Electronic Colloquium on Computational Complexity, Report No. 21 (1999)

ON THE UNIFORMITY OF DISTRIBUTION OF A CERTAIN PSEUDO-RANDOM FUNCTION

Igor E. Shparlinski*
Department of Computing, Macquarie University
NSW 2109, Australia
igor@mpce.mq.edu.au

April 8, 1999

Abstract

We show that a pseudo-random number generator, introduced recently by M. Naor and O. Reingold, possess one more attractive and useful property. Namely, it is proved that for almost all values of parameters it produces a uniformly distributed sequence. The proof is based on some recent bounds of exponential sums with exponential functions.

1991 Mathematics Subject Classification. Primary 11K45, 94A60; Secondary 11K38, 11L07, 11T23

Key words and phrases. Pseudo-random numbers, Exponential functions, Exponential sums.

*Supported in part by ARC grant A69700294.

ISSN 1433-8092

1 Introduction

Let p be an n-bit prime, $2^{n-1} \le p \le 2^n - 1$ and let l be a prime divisor of $p-1$.

Denote by $\mathbb{F}_p$ the finite field of p elements and select an element $g \in \mathbb{F}_p^*$ of multiplicative order l modulo p. We recall that $\vartheta \in \mathbb{F}_p^*$ is of multiplicative order t if and only if

$$g^i \ne 1, \ 1 \le i \le t-1, \qquad g^t = 1.$$

Then for each n-dimensional vector $\mathbf{a} = (a_1, \ldots, a_n) \in (\mathbb{Z}/l)^n$ one can define the function

$$f_{\mathbf{a}}(x) = g^{a_1^{x_1} \ldots a_n^{x_n}} \in \mathbb{F}_p,$$

where $x = x_1 \ldots x_n$ is the bit representation of an integer x, $0 \le x \le 2^n - 1$, with some extra leading zeros is necessary.

In [6] M. Naor and O. Reingold has proposed the function $f_{\mathbf{a}}(x)$ as an efficient pseudo-random function (for a randomly chosen vector $\mathbf{a} \in (\mathbb{Z}/l)^n$). It is shown in [6] that this function can be computed in parallel by threshold circuits of bounded depth and polynomial size and also has some very desirable security property, provided certain standard cryptographic assumptions hold.

Here we show that this function has one more useful feature, which comes as an additional bonus to the aforementioned cryptographic properties of $f_{\mathbf{a}}(x)$. Namely, we prove that for almost all vectors $\mathbf{a} \in (\mathbb{Z}/l)^n$, the sequence $f_{\mathbf{a}}(x)$, $x = 0, 1, \ldots, 2^n - 1$, is asymptotically uniformly distributed.

We remark that although this property does not seem to have any immediate cryptographic implications, the inverse fact, that is, non-uniformity of distribution, if it had been true, would have disastrous consequences for applications of this function. Besides this, studying the uniformity of distribution of interesting functions is a very attractive number theoretic question. Our main tool is the bound of exponential sums with exponential functions which is due to S. V. Konyagin and the author [2]. Previously known estimates, which are due to N. M. Korobov [3, 4] and H. Niederreiter [7, 8], can also be used, however they imply weaker results.

2 Preparations

We identify $\mathbb{F}_p$ with the set $\{0, \ldots, p-1\}$.

For a set $\mathcal{M} \subseteq \mathbb{F}_p$ we define the *discrepancy* $D(\mathcal{M})$ modulo p as

$$D(\mathcal{M}) = \sup_{\mathcal{I} \subseteq [0,1]} \left| \frac{N(\mathcal{I})}{\#\mathcal{M}} - |\mathcal{I}| \right|,$$

where $N(\mathcal{I})$ is the number of fractional parts $\{m/p\}$ with $m \in \mathcal{M}$ which hit the interval $\mathcal{I} = [\alpha, \beta] \subseteq [0,1]$ of length $|\mathcal{I}| = \beta - \alpha$.

We denote by $D_{\mathbf{a}}$ the discrepancy of the set $\{f_{\mathbf{a}}(x) \,|\, x = 0, 1, \ldots, 2^n - 1\}$. We show that $D_{\mathbf{a}} = o(1)$ for all except possibly $o(l^n)$ vectors $\mathbf{a} \in (\mathbb{Z}/l)^n$, provided that $l \ge p^{1/3+\varepsilon}$ with any fixed $\varepsilon > 0$.

Throughout the paper the implied constants in symbols 'O' and '$\ll$' are absolute (we recall that $A \ll B$ is equivalent to $A = O(B)$)

We also denote by $\log a$ the binary logarithm of a and

$$\mathbf{e}(a) = \exp(2\pi i a/p).$$

We need a form of the *Erdös–Turán inequality* which relates the discrepancy and exponential sums, see Corollary 1.1 to Chapter 1 of [5] or Corollary 3.11 of [8].

Lemma 1. *For any set $\mathcal{M} \subseteq \mathbb{F}_p$ the bound*

$$D(\mathcal{M}) \ll \frac{1}{p} + \frac{1}{\#\mathcal{M}} \sum_{h=1}^{p-1} \frac{1}{h} \left| \sum_{m \in \mathcal{M}} \mathbf{e}\,(hm) \right|$$

holds.

We also need the following upper bound on exponential sums with exponential functions which is essentially Theorem 3.4 of [2].

Lemma 2. *Let p be prime and let $\vartheta \in \mathbb{F}_p^*$ be of multiplicative order t modulo p. Then the bound*

$$\max_{\gcd(h,p)=1} \left| \sum_{r=0}^{t-1} \mathbf{e}(h\vartheta^r) \right| \ll B(t,p)$$

where

$$B(t,p) = \begin{cases} p^{1/2}, & \text{if } t \ge p^{2/3}; \\ p^{1/4}t^{3/8}, & \text{if } p^{1/2} \le t \le p^{2/3}; \\ p^{1/8}t^{5/8}, & \text{if } p^{1/3} \le t \le p^{1/2}; \end{cases}$$

holds.

3 Main Result

Now we are prepared to prove our main result.

Theorem 3. *For for all, except possibly $o(p^n)$, vectors $\mathbf{a} \in (\mathbb{Z}/l)^n$, the bound*

$$D_{\mathbf{a}} \le \Delta(l,p)$$

where

$$\Delta(l,p) = \begin{cases} p^{(1-\gamma)/2}l^{-1/2}\log^2 p, & \text{if } l \ge p^{\gamma}; \\ p^{1/2}l^{-1}\log^2 p, & \text{if } p^{2/3} \le l \le p^{\gamma}; \\ p^{1/4}l^{-5/8}\log^2 p, & \text{if } p^{1/2} \le l \le p^{2/3}; \\ p^{1/8}l^{-3/8}\log^2 p, & \text{if } p^{1/3} \le l \le p^{1/2}; \end{cases}$$

and $\gamma = 2.5 - \log 3 = 0.9150\ldots$, holds.

Proof. We may assume that p is large enough, in particular that $n \ge 3$. From Lemma 1 we conclude that It is easy to show that

$$\sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} D_{\mathbf{a}} \ll l^n p^{-1} + 2^{-n} \sum_{h=1}^{p-1} \frac{1}{h} W_h, \tag{1}$$

where

$$W_h = \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} \left| \sum_{x=0}^{2^n-1} \mathbf{e}\left(h f_{\mathbf{a}}(x)\right) \right|.$$

Using the Cauchy inequality, we derive

$$W_h^2 = l^n \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} \left| \sum_{x=0}^{2^n-1} \mathbf{e}\left(h f_{\mathbf{a}}(x)\right) \right|^2.$$

We recall that $|z|^2 = z\overline{z}$ for any complex z and that $\overline{\mathbf{e}(a)} = \mathbf{e}(-a)$ for any real a. Then, it is easy to see that replacing the square of the inner sum by a double sum and changing the order of summation we obtain

$$W_h^2 = l^n \sum_{x,y=0}^{2^n-1} \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} \mathbf{e}\left(h\left(f_{\mathbf{a}}(x) - f_{\mathbf{a}}(y)\right)\right).$$

If $x = y$ the inner sum is equal to l^n.

Now we consider the case $x \neq y$. We say that $x \succ y$ if $x_i \geq y_i$, $i = 1, \ldots, n$, where $x = x_1 \ldots x_n$ and $y = y_1 \ldots y_n$ are the bit representation of x and y.

We also say that integers x and y are *comparable* if either $x \succ y$ or $y \succ x$.

If $x \neq y$ and $x \succ y$ we fix i, $1 \leq i \leq n$, with $x_i = 1$, $y_i = 0$.

We see that the term $f_{\mathbf{a}}(y)$ does not depend on a_i.

Let the vector $(z_1, \ldots, z_{n-1})$ be formed by the all bits of x except x_i, that is, $z_k = x_k$ if $1 \leq k < i$ and $z_k = x_{k+1}$ if $i \leq k \leq n-1$. Therefore,

$$\left| \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} \mathbf{e}\left(h\left(f_{\mathbf{a}}(x) - f_{\mathbf{a}}(y)\right)\right) \right| \leq \sum_{\mathbf{b} \in (\mathbb{Z}/l)^{n-1}} \left| \sum_{r=0}^{l-1} \mathbf{e}\left(h\vartheta_{\mathbf{b},x}^r\right) \right|,$$

where $\mathbf{b} = (b_1, \ldots, b_{n-1})$ and

$$\vartheta_{\mathbf{b},x} = g^{b_1^{z_1} \ldots b_{n-1}^{z_{n-1}}}.$$

We see that if

$$b_1 \ldots b_{n-1} \not\equiv 0 \pmod{l}$$

then, because l is prime, $\vartheta_{\mathbf{b},x}$ is of multiplicative order l. Hence the bound of Lemma 2 applies to the inner sum. For other $O(nl^{n-2})$ vectors $\mathbf{b}$ we estimate the inner sum trivially by l.

It is easy to see that there are

$$\sum_{k=0}^{n} \binom{n}{k} 2^k = 3^n$$

pairs of (x, y), $0 \leq x, y \leq 2^n - 1$, with $x \succ y$. Thus this part of the sum can be estimated as

$$\left| \sum_{\substack{x,y=0 \\ x \neq y,\ x \succ y}}^{2^n-1} \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} \mathbf{e}\left(h\left(f_{\mathbf{a}}(x) - f_{\mathbf{a}}(y)\right)\right) \right| \ll 3^n \left(nl^{l-1} + l^{n-1}B(l,p)\right).$$

The case $x \neq y$ and $y \succ x$ can be considered quite analogously.

Finally, let us consider pairs of x and y which are not comparable. In this case there are i and j, $1 \le i, j \le n$, with $x_i = y_j = 1$ and $x_j = y_i = 0$. We see that the term $f_{\mathbf{a}}(y)$ does not depend on a_i and the term $f_{\mathbf{a}}(x)$ does not depend on a_j.

Let the vector $(z_1, \ldots, z_{n-2})$ be formed by the all bits of x except x_i and x_j that is, $z_k = x_k$ if $1 \le k < I$, $z_k = x_{k+1}$ if $I \le k < J-1$ and $z_k = x_{k+2}$ if $J-1 \le k \le n-2$, where $I = \min\{i, j\}$ and $J = \max\{i, j\}$. We also form the vector $(w_1, \ldots, w_{n-2})$ in a similar way from the all bits of y except y_i and y_j.

Therefore,

$$\left| \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} \mathbf{e}\left(h\left(f_{\mathbf{a}}(x) - f_{\mathbf{a}}(y)\right)\right) \right| \le \sum_{\mathbf{b} \in (\mathbb{Z}/l)^{n-2}} \left| \sum_{r=0}^{l-1} \mathbf{e}\left(h\lambda_{\mathbf{b},x}^r\right) \right| \left| \sum_{s=0}^{l-1} \mathbf{e}\left(h\mu_{\mathbf{b},y}^s\right) \right|,$$

where $\mathbf{b} = (b_1, \ldots, b_{n-2})$,

$$\lambda_{\mathbf{b},x} = g^{b_1^{z_1} \ldots b_{n-2}^{z_{n-2}}} \qquad \text{and} \qquad \mu_{\mathbf{b},y} = g^{b_1^{w_1} \ldots b_{n-2}^{w_{n-2}}}.$$

We see that if

$$b_1 \ldots b_{n-1} \not\equiv 0 \pmod{l}$$

then, because l is prime, $\lambda_{\mathbf{b},x}$ and $\mu_{\mathbf{b},y}$ are both of multiplicative order l. Hence the bound of Lemma 2 applies to both inner sums. For other $O(nl^{n-3})$ vectors $\mathbf{b}$ we estimate the inner sums trivially by l each.

Therefore, for each pair of x and y which are not comparable the bound

$$\left| \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} \mathbf{e}\left(h\left(f_{\mathbf{a}}(x) - f_{\mathbf{a}}(y)\right)\right) \right| \ll nl^{n-1} + l^{n-2}B(l,p)^2$$

holds.

Putting everything together and taking into account that $2^n = O(p)$ and $3^n = O(p^\alpha)$, where $\alpha = \log 3$, we derive

$$\begin{aligned} W_h^2 &\ll l^n \left(2^n l^n + 3^n \left(nl^{n-1} + l^{n-1}B(l,p)\right) + 2^{2n}\left(nl^{n-1} + l^{n-2}B(l,p)^2\right)\right) \\ &\ll pl^{2n} + np^\alpha l^{2n-1} + p^\alpha l^{2n-1}B(l,p) + np^2 l^{2n-1} + p^2 l^{2n-2}B(l,p)^2. \end{aligned}$$

It is easy to see that the terms including $B(l,p)$ dominate all other terms. Thus

$$W_h \ll p^{\alpha/2} l^{n-1/2} B(l,p)^{1/2} + p l^{n-1} B(l,p). \tag{2}$$

Combining (1) and (2), we derive

$$\sum_{\substack{\mathbf{a} \in (\mathbb{Z}/l)^n \\ D_{\mathbf{a}} \geq \Delta(l,p)}} 1 \leq \frac{1}{\Delta(l,p)} \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} D_{\mathbf{a}} \ll \frac{p^{\alpha/2-1} l^{n-1/2} B(l,p)^{1/2} + l^{n-1} B(l,p)}{\Delta(l,p)}.$$

Remarking that the first term in the numerator dominates if and only if $l \leq p^{\gamma}$, we obtain the desired result. □

4 Remarks

It is easy to see that the bound of Theorem 3 is nontrivial beginning with $l \geq p^{1/3+\varepsilon}$ with any fixed $\varepsilon > 0$. It is also useful to recall that there exist infinitely many primes p such that $p-1$ has a prime divisor $l > p^{0.677}$, see [1]. For such p and l we see that $D_{\mathbf{a}} \leq l^{-0.26}$ for almost all $\mathbf{a} \in (\mathbb{Z}/l)^n$. Moreover, it is expected that $l = (p-1)/2$ is prime for infinitely many primes p. Such pairs of p and l are of special value for cryptography. For them we deduce that $D_{\mathbf{a}} \leq l^{-0.41}$ for almost all $\mathbf{a} \in (\mathbb{Z}/l)^n$.

Analogues of Theorem 3 can also be obtained for other pseudo-random number generators from [6]. The same method can also be used to study the distribution of $f_{\mathbf{a}}(x)$ for $x = 0, 1, \ldots, N-1$ with $N \leq 2^n$.

Finally, it would also be interesting to study the distribution of k-tuples $(f_{\mathbf{a}}(x), \ldots, f_{\mathbf{a}}(x+k-1))$.

References

[1] R. C. Baker and G. Harman, 'Shifted primes without large prime factors', *Acta Arithm.*, **83** (1998), 331–361.

[2] S. V. Konyagin and I. E. Shparlinski, *Character sums with exponential functions and their applications*, Cambridge Univ. Press, Cambridge, 1999 (to appear).

[3] N. M. Korobov, 'On the distribution of digits in periodic fractions', *Math. USSR – Sbornik*, **18** (1972), 659–676.

[4] N. M. Korobov, *Exponential sums and their applications*, Kluwer Acad. Publ., Dordrecht, 1992.

[5] H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, CBMS Regional Conference Series in Math. Vol. 84, Amer. Math. Soc., Providence, 1994.

[6] M. Naor and O. Reingold, 'Number-theoretic constructions of efficient pseudo-random functions', *Proc 38th IEEE Symp. on Foundations of Comp. Sci.*, 1997, 458–467.

[7] H. Niederreiter, 'Quasi-Monte Carlo methods and pseudo-random numbers', *Bull. Amer. Math. Soc.*, **84** (1978), 957–1041.

[8] H. Niederreiter, *Random number generation and quasi–Monte Carlo methods*, SIAM, Philadelphia, 1992.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'