sum_{\mathbf{z} \in \mathbb{Z}_t^r} \sum_{\substack{w \in \mathbb{Z}_t \\ w = S_{r,t,\mathbf{z}}(w)}} 1 = \frac{1}{t^r} \sum_{w \in \mathbb{Z}_t} \sum_{\substack{\mathbf{z} \in \mathbb{Z}_t^r \\ w = S_{r,t,\mathbf{z}}(w)}} 1.$$

For

$$w = \overline{u_m \dots u_{r+1} \underbrace{0 \dots 0}_{r \text{ zeros}}} \in \mathbb{Z}_t \tag{2.2}$$

whose binary expansion end with a string of r zeros, we obviously obtain $S_{r,t,\mathbf{z}}(w) = 0$. This leaves only one possible value for $w \in \mathbb{Z}_t$ with $S_{r,t,\mathbf{z}}(w) = w$, namely, $w = 0$, in which case the inner sum is equal to t^r.

The condition (2.2) on w means that $2^r \mid w$ and thus this happens for $\lceil t/2^r \rceil$ elements $w \in \mathbb{Z}_t$.

For the remaining $t - \lceil t/2^r \rceil$ choices of $w = \overline{u_m \dots u_1}$ with

$$(u_r, \dots, u_1) \ne (0, \dots, 0)\,,$$

there is at least one non-zero entry among the first r least significant bits in its binary representation, whose index we define as i. Then the component z_i of $\mathbf{z}$ as in (1.1) is uniquely defined from the equation

$$w = S_{r,t,\mathbf{z}}(w) = \sum_{i=1}^{r} u_i z_i$$

by the other components of $\mathbf{z}$, hence there exactly t^{r-1} such choices for $\mathbf{z}$.

Therefore,

$$A(r,t) = \frac{1}{t^r} \left(t^r + \left(t - \lceil t/2^r \rceil\right) t^{r-1}\right),$$

which concludes the proof. □

In particular, we see from Theorem 2.1 that we can improve (1.2) as

$$A(r,t) < 2.$$

3 Bounding higher moments of the number of fixed points

We recall that the notation $U = O(V)$, $U \ll V$ and $V \gg U$ are equivalent to $|U| \le cV$ for some positive constant c, which throughout the paper may depend on the order of the moment ν.

Here we always assume that t is a prime number, hence $\mathbb{Z}_t = \mathbb{F}_t$ is a finite field of t elements and hence we can use linear algebra over $\mathbb{F}_t$.

Theorem 3.1 *For a prime $t > 2^r$, for any fixed integer $\nu \ge 1$ we have*

$$M_\nu(r,t) \ll \left(t/2^r\right)^{\nu-1}$$

Proof Let m be defined by (2.1), that is, m is the length of the binary expansion of t. In particular, by our assumption $t > 2^r$ we have $m \ge r$.

We start with an observation that the value of $F_{r,t}(\mathbf{z})^\nu$ is equal to the number of solutions to the system of ν equations in $m\nu$ variables $u_{i,j} \in \{0, 1\}$, $i = 1, \ldots, m$, $j = 1, \ldots, \nu$:

$$\sum_{i=1}^{m} u_{i,j} \cdot 2^{i-1} \equiv \sum_{i=1}^{r} u_{i,j} \cdot z_i \pmod{t}, \qquad j \in \{1, \cdots, \nu\}, \tag{3.1}$$

Note that the variables $u_{i,j} \in \{0, 1\}$, $i = 1, \ldots, m$, $j = 1, \ldots, \nu$, in (3.1) correspond to ν vectors $(\mathbf{u}_1, \ldots, \mathbf{u}_\nu)$ coming from binary expansions of solutions $w_1, \ldots, w_\nu \in \mathbb{F}_t$ to ν independent equations $w_j = S_{r,t,\mathbf{z}}(w_j)$, $j = 1, \ldots, \nu$.

We define $U_{\nu,r}(s)$ to be the set ν-tuples of binary vectors $(\mathbf{u}_1, \ldots, \mathbf{u}_\nu)$ for which the first r components form a matrix of rank s over $\mathbb{F}_t$, that is,

$$\operatorname{rank}_{\mathbb{F}_t} \begin{pmatrix} u_{1,1} & \ldots & u_{1,r} \\ \ldots & \ldots & \ldots \\ u_{\nu,1} & \ldots & u_{\nu,r} \end{pmatrix} = s. \tag{3.2}$$

Clearly for every ν-tuple $(\mathbf{u}_1, \ldots, \mathbf{u}_\nu) \in U_{\nu,r}(s)$ of vectors, the system of congruences (3.1) has at most t^{r-s} solutions in $\mathbf{z} \in \mathbb{Z}_t^r$.

We now switch the roles of the binary variables $u_{i,j} \in \{0, 1\}, i = 1, \ldots, m$, $j = 1, \ldots, \nu$, and the vectors $\mathbf{z} \in \mathbb{Z}_t^r$. That is, for each choice of $u_{i,j} \in \{0, 1\}, i = 1, \ldots, m$, $j = 1, \ldots, \nu$, we count the number of vectors $\mathbf{z} \in \mathbb{Z}_t^r$ satisfying (3.1).

We can then bound our summation in terms of $\#U_{\nu,r}(s)$:

$$\sum_{\mathbf{z} \in \mathbb{Z}_t^r} F_{r,t}(\mathbf{z})^\nu \le \sum_{s=0}^{\nu} \#U_{\nu,r}(s) t^{r-s}. \tag{3.3}$$

First we note that $\#U_{\nu,r}(0) = 1$ as this corresponds to the zero matrix in (3.2) and thus (3.1) implies that the remaining $m - r$ components of each of the binary vectors $(\mathbf{u}_1, \ldots, \mathbf{u}_\nu)$ also vanish. Then we have t^r choices for $\mathbf{z}$. Hence such vectors contribute in total t^r to the case $s = 0$.

To estimate $\#U_{\nu,r}(s)$ with $s \ge 1$, we note that if we fix s linearly independent vectors

$$(\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_s}), \qquad 1 \le j_1 < \ldots < j_s \le \nu,$$

in a family of vectors $(\mathbf{u}_1, \ldots, \mathbf{u}_\nu) \in U_{\nu,r}(s)$, then any other vector $\mathbf{u}_j$ belongs to the linear span of $\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_s}$ over $\mathbb{F}_t$. That is,

$$\mathbf{u}_j = a_1 \mathbf{u}_{j_1} + \ldots + a_s \mathbf{u}_{j_s} \tag{3.4}$$

for some $a_1, \ldots, a_s \in \mathbb{F}_t$. By the *Cramer rule* we have

$$a_j \equiv \frac{\Delta_j}{\Delta} \pmod{t}, \qquad j = 1, \ldots, s, \tag{3.5}$$

for some determinants $\Delta, \Delta_1, \ldots, \Delta_s$ over $\mathbb{F}_t$ forms by the components of the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_\nu$ and with $\Delta \not\equiv 0 \pmod{t}$. Since all vectors $\mathbf{u}_1, \ldots, \mathbf{u}_\nu$ are binary, we easily infer that

$$|\Delta|, |\Delta_j| \le 2^{-s} (s+1)^{(s+1)/2}. \tag{3.6}$$

see, for example, [10, Problem 523]. Thus, adjusting the signs we see from (3.5) and (3.6) that, regardless of the choice of $\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_s}$, each vector $(a_1, \ldots, a_s)$ satisfies

$$(a_1, \ldots, a_s) = \left(D_1 D^{-1}, \ldots, D_s D^{-1}\right) \pmod{t} \tag{3.7}$$

(where D^{-1} is computed modulo t) with some integers

$$D \in \left[1, 2^{-s}\,(s+1)^{(s+1)/2}\right]$$

and

$$D_j \in \left[-2^{-s}\,(s+1)^{(s+1)/2}, 2^{-s}\,(s+1)^{(s+1)/2}\right], \qquad j = 1, \ldots, s,$$

and hence there are at most

$$A_s = 2^{-s}\,(s+1)^{(s+1)/2}\left(1 + 2^{-s+1}\,(s+1)^{(s+1)/2}\right)^s \tag{3.8}$$

choices for the vector of the coefficients $(a_1, \ldots, a_s)$ in (3.4).

We emphasise that the meaning of the bound (3.8) is even if the number of possible vectors $(\mathbf{u}_1, \ldots, \mathbf{u}_\nu) \in U_{\nu,r}(s)$, and thus the number systems of relations (3.4). grows rapidly with r and t, the number of possible choices for the coefficients $(a_1, \ldots, a_s)$ can be bounded only in terms of s (and thus of ν) and therefore independently on r and t.

This implies that when $\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_s}$ are fixed to satisfy (3.2), there at most A_s possibilities to form the first r coordinates of each of the other vectors to form a ν-tuple $(\mathbf{u}_1, \ldots, \mathbf{u}_\nu) \in U_{\nu,r}(s)$, and thus at most $A_s 2^{m-r}$ possibilities for the whole vector. Since there are at most

$$\binom{\nu}{s}(2^m)^s = \binom{\nu}{s}2^{ms}$$

choices for $\mathbf{u}_{j_1}, \ldots, \mathbf{u}_{j_s}$ we obtain

$$\#U_{\nu,r}(s) \le \binom{\nu}{s}2^{ms}\left(A_s 2^{m-r}\right)^{\nu-s}.$$

Since we assume that ν is fixed and $s \le \nu$, this simplifies as

$$\#U_{\nu,r}(s) \ll 2^{ms+(m-r)(\nu-s)} = 2^{m\nu-r(\nu-s)} \ll t^\nu 2^{-r(\nu-s)}.$$

We can now substitute the above bound for $\#U_{\nu,r}(s)$ in (3.3), getting

$$\sum_{\mathbf{z}\in\mathbb{Z}_t^r} F_{r,t}(\mathbf{z})^\nu \ll t^r + \sum_{s=1}^{\nu} t^{\nu+r-s}2^{-r(\nu-s)} = t^r + t^r \sum_{s=1}^{\nu}\left(t\cdot 2^{-r}\right)^{\nu-s}.$$

Note that we have requested that $t > 2^r$, which implies $t \cdot 2^{-r} > 1$, so

$$\sum_{\mathbf{z}\in\mathbb{Z}_t^r} F_{r,t}(\mathbf{z})^\nu \ll t^r + t^r \sum_{s=1}^{\nu}\left(t\cdot 2^{-r}\right)^{\nu-s} \ll t^r\left(t\cdot 2^{-r}\right)^{\nu-1},$$

which concludes the proof. □

4 Comments

In Theorem 3.1, we have suppressed the dependence on the order of the moments ν. There are two reasons for this.

First, we do not consider ν to be an important parameter. For example, the choice of $\nu = 2$ already gives us important information and extra technical calculations do not seem to justify the importance of this. However, we provide all necessary estimates for this, if one decides to trace the dependence on ν. For example, we note that (3.6) is slightly stronger that the classical *Hadamard inequality*, which is still sufficient for our purposes, since we do not compute the explicit dependence on ν. Besides the potential contribution to computing explicit dependence on ν, we also present (3.6) because we believe it deserves to be known more broadly.

The second reason is that before computing the explicit dependence on ν, one has to attempt to improve the bound (3.8) on the number of distinct vectors which can be solutions to all non-singular systems of s linear congruences modulo t with binary coefficients. This question seems to be of independent interest and certainly deserves further investigation. Certainly one can improve (3.8) by an absolute constant, taking into account that in (3.7) we need only count $D, D_1, \ldots, D_s$ with

$$\gcd(D, D_1, \ldots, D_s) = 1.$$

However we are interested in more substantial improvements.

We also would like to note that our approach does not extend on bounding the number of short cycles. For example, we do not have any nontrivial estimate on the number 2-cycles

$$\#\{w \in \mathbb{F}_t : \ w = S_{r,t,\mathbf{z}}\left(S_{r,t,\mathbf{z}}(w)\right)\}$$

on average over $\mathbf{z} \in \mathbb{F}_t^r$, which is another interesting open question.

Acknowledgements The author is very grateful to the referees for the very careful reading of the manuscript and many very useful comments. This work was partially supported by the Australian Research Council Grant DP200100355.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

References

1. Balog A., Broughan K.A., Shparlinski I.E.: On the number of solutions of exponential congruences. Acta Arith. **148**, 93–103 (2011).
2. Blackburn S.R., Gómez-Pérez D., Gutierrez J., Shparlinski I.E.: Predicting the inversive generator. Lecture Notes Comp. Sci. **2898**, 264–275 (2003).
3. Blackburn S.R., Gómez-Pérez D., Gutierrez J., Shparlinski I.E.: Predicting nonlinear pseudorandom number generators. Math. Comput. **74**, 1471–1494 (2005).
4. Blackburn S.R., Gómez-Pérez D., Gutierrez J., Shparlinski I.E.: Reconstructing noisy polynomial evaluation in residue rings. J. Algorithms **61**, 47–90 (2006).

5. Blakley G., Borosh I.: Rivest-Shamir-Adleman public key cryptosystems do not always conceal messages. Comput. Math. Appl. **5**, 169–178 (1979).
6. Bourgain, J., Konyagin, S.V., Shparlinski, I.E.: Product sets of rationals, multiplicative translates of subgroups in residue rings and fixed points of the discrete logarithm. Intern. Math. Res. Not., **2008**, Article rnn090 (2008). (Corrigenda: Int. Math. Res. Not., **2009**, 3146–3147) (2009)
7. Bourgain J., Konyagin S.V., Shparlinski I.E.: Distribution of elements of cosets of small subgroups and applications. Intern. Math. Res. Not. **2012**, 1968–2009 (2012).
8. Chen Z., Winterhof A.: Interpolation of Fermat quotients. SIAM J. Discret. Math. **28**, 1–7 (2014).
9. Cilleruelo J., Garaev M.Z.: Congruences involving product of intervals and sets with small multiplicative doubling modulo a prime and applications. Math. Proc. Camb. Philos. Soc. **160**, 477–494 (2016).
10. Faddeev D.K., Sominskii I.S.: Problems in Higher Algebra. W. H. Freeman, San Francisco (1965).
11. Felix A.T., Kurlberg P.: On the fixed points of the map $x \mapsto x^x$ modulo a prime, II. Finite Fields Appl. **48**, 141–159 (2017).
12. Gómez-Pérez D., Gutierrez J., Ibeas Á.: Attacking the Pollard generator. IEEE Trans. Inform. Theory **52**, 5518–5523 (2006).
13. Gutierrez J.: Attacking the linear congruential generator on elliptic curves via lattice techniques. Cryptogr. Commun. **14**, 505–525 (2002).
14. Gutierrez J.: Reconstructing points of superelliptic curves over a prime finite field. Adv. Math. Commun. (2022). https://doi.org/10.3934/amc.2022022.
15. Gutierrez J., Ibeas Á.: Inferring sequences produced by a linear congruential generator on elliptic curves missing high-order bits. Des. Codes Cryptogr. **41**, 199–212 (2007).
16. Holden, J., Moree, P.:New conjectures and results for small cycles of the discrete logarithm. In: Proceedings of the High Primes and Misdemeanours: Lectures in Honour of the 60th Birthday of Hugh Cowie Williams, Fields Institute Communications, vol. 41, pp. 245–254. Amer. Math. Soc., Providence, RI (2004)
17. Holden J., Moree P.: Some heuristics and and results for small cycles of the discrete logarithm. Math. Comput. **75**, 419–449 (2006).
18. Impagliazzo R., Naor M.: Efficient cryptographic schemes provably as secure as subset sum. J. Cryptol. **9**, 199–216 (1996).
19. Krawczyk H.: How to predict congruential generators. J. Algorithms **13**, 527–545 (1992).
20. Kurlberg P., Luca F., Shparlinski I.E.: On the fixed points of the map $x \mapsto x^x$ modulo a prime. Math. Res. Lett. **22**, 141–168 (2015).
21. Lagarias, J. C.: Pseudorandom number generators in cryptography and number theory. in: Proceedings of the Proceedings of Symposia in Applied Mathematics, vol.42, pp. 115–143. Amer. Math. Soc., Providence, RI (1990)
22. Ostafe A., Shparlinski I.E.: Pseudorandomness and dynamics of Fermat quotients. SIAM J. Discret. Math. **25**, 50–71 (2011).
23. Rueppel R.A.: Analysis and Design of Stream Ciphers. Springer-Verlag, Berlin (1986).
24. Rueppel R.A.: Stream Ciphers, Contemporary Cryptology: The Science of Information Integrity, pp. 65–134. IEEE Press, New York (1992).
25. Rueppel R.A., Massey J.L.: Knapsack as a nonlinear function. In: Proceedings of the IEEE International Symposium on Information Theory, p. 46. IEEE Press, NY (1985)
26. Shparlinski I.E.: Cryptographic Applications of Analytic Number Theory. Birkhäuser, Basel (2003).
27. Shparlinski I.E.: Dynamical systems of non-algebraic origin: fixed points and orbit lengths. Contemp. Math. **669**, 261–283 (2016).