MATHEMATICS OF COMPUTATION
Volume 83, Number 289, September 2014, Pages 2429–2434
S 0025-5718(2014)02794-4
Article electronically published on January 17, 2014

ON THE MULTIDIMENSIONAL DISTRIBUTION OF THE NAOR–REINGOLD PSEUDO-RANDOM FUNCTION

SAN LING, IGOR SHPARLINSKI, AND HUAXIONG WANG

Abstract. We show that the pseudo-random number function, introduced by M. Naor and O. Reingold (FOCS, 1997), possesses 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 character sums with exponential functions.

1. Introduction

Let p be an n-bit prime, so $2^{n-1} \le p \le 2^n - 1$, 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. We recall that $\vartheta \in \mathbb{F}_p^*$ is of multiplicative order t if and only if

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

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 x_n being the least significant bit and some extra leading zeros if necessary, namely,

$$x = \sum_{i=1}^{n} x_i 2^{n-i}.$$

In [7] M. Naor and O. Reingold have 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 [7] that this function can be computed in parallel by threshold circuits of bounded depth and polynomial size and also has some very desirable security properties, provided certain standard cryptographic assumptions (for example, the decisional Diffie-Hellman assumption) hold.

It has been shown in [11] 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 modulo p. An exponential lower bound on the *linear complexity* of this generator has been obtained in [6, 10]. In [1] this bound has been extended to *nonlinear complexity*.

Received by the editor July 28, 2012 and, in revised form, December 17, 2012.

2010 *Mathematics Subject Classification.* Primary 11K45, 11T23, 65C10, 94A60.

Key words and phrases. Naor-Reingold pseudo-random function, discrepancy, exponential sums.

During the preparation of this paper, the authors were supported by NRF Grant CRP2-2007-03 (Singapore).

The second author was supported in part by ARC Grant DP1092835 (Australia).

For the *elliptic curve* version of this generator similar results have been obtained in [3, 9, 12].

Here we show that given an integer $s \ge 1$, for almost all vectors $\mathbf{a} \in (\mathbb{Z}/l)^n$, the sequence of vectors $(f_{\mathbf{a}}(x), \ldots, f_{\mathbf{a}}(x+s-1))$, $x = 0, 1, \ldots, 2^n - 1$, is asymptotically uniformly distributed modulo p, generalizing the result in [11] to the multidimensional case. Our main tool is the bound of character sums with linear combinations of exponential functions, which is due to Bourgain [2]. As far as we know, this result has never been used in the theory of pseudo-random functions and number generators. So we believe that the introduction of this new technique in the area is of independent interest and may have further applications.

2. Preparations

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

For an integer $s \ge 1$ and an N-term sequence $\mathcal{M}$ of s-dimensional vectors over $\mathbb{F}_p$,

$$\mathcal{M} = \{\mathbf{m}_n = (m_{0,j}, \ldots, m_{s-1,j}) \in \mathbb{F}_p^s \ : \ j = 1, \ldots, N\}, \tag{1}$$

we define the *discrepancy* $D(\mathcal{M})$ modulo p as

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

where $N(\mathcal{B})$ is the number of vectors of fractional parts,

$$(\{m_{0,j}/p\}, \ldots, \{m_{s-1,j}/p\}), \quad j = 1, \ldots, N,$$

inside of the box $\mathcal{B} = [\alpha_0, \beta_0] \times \cdots \times [\alpha_{s-1}, \beta_{s-1}] \subseteq [0,1]^s$ of volume $|\mathcal{B}| = (\beta_0 - \alpha_0) \ldots (\beta_{s-1} - \alpha_{s-1})$.

We denote by $D_{l,p,s}(\mathbf{a})$ the discrepancy modulo p of the set

$$\{(f_{\mathbf{a}}(x), \ldots, f_{\mathbf{a}}(x+s-1)) \ : \ x = 0, 1, \ldots, 2^n - s\}.$$

We show that $D_{l,p,s}(\mathbf{a}) = o(1)$ for all but possibly $o(l^n)$ vectors $\mathbf{a} \in (\mathbb{Z}/l)^n$, provided that $l \ge p^{\varepsilon}$ with any fixed $\varepsilon > 0$.

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

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

$$\mathbf{e}_p(a) = \exp(2\pi i a/p), \qquad a \in \mathbb{F}_p.$$

Thus $\mathbf{e}_p(a)$ is a nontrivial additive character of $\mathbb{F}_p$.

We need the following version of the celebrated *Koksma–Szüsz inequality*; see [4, Theorem 1.21]:

Lemma 1. *For any sequence $\mathcal{M}$ of the form* (1)*, the following bound holds:*

$$D(\mathcal{M}) \ll \frac{1}{p} + \frac{1}{N} \sum_{(h_0, \ldots, h_{s-1}) \in \mathbb{F}_p^s} \prod_{i=0}^{s-1} \frac{1}{|h_i| + 1} \left| \sum_{j=1}^{N} \mathbf{e}_p\left(h_0 m_{0,j} + \cdots + h_{s-1} m_{s-1,j}\right) \right|.$$

We also need the following upper bound on character sums with exponential functions which is due to Bourgain [2, Theorem 2].

Lemma 2. *Let $\varepsilon > 0$ be a fixed real number and let $\vartheta_0, \ldots, \vartheta_{s-1} \in \mathbb{F}_p^*$ be pairwise distinct elements of prime multiplicative order $l > p^{\varepsilon}$ modulo p. Then*

$$\max_{\gcd(h_0,\ldots,h_{s-1},p)=1} \left| \sum_{a \in \mathbb{Z}/l} \mathbf{e}_p\left(h_0\vartheta_0^a + \cdots + h_{s-1}\vartheta_{s-1}^a\right) \right| \ll l^{1-\delta},$$

where $\delta > 0$ depends only on s and $\varepsilon > 0$.

3. Main Result

Now we are prepared to prove our main result.

Theorem 3. *For any $\varepsilon > 0$ there is some $\eta > 0$ such that for $l > p^{\varepsilon}$ we have*

$$\frac{1}{l^n} \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} D_{l,p,s}(\mathbf{a}) \ll l^{-\eta}.$$

Proof. We may assume that p is large enough. From Lemma 1 and the triangle inequality we conclude that

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

where

$$\sigma(\mathbf{a}) = \frac{1}{2^n} \sum_{(h_0,\ldots,h_{s-1}) \in \mathbb{F}_p^s} \prod_{i=0}^{s-1} \frac{1}{|h_i|+1} \left| \sum_{x=0}^{2^n-s} \mathbf{e}_p\left(\sum_{i=0}^{s-1} h_i f_{\mathbf{a}}(x+i)\right) \right|.$$

Writing $|h_i|+1 = (|h_i|+1)^{1/2}(|h_i|+1)^{1/2}$, $i = 0, \ldots, s-1$, and using the Cauchy inequality again we derive

$$\begin{aligned}
\sigma(\mathbf{a})^2 &\le \frac{1}{2^{2n}} \sum_{(j_0,\ldots,j_{s-1}) \in \mathbb{F}_p^s} \prod_{i=0}^{s-1} \frac{1}{|j_i|+1} \\
&\qquad \sum_{(h_0,\ldots,h_{s-1}) \in \mathbb{F}_p^s} \prod_{i=0}^{s-1} \frac{1}{|h_i|+1} \left| \sum_{x=0}^{2^n-s} \mathbf{e}_p\left(\sum_{i=0}^{s-1} h_i f_{\mathbf{a}}(x+i)\right) \right|^2 \\
&\le \frac{(\log p)^s}{2^{2n}} \sum_{(h_0,\ldots,h_{s-1}) \in \mathbb{F}_p^s} \prod_{i=0}^{s-1} \frac{1}{|h_i|+1} \left| \sum_{x=0}^{2^n-s} \mathbf{e}_p\left(\sum_{i=0}^{s-1} h_i f_{\mathbf{a}}(x+i)\right) \right|^2.
\end{aligned}$$

Recalling (2), we obtain

$$\begin{aligned}
&\frac{1}{l^n} \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} D_{l,p,s}(\mathbf{a})^2 \\
&\qquad \ll \frac{1}{p^2} + \frac{(\log p)^s}{l^n 2^{2n}} \sum_{(h_0,\ldots,h_{s-1}) \in \mathbb{F}_p^s} \prod_{i=0}^{s-1} \frac{1}{|h_i|+1} W(h_0, \ldots, h_{s-1}),
\end{aligned} \tag{3}$$

where

$$W(h_0, \ldots, h_{s-1}) = \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} \left| \sum_{x=0}^{2^n-s} \mathbf{e}_p\left(\sum_{i=0}^{s-1} h_i f_{\mathbf{a}}(x+i)\right) \right|^2.$$

We now fix some real parameter $\gamma > 0$ and set $r = \lfloor \gamma n \rfloor$. For an n-bit integer $x = x_1 \dots x_n$ we now define the vectors

$$u(x) = x_1 \dots x_{n-r} \qquad \text{and} \qquad v(x) = x_{n-r+1} \dots x_n.$$

Let $\mathcal{Z}_r$ be the set of $x \in [0, 2^n - s]$ for which $u(x) \neq u(x+i)$ for at least one $i = 1, \dots, s-1$ and let $\mathcal{X}_r$ be the set of the remaining integers from the same interval $[0, 2^n - s]$. Thus, by the Cauchy inequality

$$W(h_0, \dots, h_{s-1}) \ll \widetilde{W}(h_0, \dots, h_{s-1}) + l^n (\#\mathcal{Z}_r)^2, \tag{4}$$

where

$$\widetilde{W}(h_0, \dots, h_{s-1}) = \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} \left| \sum_{x \in \mathcal{X}_r} \mathbf{e}_p \left(\sum_{i=0}^{s-1} h_i f_{\mathbf{a}}(x+i) \right) \right|^2.$$

Assume that

$$r > \lceil \log s \rceil + 1. \tag{5}$$

Then we have

$$\#\mathcal{Z}_r \leq s 2^{n-r+1}. \tag{6}$$

To see this, consider $w = x_{n-r+1} \dots x_{n-\lceil \log s \rceil + 1}$, then $u(x) \neq u(x+i)$ possibly occurs only when all the components of w are 1, that is, $x_{n-r+1} = \dots = x_{\lceil \log s \rceil + 1} = 1$. Therefore, $\#\mathcal{Z}_r \leq 2^{n-r+\lceil \log s \rceil}$, which implies (6).

We recall that $|z|^2 = z\overline{z}$ for any complex z and that $\overline{\mathbf{e}_p(a)} = \mathbf{e}_p(-a)$ 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

$$\widetilde{W}(h_0, \dots, h_{s-1}) = \sum_{x, y \in \mathcal{X}_r} \sum_{\mathbf{a} \in (\mathbb{Z}/l)^n} \mathbf{e}_p \left(\sum_{i=0}^{s-1} h_i (f_{\mathbf{a}}(x+i) - f_{\mathbf{a}}(y+i)) \right).$$

Note that for $x \in \mathcal{X}_r$ we have

$$\sum_{i=0}^{s-1} h_i f_{\mathbf{a}}(x+i) = \sum_{i=0}^{s-1} h_i g^{a_1^{u_1} \dots a_{n-r}^{u_{n-r}} a_{n-r+1}^{v_{i,1}} \dots a_n^{v_{i,r}}}, \tag{7}$$

where

$$u(x) = \dots = u(x+s-1) = u_1 \dots u_{n-r},$$

and

$$v(x+i) = v_{i,1} \dots v_{i,r}, \qquad i = 0, \dots, s-1.$$

If $u(x) = u(y)$ we estimate the inner sum trivially as l^n. There are $O(2^{n+r})$ such pairs of x and y.

If $u(x) \neq u(y)$, we fix any index ν, $1 \leq \nu \leq n-r$, with $x_\nu = 1$, $y_\nu = 0$ or with $x_\nu = 0$, $y_\nu = 1$. Furthermore, changing the names of variables permuting the indices, and changing the roles of x and y we may assume that $\nu = 1$, $x_1 = 1$, $y_1 = 0$.

We see that the term $f_{\mathbf{a}}(y)$ does not depend on a_1. Therefore, we see from (7) that in this case

$$\left|\sum_{\mathbf{a}\in(\mathbb{Z}/l)^n} \mathbf{e}_p\left(\sum_{i=0}^{s-1} h_i(f_{\mathbf{a}}(x+i)-f_{\mathbf{a}}(y+i))\right)\right| \le \sum_{a_2,\dots,a_n\in(\mathbb{Z}/l)^{n-1}} \left|\sum_{a\in\mathbb{Z}/l} \mathbf{e}_p\left(\sum_{i=0}^{s-1} h_i\vartheta_i(a_2,\dots,a_n)^a\right)\right|, \tag{8}$$

where

$$\vartheta_i(a_2,\dots,a_n) = g^{a_2^{u_2}\dots a_{n-r}^{u_{n-r}} a_{n-r+1}^{v_{i,1}}\dots a_n^{v_{i,r}}}, \qquad i=0,\dots,s-1.$$

Clearly, for $x\in\mathcal{X}_r$, the binary vectors $v(x+i)$, $i=0,\dots,s-1$, are pairwise distinct (as $u(x)=\dots=u(x+s-1)$). Hence the products

$$a_{n-r+1}^{v_{i,1}}\dots a_n^{v_{i,r}}, \qquad i=0,\dots,s-1, \tag{9}$$

are pairwise distinct nonzero elements of $\mathbb{Z}/l$ for all but $O(q^{r-1})$ values of $(a_{n-r+1},\dots,a_n)$. If the products (9) are pairwise distinct, then we apply Lemma 2. Otherwise we estimate the sum over a on the right-hand side of (8) trivially as l. Hence, if $u(x)\ne u(y)$, then

$$\left|\sum_{\mathbf{a}\in(\mathbb{Z}/l)^n} \mathbf{e}_p\left(\sum_{i=0}^{s-1} h_i(f_{\mathbf{a}}(x+i)-f_{\mathbf{a}}(y+i))\right)\right| \ll l^{n-1}\cdot l^{1-\delta} + l^{n-2}\cdot l \ll l^{n-\delta}$$

(we can certainly assume that $\delta<1$). Therefore,

$$\widetilde{W}(h_0,\dots,h_{s-1}) \le 2^{n+r}l^n + 2^{2n}l^{n-\delta},$$

which together with (6), after substitution in (4), implies the bound:

$$W(h_0,\dots,h_{s-1}) \ll 2^{n+r}l^n + 2^{2n}l^{n-\delta} + 2^{2n-2r}l^n.$$

Taking r to satisfy the inequality,

$$2^{2r} \le l^{\delta} < 2^{2(r+1)},$$

thus for a sufficiently large p the inequality (5) holds, we derive

$$W(h_0,\dots,h_{s-1}) \ll 2^n l^{n+\delta/2} + 2^{2n}l^{n-\delta}.$$

Without loss of generality, we can assume that $\delta<2/3$. Then, since $l<p\le 2^n$, we have

$$2^n l^{n+\delta/2} \ll 2^{2n}l^{n-\delta}.$$

Hence

$$W(h_0,\dots,h_{s-1}) \ll 2^{2n}l^{n-\delta},$$

which after inserting in (3), yields

$$\frac{1}{l^n}\sum_{\mathbf{a}\in(\mathbb{Z}/l)^n} D_{l,p,s}(\mathbf{a})^2 \ll p^{-2} + l^{n-\delta}(\log p)^{2s},$$

which together with the Cauchy inequality concludes the proof. □

4. Remarks

It is certainly interesting to get an explicit form of the bound of Theorem 3, even only for reasonably large values of ε (say, for $\varepsilon \geq 2/3$). In turn this leads to the question of obtaining an explicit form of Lemma 2, which is an important (and quite feasible) question of independent interest.

References

[1] W. Banks, F. Griffin, D. Lieman and I. E. Shparlinski, *Non-linear complexity of the Naor–Reingold pseudo-random function*, Proc. the 2nd Intern. Conf. on Information Security and Cryptology, Seoul, 1999, Lect. Notes in Comp. Sci., 1787, Springer-Verlag, Berlin, 2000, 53–59.

[2] J. Bourgain, *Mordell's exponential sum estimate revisited*, J. Amer. Math. Soc. **18** (2005), no. 2, 477–499, DOI 10.1090/S0894-0347-05-00476-5. MR2137982 (2006b:11099)

[3] Marcos Cruz, Domingo Gómez, and Daniel Sadornil, *On the linear complexity of the Naor-Reingold sequence with elliptic curves*, Finite Fields Appl. **16** (2010), no. 5, 329–333, DOI 10.1016/j.ffa.2010.05.005. MR2678621 (2011g:14062)

[4] Michael Drmota and Robert F. Tichy, *Sequences, discrepancies and applications*, Lecture Notes in Mathematics, vol. 1651, Springer-Verlag, Berlin, 1997. MR1470456 (98j:11057)

[5] Domingo Gómez, Jaime Gutierrez, and Álvar Ibeas, *On the linear complexity of the Naor-Reingold sequence*, Inform. Process. Lett. **111** (2011), no. 17, 854–856, DOI 10.1016/j.ipl.2011.05.017. MR2849394 (2012e:11145)

[6] F. Griffin and I. E. Shparlinski, *On the linear complexity of the Naor–Reingold pseudo-random function*, Proc. 2nd Intern. Conf. on Information and Communication Security, Sydney, 1999, Lect. Notes in Comp. Sci., 1726, Springer-Verlag, Berlin, 1999, 301–308.

[7] 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.

[8] Harald Niederreiter, *Quasi-Monte Carlo methods and pseudo-random numbers*, Bull. Amer. Math. Soc. **84** (1978), no. 6, 957–1041, DOI 10.1090/S0002-9904-1978-14532-7. MR508447 (80d:65016)

[9] Igor E. Shparlinski, *On the Naor-Reingold pseudo-random function from elliptic curves*, Appl. Algebra Engrg. Comm. Comput. **11** (2000), no. 1, 27–34, DOI 10.1007/s002000000023. MR1817696 (2001m:65015)

[10] Igor E. Shparlinski, *Linear complexity of the Naor-Reingold pseudo-random function*, Inform. Process. Lett. **76** (2000), no. 3, 95–99, DOI 10.1016/S0020-0190(00)00133-2. MR1801380

[11] Igor E. Shparlinski, *On the uniformity of distribution of the Naor-Reingold pseudo-random function*, Finite Fields Appl. **7** (2001), no. 2, 318–326, DOI 10.1006/ffta.2000.0291. MR1826340 (2002b:11104)

[12] Igor E. Shparlinski and Joseph H. Silverman, *On the linear complexity of the Naor-Reingold pseudo-random function from elliptic curves*, Des. Codes Cryptogr. **24** (2001), no. 3, 279–289, DOI 10.1023/A:1011223204345. MR1857142 (2002j:11087)

Division of Mathematical Sciences, School of Physical & Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore

E-mail address: lingsan@ntu.edu.sg

Department of Computing, Macquarie University, Sydney NSW 2109, Australia

E-mail address: igor.shparlinski@mq.edu.au

Division of Mathematical Sciences, School of Physical & Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore

E-mail address: hxwang@ntu.edu.sg