Distribution Properties of Binary Sequences Derived from Primitive Sequences Modulo Square-free Odd Integers

Qun-Xiong Zheng[1,2,3(✉)], Dongdai Lin[1], and Wen-Feng Qi[2,3]

[1] State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China
qunxiong_zheng@163.com, ddlin@iie.ac.cn

[2] State Key Laboratory of Cryptology, P.O. Box 5159, Beijing 100878, China
wenfeng.qi@263.net

[3] National Digital Switching System Engineering and Technological Research Center, Zhengzhou 450001, China

Abstract. Recently, a class of nonlinear sequences, modular reductions of primitive sequences over integer residue rings, was proposed and has attracted much attention. In particular, modulo 2 reductions of primitive sequences over $\mathbf{Z}/(2^{31}-1)$ were used in the ZUC algorithm. In this paper, we study the distribution properties of modulo 2 reductions of primitive sequences over $\mathbf{Z}/(M)$, where M is a square-free odd integer. Let $\underline{a}$ be a primitive sequence of order n over $\mathbf{Z}/(M)$ with period T and $[\underline{a}]_{\operatorname{mod} 2}$ the modulo 2 reduction of $\underline{a}$. With the estimate of exponential sums over $\mathbf{Z}/(M)$, the proportion f_s of occurrences of s within a segment of $[\underline{a}]_{\operatorname{mod} 2}$ of length μT is estimated, where $s \in \{0,1\}$ and $0 < \mu \leq 1$. Based on this estimate, it is further shown that for given M and μ, f_s tends to $\frac{M+1-2s}{2M}$ as $n \to \infty$. This result implies that there exists a small imbalance between 0 and 1 in $[\underline{a}]_{\operatorname{mod} 2}$, which should be taken into full consideration in the design of stream ciphers based on $[\underline{a}]_{\operatorname{mod} 2}$.

Keywords: Integer residue ring · Primitive sequence · Modular reduction · 0, 1 distribution · ZUC algorithm

1 Introduction

For an integer $m \geq 2$, let $\mathbf{Z}/(m)$ denote the integer residue ring modulo m. The set $\{0, 1, \ldots, m-1\}$ is always chosen as the complete set of representatives for

This work was supported by NSF of China (Nos. 61872383, 61402524, 61872359 and 61602510). The work of Qun-Xiong Zheng was also supported by Young Elite Scientists Sponsorship Program by CAST (2016QNRC001) and by National Postdoctoral Program for Innovative Talents (BX201600188) and by China Postdoctoral Science Foundation funded project (2017M611035).

F. Guo et al. (Eds.): Inscrypt 2018, LNCS 11449, pp. 568–585, 2019.
https://doi.org/10.1007/978-3-030-14234-6_30

the elements of the ring $\mathbf{Z}/(m)$. Thus a sequence $\underline{a}$ over $\mathbf{Z}/(m)$ is usually seen as an integer sequence over $\{0, 1, \ldots, m-1\}$. Moreover, for an integer a and a positive integer $b \geq 2$, let us denote the least nonnegative residue of a modulo b by $[a]_{\bmod b}$, and similarly, for a sequence $\underline{a} = (a(t))_{t\geq 0}$ over $\mathbf{Z}/(m)$, denote $[\underline{a}]_{\bmod b} = ([a(t)]_{\bmod b})_{t\geq 0}$.

Let p be a prime number and e a positive integer. During the past two decades, the maximal period linear recurring sequences over $\mathbf{Z}/(p^e)$, called primitive sequences over $\mathbf{Z}/(p^e)$, have been paid much attention. An enormous amount of effort is directed toward the study of finding useful mappings to derive good pseudorandom sequences from primitive sequences over $\mathbf{Z}/(p^e)$, which are called compression mappings in literature, and proving that they are injective. Generally there are two kinds of compression mappings: one is based on e-variable functions over $\mathbf{Z}/(p)$ [10,15–17,20,21]; the other is based on the modular arithmetic [13,22]. Besides, the pseudorandom properties of these compression sequences are also extensively studied, such as periodicity [7,13], linear complexity [3,6,15] and distribution properties [2,8,12,23].

Recently research interests on primitive sequences over $\mathbf{Z}/(p^e)$ are further extended to primitive sequences over $\mathbf{Z}/(M)$ [4,9,24–27], where M is a square-free odd integer. One of important reasons for this is that the period of a primitive sequence $\underline{a}$ of order n over $\mathbf{Z}/(p^e)$ is undesirable if $e \geq 2$. Recall that the period $per(\underline{a})$ of a primitive sequence $\underline{a}$ of order n over $\mathbf{Z}/(p^e)$ is equal to $p^{e-1} \cdot (p^n - 1) \approx p^{e+n-1}$ [18]. It can be seen that for a fixed prime power p^e with $e \geq 2$, the period $per(\underline{a})$ increases slowly and far less than $p^{e\cdot n}$ as n increases. Therefore, to meet the requirement of long period in practical applications, n should be chosen large enough, which will be high resource consumption in hardware and software implementation. For example, to generate a sequence with period not less than 2^{64} over $\mathbf{Z}/(2^8)$, $\mathbf{Z}/(2^{16})$ and $\mathbf{Z}/(2^{32})$, the number of bit-registers required must be larger than 456, 784 and 1056, respectively. However for many choices of M, primitive sequences over $\mathbf{Z}/(M)$ have no such periodic weakness. For cryptographic applications, the moduli of the form $2^e - 1$ have attracted much attention since the operation "$\bmod 2^e - 1$" can be efficiently implemented both in hardware and software, and this offers new possibilities for advancement in the solution of applying linear recurring sequences over integer residue rings. For instance, primitive sequences over $\mathbf{Z}/(2^{31} - 1)$ are used to design the ZUC algorithm, a stream cipher that is the core of the standardised 3GPP confidentiality algorithm 128-EEA3 and the 3GPP integrity algorithm 128-EIA3, see [28].

By applying the operation $\bmod 2$ to primitive sequences over $\mathbf{Z}/(M)$, one can easily obtain a class of binary sequences, called modulo 2 reductions of primitive sequences over $\mathbf{Z}/(M)$. It is thought that the operation $\bmod 2$ destroys the original linear recurrence relation of primitive sequences over $\mathbf{Z}/(M)$ and the obtained binary sequences should have many desirable cryptographic properties if the modulus M and the order n are carefully chosen. One of the most interesting properties is the so-called "modulo 2 distinctness". Some progress has been made on the modulo 2 distinctness, see, for example, [9,26]. From the viewpoint of cryptographic applications, it is naturally interested in the pseudorandom properties of modulo 2 reductions of primitive sequences over $\mathbf{Z}/(M)$. However, so far few result was obtained. In [25], to study the modulo 2 distinctness of primitive sequences over $\mathbf{Z}/(M)$, two distribution properties of primitive sequences over $\mathbf{Z}/(M)$ are investigated. One is to determine whether there is an integer $t \geq 0$ such that $a(t) = s$ for a given element $s \in \mathbf{Z}/(M)$ and a given primitive sequence $\underline{a}$ of order n over $\mathbf{Z}/(M)$. The other is to determine whether there is an integer $t \geq 0$ such that $a(t)$ is an even number for a given primitive sequence $\underline{a}$ of order 1 over $\mathbf{Z}/(M)$. In [9], Hu and Wang studied whether there is an integer $t \geq 0$ such that $a(t) = a$ and $b(t) = b$, for two given elements $a, b \in \mathbf{Z}/(M)$ and two given primitive sequences $\underline{a}, \underline{b}$ generated by a same primitive polynomial over $\mathbf{Z}/(M)$.

In this paper, we study the distribution properties of the binary sequence $[\underline{a}]_{\bmod 2}$, where $\underline{a}$ is a primitive sequence of order n over $\mathbf{Z}/(M)$ with period T. With the estimate of exponential sums over $\mathbf{Z}/(M)$, the proportion f_s of occurrences of s within a segment of $[\underline{a}]_{\bmod 2}$ of length μT is estimated, where $s \in \{0, 1\}$ and $0 < \mu \leq 1$. Based on this estimate, it is further shown that for given M and μ, f_s tends to $(M + 1 - 2s)/2M$ as $n \to \infty$. Generally speaking, if n is not too small (for example, $n \geq 3$ for $M = 2^{32} - 1$), then the value of f_s is very close to that of $(M + 1 - 2s)/2M$. This implies that there always exists a small imbalance (about $1/M$) between 0 and 1 in $[\underline{a}]_{\bmod 2}$. In order to provide a good resistance against the distinguishing attacks, such imbalance should be taken into full consideration in the design of stream ciphers based on $[\underline{a}]_{\bmod 2}$. Fortunately, by introducing a moderate amount of exclusive or operations, the imbalance of $0, 1$ will be reduced to a small enough extent.

The rest of this paper is organized as follows. Section 2 presents some necessary preliminaries. Section 3 gives the main results of this paper. Finally, conclusions are drawn in Sect. 4.

2 Preliminaries

2.1 Primitive Polynomials and Primitive Sequences over Integer Residue Rings

Let m be an integer greater than 1. If a sequence $\underline{a} = (a(t))_{t \geq 0}$ over $\mathbf{Z}/(m)$ satisfies

$$a(t) = [c_{n-1}a(t-1) + \cdots + c_1 a(t-n+1) + c_0 a(t-n)]_{\bmod m} \tag{1}$$

for all integers $t \geq n$, where n is a positive integer and $c_0, c_1, \ldots, c_{n-1} \in \mathbf{Z}/(m)$ are constant coefficients, then $\underline{a}$ is called a *linear recurring sequence* of order n over $\mathbf{Z}/(m)$ generated by $f(x) = x^n - c_{n-1}x^{n-1} - \cdots - c_0$ (or $\underline{a}$ is a sequence of order n over $\mathbf{Z}/(m)$ in short). For convenience, the set of sequences generated by $f(x)$ over $\mathbf{Z}/(m)$ is generally denoted by $G(f(x), m)$. Particular interests for cryptography are the maximal period linear recurring sequences also called primitive sequences over $\mathbf{Z}/(m)$, which are generated by primitive polynomials over $\mathbf{Z}/(m)$. Next we introduce the definitions of primitive polynomials and primitive sequences over $\mathbf{Z}/(m)$.

Let $f(x)$ be a monic polynomial of degree n over $\mathbf{Z}/(m)$. If $\gcd(f(0), m) = 1$, then there exists a positive integer T such that $x^T - 1$ is divisible by $f(x)$ in $\mathbf{Z}/(m)[x]$. The minimum of such T is called the period of $f(x)$ over $\mathbf{Z}/(m)$ and denoted by $per(f(x), m)$. For the case that m is a prime power, say $m = p^e$, it is known that $per(f(x), p^e) \leq p^{e-1}(p^n - 1)$, see [18]. If $per(f(x), p^e) = p^{e-1}(p^n - 1)$, then $f(x)$ is called a *primitive polynomial* of degree n over $\mathbf{Z}/(p^e)$. A sequence $\underline{a}$ over $\mathbf{Z}/(p^e)$ is called a *primitive sequence* of order n if $\underline{a}$ is generated by a primitive polynomial of degree n over $\mathbf{Z}/(p^e)$ and $[\underline{a}]_{\bmod p}$ is not an all-zero sequence. A primitive sequence $\underline{a}$ of order n over $\mathbf{Z}/(p^e)$ is (strictly) periodic and the period $per(\underline{a})$ is equal to $p^{e-1}(p^n - 1)$, see [18]. For the case of a general integer m, assume $m = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}$ is the canonical factorization of m. A monic polynomial $f(x)$ of degree n over $\mathbf{Z}/(m)$ is called a *primitive polynomial* if for every $k \in \{1, 2, \ldots, r\}$, $f(x)$ is a primitive polynomial of degree n over $\mathbf{Z}/(p_k^{e_k})$. A sequence $\underline{a}$ over $\mathbf{Z}/(m)$ is called a *primitive sequence* of order n if $\underline{a}$ is generated by a primitive polynomial of degree n over $\mathbf{Z}/(m)$ and $[\underline{a}]_{\bmod p_k}$ is not an all-zero sequence for every $k \in \{1, 2, \ldots, r\}$, that is, $[\underline{a}]_{\bmod p_k^{e_k}}$ is a primitive sequence of order n over $\mathbf{Z}/(p_k^{e_k})$. It can be seen that the period of a primitive polynomial of degree n over $\mathbf{Z}/(m)$ and that of a primitive sequence of order n over $\mathbf{Z}/(m)$ are both equal to

$$\operatorname{lcm}\left(p_1^{e_1-1}(p_1^n - 1), p_2^{e_2-1}(p_2^n - 1), \ldots, p_r^{e_r-1}(p_r^n - 1)\right).$$

For convenience, the set of primitive sequences generated by a primitive polynomial $f(x)$ over $\mathbf{Z}/(m)$ is generally denoted by $G'(f(x), m)$.

2.2 Exponential Sums over Integer Residue Rings

Let m be a positive integer greater than 1, and let $e_m(\cdot)$ be the canonical additive character over $\mathbf{Z}/(m)$ given by $e_m(a) = e^{2\pi i a/m}$, where a is an integer. For an integer c, it is well-known that

$$\sum_{a=0}^{m-1} e_m(ca) = \begin{cases} m, & \text{if } m \mid c; \\ 0, & \text{otherwise.} \end{cases}$$

The following Lemma 1 is cited from [5, Theorem 1].

Lemma 1. *([5, Theorem 1]) Let $D \geq 1$, $m \geq 2$ and $g = \gcd(m, D)$. Then we have*

$$\sum_{a=1}^{m-1} \left| \frac{\sin \pi a D/m}{\sin \pi a/m} \right| < \frac{4}{\pi^2} m \ln m + 0.38m + 0.608 + 0.116 \frac{g^2}{m},$$

where $\ln(m)$ *is the natural logarithm of* m.

The following Lemma 2 is an improvement of a well-known result of Korobov [14, Theorem 13].

Lemma 2. *Let $\underline{a}$ be a primitive sequence of order n over $\mathbf{Z}/(m)$ with period T. Then for any integer h we have*

$$\left|\sum_{t=0}^{T-1} e_m\left(a\left(t\right)\right) e_T\left(ht\right)\right| \leq m^{\frac{n}{2}}. \tag{2}$$

In particular,

$$\left|\sum_{t=0}^{T-1} e_m\left(a\left(t\right)\right)\right| \leq m^{\frac{n}{2}}.$$

Moreover, we have

$$\left|\sum_{t=k}^{k+L-1} e_m\left(a\left(t\right)\right)\right| \leq m^{\frac{n}{2}}\left(\frac{4\ln T}{\pi^2}+0.409+\frac{L+1}{T}\right) \tag{3}$$

for any integer $k \geq 0$ and $0 < L < T$.

Proof. Since the inequality (2) has been proved in [14, Theorem 13], we only prove the inequality (3). We start from the identity

$$\sum_{t=k}^{k+L-1} e_m\left(a\left(t\right)\right) = \sum_{t=k}^{k+T-1} e_m\left(a\left(t\right)\right) \sum_{j=0}^{L-1} \frac{1}{T} \sum_{h=0}^{T-1} e_T\left(h\left(t-k-j\right)\right) \text{ for } k \geq 0 \text{ and } 0 < L < T,$$

which is valid since the sum over j is 1 for $k \leq t \leq k+L-1$ and 0 for $k+L \leq t \leq k+T-1$. Rearranging terms, we get

$$\sum_{t=k}^{k+L-1} e_m\left(a\left(t\right)\right) = \frac{1}{T}\sum_{h=0}^{T-1}\left(\sum_{t=k}^{k+T-1} e_m\left(a\left(t\right)\right) e_T\left(ht\right)\right)\left(\sum_{j=0}^{L-1} e_T\left(-h\left(k+j\right)\right)\right),$$

and so we obtain

$$\begin{aligned}\left|\sum_{t=k}^{k+L-1} e_m\left(a\left(t\right)\right)\right| &\leq \frac{1}{T}\sum_{h=0}^{T-1}\left|\sum_{t=k}^{k+T-1} e_m\left(a\left(t\right)\right) e_T\left(ht\right)\right|\left|\sum_{j=0}^{L-1} e_T\left(-h\left(k+j\right)\right)\right| \\ &= \frac{1}{T}\sum_{h=0}^{T-1}\left|\sum_{t=0}^{T-1} e_m\left(a\left(t\right)\right) e_T\left(ht\right)\right|\left|\sum_{j=0}^{L-1} e_T\left(-hj\right)\right|.\end{aligned}$$

Then by the inequality (2) we get

$$\begin{aligned}\left|\sum_{t=k}^{k+L-1} e_m\left(a\left(t\right)\right)\right| &\leq \frac{m^{\frac{n}{2}}}{T}\sum_{h=0}^{T-1}\left|\sum_{j=0}^{L-1} e_T\left(-hj\right)\right| \\ &= \frac{m^{\frac{n}{2}}}{T}\sum_{h=0}^{T-1}\left|\sum_{j=0}^{L-1} e_T\left(hj\right)\right| \\ &\leq \frac{m^{\frac{n}{2}}}{T}\sum_{h=1}^{T-1}\left|\sum_{j=0}^{L-1} e_T\left(hj\right)\right| + \frac{L\cdot m^{\frac{n}{2}}}{T}.\end{aligned} \tag{4}$$

We note that

$$\begin{aligned}\sum_{h=1}^{T-1}\left|\sum_{j=0}^{L-1} e_T\left(hj\right)\right| &= \sum_{h=1}^{T-1}\left|\frac{e_T(hL)-1}{e_T(h)-1}\right| \\ &= \sum_{h=1}^{T-1}\left|\frac{\cos(2\pi hL/T)+i\sin(2\pi hL/T)-1}{\cos(2\pi h/T)+i\sin(2\pi h/T)-1}\right| \\ &= \sum_{h=1}^{T-1}\left|\frac{-2\sin^2(\pi hL/T)+2i\sin(\pi hL/T)\cos(\pi hL/T)}{-2\sin^2(\pi h/T)+2i\sin(\pi h/T)\cos(\pi h/T)}\right| \\ &= \sum_{h=1}^{T-1}\left|\frac{\sin(\pi hL/T)}{\sin(\pi h/T)}\right|,\end{aligned}$$

and so an application of Lemma 3 yields

$$\begin{aligned}\sum_{h=1}^{T-1}\left|\sum_{j=0}^{L-1} e_T\left(hj\right)\right| &< \frac{4}{\pi^2}T\ln T+0.38T+0.608+0.116\frac{g^2}{T} \\ &< \frac{4}{\pi^2}T\ln T+0.409T+1, \qquad (5)\end{aligned}$$

where the last inequality follows from the fact that $g=\gcd(L,T)<T/2$. Combining the inequalities (4) and (5), we get the desired result.

3 Main Results

Throughout the rest of this paper, we always assume that $M>1$ is a given square-free odd integer and $M=p_1p_2\cdots p_r$ is the canonical factorization of M. Let $d>1$ be a divisor of M. Suppose $d=p_{i_1}\cdots p_{i_k}$ with $1\le i_1<\cdots<i_k\le r$. Let us denote by $\lambda_n(d)$ the period of primitive sequences of order n over $\mathbf{Z}/(d)$, that is,

$$\lambda_n(d)=\mathrm{lcm}(p_{i_1}^n-1,\ldots,p_{i_k}^n-1).$$

The main results of this paper are stated in the following Theorems 1 and 2.

Theorem 1. *Let $\underline{a}$ be a primitive sequence of order n over $\mathbf{Z}/(M)$ with period $T=\lambda_n\left(M\right)$, and let $\underline{b}=\left[\underline{a}\right]_{mod\,2}$. For $s\in\{0,1\}$, denote by $N\left(\underline{b}^T,s\right)$ the number of t, $0\le t\le T-1$, with $b\left(t\right)=s$. Then we have*

$$\frac{N\left(\underline{b}^T,s\right)}{T}=\frac{1}{2}+\left(-1\right)^s\cdot\frac{M^{n-1}-1}{2M^n-2} \qquad (6)$$

if M is an odd prime number; and

$$\left|\frac{N\left(\underline{b}^T,s\right)}{T}-\frac{M+1-2s}{2M}\right|<\frac{1}{M}\sum_{\substack{d|M\\ d>1}}\frac{d^{n/2}}{\lambda_n\left(d\right)}\cdot\left(\frac{d\ln d}{\pi}+0.538d\right) \qquad (7)$$

if M has at least two different prime divisors.

Theorem 2. *Let $\underline{a}$ be a primitive sequence of order n over $\mathbf{Z}/(M)$ with period $T = \lambda_n(M)$. Let $\underline{b} = [\underline{a}]_{mod\,2}$ and $\underline{b}^L = (b(k), b(k+1), \ldots, b(k+L-1))$ a segment of $\underline{b}$ with length $L = \mu T$, where $0 \le k \le T-1$ and $0 < \mu < 1$. For $s \in \{0,1\}$, denote by $N\left(\underline{b}^L, s\right)$ the number of t, $0 \le t \le L-1$, with $b(k+t) = s$. Then we have*

$$\left|\frac{N\left(\underline{b}^L, s\right)}{L} - \frac{M+1-2s}{2M}\right| < \frac{1}{M}\sum_{\substack{d|M\\ d>1}} \frac{d^{\frac{n}{2}}}{\lambda_n(d)} C(d, L),$$

where

$$C(d, L) = \left(\frac{\lambda_n(d)}{L}\left(\frac{4}{\pi^2}\ln\lambda_n(d) + 0.409\right) + \frac{L+1}{L}\right)\left(\frac{d\ln d}{\pi} + 0.538d\right).$$

The rest of this section is divided into three subsections. Subsects. 3.1 and 3.2 are mainly devoted to the proof of Theorem 1 and the proof of Theorem 2, respectively. Finally, as an example, an application of Theorems 1 and 2 to the modulo 2 reductions of primitive sequences over $\mathbf{Z}/(2^{32}-1)$ is given in Subsect. 3.3.

3.1 The Proof of Theorem 1

We first collect two well-known results on trigonometric functions in Lemma 3. The first result can be found in [19] and the second result can be found in [11, p. 447].

Lemma 3. *Let $\tan x = \sin x/\cos x$, $\sec x = 1/\cos x$ and $\csc x = 1/\sin x$ be the tangent function, the secant function and the cosecant function, respectively. Then we have:*

(1) $\int \sec x dx = \ln|\sec x + \tan x| + C$, *where C is the constant of integration;*
(2) $\csc(\pi/m) \le m/3$ *if $m \ge 6$.*

Lemma 4. *For an odd integer $m > 1$, we have*

$$\sum_{h=1}^{m-1}\left|\sec\frac{h\pi}{m}\right| < \frac{2}{\pi}m\ln m + 1.076m.$$

Proof. It can be directly verified that the lemma holds for $m = 3$ or 5. Therefore, we assume that $m \ge 7$. Since it is clear that

$$\sec\frac{h\pi}{m} = -\sec\frac{(m-h)\cdot\pi}{m} > 0 \text{ for } 1 \le h \le \frac{m-1}{2},$$

we obtain

$$\sum_{h=1}^{m-1}\left|\sec\frac{h\pi}{m}\right| = 2\sum_{h=1}^{\frac{m-1}{2}}\sec\frac{h\pi}{m}. \tag{8}$$

Note that the convexity of the function $\sec x$ implies that

$$\int_{u-\frac{\theta}{2}}^{u+\frac{\theta}{2}} \sec x dx > \theta \cdot \sec u \text{ for } \frac{\theta}{2} < u < \frac{\pi}{2} - \frac{\theta}{2}.$$

Thus by taking $\theta = \frac{\pi}{m}$ we get

$$\begin{aligned}
\sum_{h=1}^{\frac{m-3}{2}} \sec \frac{h\pi}{m} &< \frac{m}{\pi} \sum_{h=1}^{\frac{m-3}{2}} \int_{\frac{h\pi}{m}-\frac{\pi}{2m}}^{\frac{h\pi}{m}+\frac{\pi}{2m}} \sec x dx \\
&= \frac{m}{\pi} \int_{\frac{\pi}{2m}}^{\frac{(m-2)\cdot\pi}{2m}} \sec x dx \\
&< \frac{m}{\pi} \int_{0}^{\frac{(m-2)\cdot\pi}{2m}} \sec x dx \\
&= \frac{m}{\pi} \cdot \ln \left(\sec \frac{(m-2)\cdot\pi}{2m} + \tan \frac{(m-2)\cdot\pi}{2m} \right) \\
&= \frac{m}{\pi} \cdot \left(\ln \left(\sec \frac{(m-2)\cdot\pi}{2m} \right) + \ln \left(1 + \sin \frac{(m-2)\cdot\pi}{2m} \right) \right) \\
&< \frac{m}{\pi} \cdot \left(\ln \left(\csc \frac{\pi}{m} \right) + \ln 2 \right). \qquad (9)
\end{aligned}$$

By combining (8) and (9), we obtain

$$\begin{aligned}
\sum_{h=1}^{m-1} \left| \sec \frac{h\pi}{m} \right| &< \frac{2m}{\pi} \cdot \left(\ln \left(\csc \frac{\pi}{m} \right) + \ln 2 \right) + 2 \sec \frac{(m-1)\,\pi}{2m} \\
&= \frac{2m}{\pi} \cdot \left(\ln \left(\csc \frac{\pi}{m} \right) + \ln 2 \right) + 2 \csc \frac{\pi}{2m}. \qquad (10)
\end{aligned}$$

By applying $\csc(\pi/m) \leq m/3$ to the right-hand side of (10) we get

$$\sum_{h=1}^{m-1} \left| \sec \frac{h\pi}{m} \right| < \frac{2}{\pi} m \ln m + \frac{2m}{\pi} (\ln 2 - \ln 3) + \frac{4}{3} m < \frac{2}{\pi} m \ln m + 1.076m.$$

This completes the proof.

Now we start to prove Theorem 1.

Proof (Proof of Theorem 1). If M is an odd prime number, then (6) immediately follows from the theory of m-sequences over finite fields (see, for example, [11]). Next we will prove the equality (7). Note that

$$\left| \frac{N\left(\underline{b}^T, 0\right)}{T} - \frac{M+1}{2M} \right| = \left| 1 - \frac{N\left(\underline{b}^T, 1\right)}{T} - \frac{M+1}{2M} \right| = \left| \frac{N\left(\underline{b}^T, 1\right)}{T} - \frac{M-1}{2M} \right|,$$

and so it suffices to show (7) holds for the case that $s=0$, that is,

$$\left|\frac{N\left(\underline{b}^T,0\right)}{T}-\frac{M+1}{2M}\right|<\frac{1}{M}\sum_{\substack{d|M\\d>1}}\frac{d^{n/2}}{\lambda_n\left(d\right)}\cdot\left(\frac{d\ln d}{\pi}+0.538d\right). \tag{11}$$

Since

$$\begin{aligned}N\left(\underline{b}^T,0\right)&=\sum_{t=0}^{T-1}\sum_{x=0}^{\frac{M-1}{2}}\left(\frac{1}{M}\sum_{h=0}^{M-1}e_M\left(h\left(a\left(t\right)-2x\right)\right)\right)\\&=\frac{1}{M}\sum_{h=0}^{M-1}\left(\sum_{t=0}^{T-1}e_M\left(ha\left(t\right)\right)\cdot\sum_{x=0}^{\frac{M-1}{2}}e_M\left(-2hx\right)\right)\\&=\frac{T\cdot\left(M+1\right)}{2M}+\frac{1}{M}\sum_{h=1}^{M-1}\left(\sum_{t=0}^{T-1}e_M\left(ha\left(t\right)\right)\cdot\sum_{x=0}^{\frac{M-1}{2}}e_M\left(-2hx\right)\right),\end{aligned}$$

we get

$$\left|\frac{N\left(\underline{b}^T,0\right)}{T}-\frac{M+1}{2M}\right|\le\frac{1}{MT}\sum_{h=1}^{M-1}\left|\sum_{t=0}^{T-1}e_M\left(ha\left(t\right)\right)\right|\cdot\left|\sum_{x=0}^{\frac{M-1}{2}}e_M\left(-2hx\right)\right|. \tag{12}$$

We note that

$$\begin{aligned}\sum_{x=0}^{\frac{M-1}{2}}e_M\left(-2hx\right)&=\frac{\left(e^{-\frac{4h\pi i}{M}}\right)^{\frac{M+1}{2}}-1}{e^{-\frac{4h\pi i}{M}}-1}\\&=\frac{e^{-\frac{2h\pi i}{M}}-1}{e^{-\frac{4h\pi i}{M}}-1}\\&=\frac{1}{e^{-\frac{2h\pi i}{M}}+1}\\&=\frac{1}{\cos\frac{2h\pi}{M}-i\sin\frac{2h\pi}{M}+1}\\&=\frac{1}{2\cos^2\frac{h\pi}{M}-2i\sin\frac{h\pi}{M}\cos\frac{h\pi}{M}}\\&=\frac{1}{2}\cdot\frac{\sec\frac{h\pi}{M}}{e^{-\frac{h\pi i}{M}}}.\end{aligned} \tag{13}$$

Applying (13) to (12) we obtain

$$\begin{aligned}\left|\frac{N\left(\underline{b}^T,0\right)}{T}-\frac{M+1}{2M}\right| &\le \frac{1}{2MT}\sum_{h=1}^{M-1}\left|\sum_{t=0}^{T-1}e_M\left(ha\left(t\right)\right)\right|\cdot\left|\sec\frac{h\pi}{M}\right| \\ &= \frac{1}{2MT}\sum_{\substack{d\mid M\\ d>1}}\sum_{\substack{1\le h\le M-1\\ \gcd(h,M)=M/d}}\left|\sum_{t=0}^{T-1}e_M\left(ha\left(t\right)\right)\right|\cdot\left|\sec\frac{h\pi}{M}\right| \\ &= \frac{1}{2MT}\sum_{\substack{d\mid M\\ d>1}}\sum_{\substack{1\le h\le d-1\\ \gcd(h,d)=1}}\left|\sum_{t=0}^{T-1}e_d\left(ha\left(t\right)\right)\right|\cdot\left|\sec\frac{h\pi}{d}\right|. \qquad (14)\end{aligned}$$

Note that given a divisor $d>1$ of M, $[h\underline{a}]_{\bmod d}$ is a primitive sequence over $\mathbf{Z}/(d)$ with period $\lambda_n\left(d\right)$ for every integer h coprime with d, and so it follows from Lemma 2 that

$$\left|\sum_{t=0}^{T-1}e_d\left(ha\left(t\right)\right)\right| = \left|\frac{T}{\lambda_n\left(d\right)}\cdot\sum_{t=0}^{\lambda_n(d)-1}e_d\left(ha\left(t\right)\right)\right| \le \frac{T\cdot d^{n/2}}{\lambda_n\left(d\right)}. \qquad (15)$$

Combining (14) and (15) yields

$$\begin{aligned}\left|\frac{N\left(\underline{b}^T,0\right)}{T}-\frac{M+1}{2M}\right| &\le \frac{1}{2MT}\sum_{\substack{d\mid M\\ d>1}}\sum_{\substack{1\le h\le d-1\\ \gcd(h,d)=1}}\frac{T\cdot d^{n/2}}{\lambda_n\left(d\right)}\cdot\left|\sec\frac{h\pi}{d}\right| \\ &= \frac{1}{2M}\sum_{\substack{d\mid M\\ d>1}}\frac{d^{n/2}}{\lambda_n\left(d\right)}\cdot\sum_{\substack{1\le h\le d-1\\ \gcd(h,d)=1}}\left|\sec\frac{h\pi}{d}\right| \\ &\le \frac{1}{2M}\sum_{\substack{d\mid M\\ d>1}}\frac{d^{n/2}}{\lambda_n\left(d\right)}\cdot\sum_{h=1}^{d-1}\left|\sec\frac{h\pi}{d}\right|, \qquad (16)\end{aligned}$$

and so (11) follows from (16) and Lemma 4.

Generally speaking, if n is sufficiently large, then the right-hand side of (6) is sufficiently small, and so the value of $N\left(\underline{b}^T,s\right)/T$ is very close to that of $\left(M+1-2s\right)/2M$ (for more details, see Table 1). In fact, we can give a more theoretical result on the asymptotic property of $N\left(\underline{b}^T,s\right)/T$ as $n\to\infty$.

Corollary 1. *Let $\underline{a}$ be a primitive sequence of order n over $\mathbf{Z}/(M)$ with period $T=\lambda_n\left(M\right)$, and let $\underline{b}=[\underline{a}]_{mod\,2}$. Then for $s\in\{0,1\}$ we have*

$$\lim_{n\to\infty}\frac{N\left(\underline{b}^T,s\right)}{T}=\frac{M+1-2s}{2M}.$$

To prove Corollary 1, we first introduce a result of Bugeaud, Corvaja and Zannier [1].

Lemma 5. *([1, Theorem 1]) If $a < b$ are two integers greater than 1 which are multiplicatively independent (that is, the only integer solution (x, y) of the equation $a^x b^y = 1$ is $(x, y) = (0, 0)$), then for any given real number $\varepsilon > 0$, there exists an integer N_ε such that*

$$\gcd\left(a^n - 1, b^n - 1\right) < a^{n\varepsilon} \text{ for all integers } n > N_\varepsilon.$$

Remark 1. Note that a and b are multiplicatively independent if $\gcd(a, b) = 1$.

Proof (Proof of Corollary 1). Since Corollary 1 is obvious true for the case that M is an odd prime number, we assume that $M = p_1 p_2 \cdots p_r$ is the canonical factorization of M with $r \geq 2$ and $3 \leq p_1 < p_2 < \cdots < p_r$. Note that the inequality

$$\frac{d \ln d}{\pi} + 0.538d \leq \frac{M \ln M}{\pi} + 0.538M$$

holds for any divisor d of M, and so by Theorem 1 we get

$$\left| \frac{N\left(\underline{v}^T, s\right)}{T} - \frac{M + 1 - 2s}{2M} \right| < \left(\frac{M \ln M}{\pi} + 0.538M \right) \cdot \sum_{\substack{d \mid M \\ d > 1}} \frac{d^{n/2}}{\lambda_n(d)}.$$

Therefore to prove Corollary 1, it suffices to show that

$$\lim_{n \to \infty} \sum_{\substack{d \mid M \\ d > 1}} \frac{d^{n/2}}{\lambda_n(d)} = 0,$$

that is

$$\lim_{n \to \infty} \sum_{k=1}^{r} \sum_{1 \leq i_1 < \cdots < i_k \leq r} \frac{\prod_{j=1}^{k} p_{i_j}^{n/2}}{\operatorname{lcm}\left(p_{i_1}^n - 1, p_{i_2}^n - 1, \ldots, p_{i_k}^n - 1\right)} = 0. \tag{17}$$

Given a real number $\varepsilon > 0$. For any $1 \leq u < v \leq r$, it follows from Lemma 5 and Remark 1 that there exists an integer $N_\varepsilon^{(u,v)}$ such that

$$\gcd\left(p_u^n - 1, p_v^n - 1\right) < p_u^{n\varepsilon} \text{ for all integers } n > N_\varepsilon^{(u,v)}.$$

Set

$$N_\varepsilon = \max\{ \left\lceil \frac{\ln p_u}{\ln p_1} \cdot N_\varepsilon^{(u,v)} \right\rceil \mid 1 \leq u < v \leq r \},$$

where $\lceil a \rceil$ denotes the smallest integer greater than or equal to a. Then it is clear that

$$\gcd\left(p_u^n - 1, p_v^n - 1\right) < p_1^{n\varepsilon},\ 1 \leq u < v \leq r \text{ and } n > N_\varepsilon. \tag{18}$$

Let $2 \le k \le r$ and $1 \le i_1 < \cdots < i_k \le r$. It follows from (18) that if $n > N_\varepsilon$, then

$$\begin{aligned}\operatorname{lcm}\left(p_{i_1}^n - 1, p_{i_2}^n - 1, \ldots, p_{i_k}^n - 1\right) &\ge \frac{\prod_{j=1}^k (p_{i_j}^n - 1)}{\prod_{1\le j<l\le k} \gcd(p_{i_j}^n - 1, p_{i_l}^n - 1)} \\ &> p_1^{-k^2 n\varepsilon/2} \cdot \prod_{j=1}^k (p_{i_j}^n - 1) \\ &\ge p_1^{-r^2 n\varepsilon/2} \cdot \prod_{j=1}^k (p_{i_j}^n - 1).\end{aligned}$$

Consequently, we have

$$\begin{aligned}\frac{\prod_{j=1}^k p_{i_j}^{n/2}}{\operatorname{lcm}\left(p_{i_1}^n - 1, p_{i_2}^n - 1, \ldots, p_{i_k}^n - 1\right)} &\le p_1^{r^2 n\varepsilon/2} \cdot \prod_{j=1}^k \frac{p_{i_j}^{n/2}}{p_{i_j}^n - 1} \\ &< p_1^{r^2 n\varepsilon/2} \cdot \prod_{j=1}^k p_{i_j}^{1-n/2} \\ &\le p_1^{r^2 n\varepsilon/2} \cdot M \cdot \prod_{j=1}^k p_{i_j}^{-n/2}. \qquad (19)\end{aligned}$$

Note that $k \ge 2$ and $p_{i_j} \ge p_1$ for $1 \le j \le k$, and so (19) yields

$$\begin{aligned}\frac{\prod_{j=1}^k p_{i_j}^{n/2}}{\operatorname{lcm}\left(p_{i_1}^n - 1, p_{i_2}^n - 1, \ldots, p_{i_k}^n - 1\right)} &< p_1^{r^2 n\varepsilon/2} \cdot M \cdot p_1^{-nk/2} \\ &\le p_1^{r^2 n\varepsilon/2} \cdot M \cdot p_1^{-n} \\ &= M \cdot p_1^{-\frac{n}{2}\cdot(2-r^2\varepsilon)}.\end{aligned}$$

Hence it can be seen that

$$\begin{aligned}&\sum_{k=1}^r \sum_{1\le i_1<\cdots<i_k\le r} \frac{\prod_{j=1}^k p_{i_j}^{n/2}}{\operatorname{lcm}\left(p_{i_1}^n - 1, p_{i_2}^n - 1, \ldots, p_{i_k}^n - 1\right)} \\ &= \sum_{i=1}^r \frac{p_i^{n/2}}{p_i^n - 1} + \sum_{k=2}^r \sum_{1\le i_1<\cdots<i_k\le r} \frac{\prod_{j=1}^k p_{i_j}^{n/2}}{\operatorname{lcm}\left(p_{i_1}^n - 1, p_{i_2}^n - 1, \ldots, p_{i_k}^n - 1\right)} \\ &< \frac{r}{p_1^{n/2} - 1} + 2^r \cdot M \cdot p_1^{-\frac{n}{2}\cdot\left(2-r^2\varepsilon\right)}.\end{aligned}$$

Then choosing $\varepsilon < r^{-2}$, we get

$$0 \le \sum_{k=1}^r \sum_{1\le i_1<\cdots<i_k\le r} \frac{\prod_{j=1}^k p_{i_j}^{n/2}}{\operatorname{lcm}\left(p_{i_1}^n - 1, p_{i_2}^n - 1, \ldots, p_{i_k}^n - 1\right)} < \frac{r}{p_1^{n/2} - 1} + 2^r \cdot M \cdot p_1^{-n/2}. \qquad (20)$$

Since r, M and p_1 are all fixed integers with $p_1 \ge 3$, we get

$$\lim_{n\to\infty} \frac{r}{p_1^{n/2} - 1} + 2^r \cdot M \cdot p_1^{-n/2} = 0,$$

and so (17) follows from (20).

3.2 The Proof of Theorem 2

Proof (Proof of Theorem 2). Since

$$\left|\frac{N\left(\underline{b}^L,0\right)}{L}-\frac{M+1}{2M}\right|=\left|\frac{N\left(\underline{b}^L,1\right)}{L}-\frac{M-1}{2M}\right|,$$

it suffices to show that

$$\left|\frac{N\left(\underline{b}^L,0\right)}{L}-\frac{M+1}{2M}\right|<\frac{1}{M}\sum_{\substack{d\mid M\\ d>1}}\frac{d^{\frac{n}{2}}}{\lambda_n\left(d\right)}C\left(d,L\right). \tag{21}$$

First, it is clear that

$$\begin{aligned}N\left(\underline{b}^L,0\right)&=\sum_{t=k}^{k+L-1}\sum_{x=0}^{\frac{M-1}{2}}\left(\frac{1}{M}\sum_{h=0}^{M-1}e_M\left(h\left(a\left(t\right)-2x\right)\right)\right)\\&=\frac{1}{M}\sum_{h=0}^{M-1}\left(\sum_{t=k}^{k+L-1}e_M\left(ha\left(t\right)\right)\cdot\sum_{x=0}^{\frac{M-1}{2}}e_M\left(-2hx\right)\right)\\&=\frac{L\cdot\left(M+1\right)}{2M}+\frac{1}{M}\sum_{h=1}^{M-1}\left(\sum_{t=k}^{k+L-1}e_M\left(ha\left(t\right)\right)\cdot\sum_{x=0}^{\frac{M-1}{2}}e_M\left(-2hx\right)\right).\end{aligned}$$

Then proceed as in the proof of Theorem 1, we can get

$$\left|\frac{N\left(\underline{b}^T,0\right)}{L}-\frac{M+1}{2M}\right|\le\frac{1}{2ML}\sum_{\substack{d\mid M\\ d>1}}\sum_{\substack{1\le h\le d-1\\ \gcd(h,d)=1}}\left|\sum_{t=k}^{k+L-1}e_d\left(ha\left(t\right)\right)\right|\cdot\left|\sec\frac{h\pi}{d}\right|. \tag{22}$$

Note that given a divisor $d>1$ of M, $[h\underline{a}]_{\bmod d}$ is a primitive sequence over $\mathbf{Z}/(d)$ with period $\lambda_n\left(d\right)$ for every integer h coprime with d, and so by Lemma 2 we have

$$\begin{aligned}\left|\sum_{t=k}^{k+L-1}e_d\left(ha\left(t\right)\right)\right|&=\left|\sum_{t=k}^{k+[L]_{\bmod \lambda_n(d)}-1}e_d\left(ha\left(t\right)\right)+\left\lfloor\frac{L}{\lambda_n\left(d\right)}\right\rfloor\cdot\sum_{t=0}^{\lambda_n(d)-1}e_d\left(ha\left(t\right)\right)\right|\\&\le\left|\sum_{t=k}^{k+[L]_{\bmod \lambda_n(d)}-1}e_d\left(ha\left(t\right)\right)\right|+\left\lfloor\frac{L}{\lambda_n\left(d\right)}\right\rfloor\cdot\left|\sum_{t=0}^{\lambda_n(d)-1}e_d\left(ha\left(t\right)\right)\right|\\&\le d^{\frac{n}{2}}\left(\frac{4\ln\lambda_n\left(d\right)}{\pi^2}+0.409+\frac{[L]_{\bmod \lambda_n(d)}+1}{\lambda_n\left(d\right)}\right)+\left\lfloor\frac{L}{\lambda_n\left(d\right)}\right\rfloor\cdot d^{\frac{n}{2}}\\&=d^{\frac{n}{2}}\left(\frac{4\ln\lambda_n\left(d\right)}{\pi^2}+0.409+\frac{L+1}{\lambda_n\left(d\right)}\right),\end{aligned} \tag{23}$$

where $\lfloor a \rfloor$ denotes the largest integer smaller than or equal to a. Combining (22) and (23) we get

$$\begin{aligned}
&\left| \frac{N\left(\underline{b}^T, 0\right)}{L} - \frac{M+1}{2M} \right| \\
&\leq \frac{1}{2M} \sum_{\substack{d|M \\ d>1}} \sum_{\substack{1\leq h\leq d-1 \\ \gcd(h,d)=1}} \frac{d^{\frac{n}{2}}}{\lambda_n(d)} \left(\frac{\lambda_n(d)}{L} \left(\frac{4}{\pi^2} \ln \lambda_n(d) + 0.409 \right) + \frac{L+1}{L} \right) \cdot \left| \sec \frac{h\pi}{d} \right| \\
&\leq \frac{1}{2M} \sum_{\substack{d|M \\ d>1}} \frac{d^{\frac{n}{2}}}{\lambda_n(d)} \left(\frac{\lambda_n(d)}{L} \left(\frac{4}{\pi^2} \ln \lambda_n(d) + 0.409 \right) + \frac{L+1}{L} \right) \cdot \sum_{h=1}^{d-1} \left| \sec \frac{h\pi}{d} \right|. \quad (24)
\end{aligned}$$

and so (21) follows from (24) and Lemma 4.

Similar to Corollary 1, we can give the asymptotic property of $\frac{N\left(\underline{b}^L, s\right)}{L}$ as $n \to \infty$.

Corollary 2. *Let $\underline{a}$ be a primitive sequence of order n over $\mathbf{Z}/(M)$ with period $T = \lambda_n(M)$. Let $\underline{b} = [\underline{a}]_{mod\,2}$ and $\underline{b}^L = (b(k), b(k+1), \ldots, b(k+L-1))$ a segment of $\underline{b}$ with length $L = \mu T$, where $0 \leq k \leq T-1$ and $0 < \mu < 1$. Then for $s \in \{0, 1\}$ we have*

$$\lim_{n\to\infty} \frac{N\left(\underline{b}^L, s\right)}{L} = \frac{M+1-2s}{2M}.$$

Proof. Since

$$\begin{aligned}
\frac{\lambda_n(d)}{L} \left(\frac{4}{\pi^2} \ln \lambda_n(d) + 0.409 \right) + \frac{L+1}{L} &< \frac{\lambda_n(M)}{L} \left(\frac{4}{\pi^2} \ln M^n + 0.409 \right) + 2 \\
&= \frac{1}{\mu} \left(\frac{4}{\pi^2} n \ln M + 0.409 \right) + 2 \\
&< \left(\frac{4 \ln M}{\mu \pi^2} + \frac{3}{\mu} \right) \cdot n
\end{aligned}$$

and

$$\frac{d \ln d}{\pi} + 0.538d \leq \frac{M \ln M}{\pi} + 0.538M$$

hold for any divisor d of M with $d > 1$, it follows from Theorem 2 that

$$\left| \frac{N\left(\underline{b}^L, s\right)}{L} - \frac{M+1-2s}{2M} \right| < D_\mu(M) \sum_{\substack{d|M \\ d>1}} \frac{n d^{\frac{n}{2}}}{\lambda_n(d)},$$

where

$$D_\mu(M) = \left(\frac{4 \ln M}{\mu \pi^2} + \frac{3}{\mu} \right) \cdot \left(\frac{M \ln M}{\pi} + 0.538M \right)$$

is a constant only depended on M and μ. Therefore to prove Corollary 2, it suffices to show that

$$\lim_{n\to\infty}\sum_{\substack{d|M\\ d>1}}\frac{nd^{\frac{n}{2}}}{\lambda_n\left(d\right)}=0,$$

that is

$$\lim_{n\to\infty}\sum_{k=1}^{r}\sum_{1\le i_1<\cdots<i_k\le r}\frac{n\prod_{j=1}^{k}p_{i_j}^{n/2}}{\operatorname{lcm}\left(p_{i_1}^n-1,p_{i_2}^n-1,\ldots,p_{i_k}^n-1\right)}=0. \tag{25}$$

Proceed as in the proof of Corollary 1 (but substitute $\prod_{j=1}^{k}p_{i_j}^{n/2}$ by $n\prod_{j=1}^{k}p_{i_j}^{n/2}$), finally we can get

$$0\le\sum_{k=1}^{r}\sum_{1\le i_1<\cdots<i_k\le r}\frac{n\prod_{j=1}^{k}p_{i_j}^{n/2}}{\operatorname{lcm}\left(p_{i_1}^n-1,p_{i_2}^n-1,\ldots,p_{i_k}^n-1\right)}<\frac{rn}{p_1^{n/2}-1}+2^rM\cdot np_1^{-n/2}. \tag{26}$$

Since r, M and p_1 are all fixed integers with $p_1\ge 3$, we get

$$\lim_{n\to\infty}\frac{rn}{p_1^{n/2}-1}+2^rM\cdot np_1^{-n/2}=0,$$

and so (25) follows from (26).

3.3 An Example: Element Distribution of Modulo 2 Reductions of Primitive Sequences over $\mathbf{Z}/(2^{32}-1)$

Let $\underline{a}$ be a primitive sequence of order n over $\mathbf{Z}/(2^{32}-1)$ with period $T=\lambda_n\left(2^{32}-1\right)$. Let $\underline{b}=[\underline{a}]_{\bmod 2}$ and $\underline{b}^L=(b\left(k\right),b\left(k+1\right),\ldots,b\left(k+L-1\right))$ a segment of $\underline{b}$ with length $L=\mu T$, where $0\le k<T$ and $0<\mu\le 1$. Then it follows from Theorems 1 and 2 that

$$\left|\frac{N\left(\underline{b}^L,s\right)}{L}-\frac{2^{31}-s}{2^{32}-1}\right|<\varLambda_n\left(\mu\right), \tag{27}$$

where

$$\varLambda_n\left(\mu\right)=\begin{cases}\frac{1}{2^{32}-1}\sum\limits_{\substack{d|2^{32}-1\\ d>1}}\frac{d^{n/2}}{\lambda_n(d)}B\left(d\right), \text{ if } \mu=1;\\ \frac{1}{2^{32}-1}\sum\limits_{\substack{d|2^{32}-1\\ d>1}}\frac{d^{n/2}}{\lambda_n(d)}C\left(d,\mu T\right), \text{ if } 0<\mu<1,\end{cases}$$

with

$$B\left(d\right)=\frac{d\ln d}{\pi}+0.538d$$

and

$$C\left(d,\mu T\right)=\left(\frac{\lambda_n\left(d\right)}{L}\left(\frac{4}{\pi^2}\ln\lambda_n\left(d\right)+0.409\right)\right)\cdot B\left(d\right).$$

The values of $\Lambda_n(\mu)$ are calculated and listed in Table 1 for $1 \leq n \leq 10$ and $\mu \in \{1, 1/2, 1/4, 1/8\}$. It can be seen from Table 1 that the estimate of (27) is nontrivial if (1) $\mu = 1$ and $n \geq 2$; or (2) $0 < \mu < 1$ and $n \geq 3$. Moreover for any $\mu \in \{1, 1/2, 1/4, 1/8\}$, the value of $\Lambda_n(\mu)$ is very close to 0 if $n \geq 3$, which is consistent with the results of Corollarys 1 and 2.

Table 1. The values of $\Lambda_n(\mu)$ for $1 \leq n \leq 10$ and $\mu \in \{1, 1/2, 1/4, 1/8\}$

n	$\Lambda_n(1)$	$\Lambda_n(1/2)$	$\Lambda_n(1/4)$	$\Lambda_n(1/8)$
1	9.867	5.580×10^{2}	1.106×10^{3}	2.202×10^{3}
2	2.035×10^{-1}	2.640×10	5.260×10	1.050×10^{2}
3	5.877×10^{-9}	2.825×10^{-7}	5.591×10^{-7}	1.112×10^{-6}
4	1.534×10^{-7}	2.317×10^{-5}	4.618×10^{-5}	9.220×10^{-5}
5	2.125×10^{-10}	2.125×10^{-10}	2.125×10^{-10}	2.125×10^{-10}
6	5.850×10^{-9}	1.326×10^{-7}	2.600×10^{-7}	5.140×10^{-7}
7	1.875×10^{-11}	1.875×10^{-11}	1.875×10^{-11}	1.875×10^{-11}
8	1.384×10^{-11}	1.384×10^{-11}	1.384×10^{-11}	1.384×10^{-11}
9	5.368×10^{-12}	5.368×10^{-12}	5.368×10^{-12}	5.368×10^{-12}
10	3.891×10^{-12}	3.891×10^{-12}	3.891×10^{-12}	3.891×10^{-12}

4 Conclusions

In this paper, the distribution properties of modulo 2 reductions of primitive sequences modulo square-free odd integers are studied. Let M be a square-free odd integer, n a positive integer, and $\underline{a}$ a primitive sequence of order n over $\mathbf{Z}/(M)$ with period T. For $s \in \{0, 1\}$ and $0 < \mu \leq 1$, denote by f_s the proportion of occurrences of s within a segment of the binary sequence $[\underline{a}]_{\bmod 2}$, the modulo 2 reduction of $\underline{a}$, of length μT. Then it is shown that the difference of f_s from the average value $\frac{M+1-2s}{2M}$ tends to 0 as $n \to \infty$. Note that $\frac{M+1}{2M}$ differs from $\frac{M-1}{2M}$ by $\frac{1}{M}$. This implies that there exists a small imbalance between 0 and 1 occurring in the binary sequence $[\underline{a}]_{\bmod 2}$, and the bias of f_0 and f_1 is about $\frac{1}{M}$. To provide a good resistance against the distinguishing attacks, such imbalance should be taken into full consideration in the design of stream ciphers based on $[\underline{a}]_{\bmod 2}$. A simple method is to introduce the exclusive or operation. A bitwise exclusive or of several phase-shifts of $[\underline{a}]_{\bmod 2}$ will has smaller bias than $[\underline{a}]_{\bmod 2}$. Therefore, by introducing a moderate amount of exclusive or operations, the imbalance of $0, 1$ will be reduced to a small enough extent. In the future we will be interested in other pseudorandom properties of $[\underline{a}]_{\bmod 2}$, such as the linear complexity of $[\underline{a}]_{\bmod 2}$.

References

1. Bugeaud, Y., Corvaja, P., Zannier, U.: An upper bound for the G.C.D. of $a^n - 1$ and $b^n - 1$. Math. Z. **243**, 79–84 (2003)
2. Bylkov, D.N., Kamlovskii, O.V.: Occurrence indices of elements in linear recurrence sequences over primary residue rings. Probl. Inf. Transm. **44**, 161–168 (2008)
3. Chan, A.H., Games, R.A.: On the linear span of binary sequences obtained from finite geometries. In: Odlyzko, A.M. (ed.) CRYPTO 1986. LNCS, vol. 263, pp. 405–417. Springer, Heidelberg (1987). https://doi.org/10.1007/3-540-47721-7_29
4. Chen, H.J., Qi, W.F.: On the distinctness of maximal length sequences over $\mathbf{Z}/(pq)$ modulo 2. Finite Fields Appl. **15**(1), 23–39 (2009)
5. Cochrane, T.: On a trigonometric inequality of Vinogradov. J. Number Theory **27**(1), 9–16 (1987)
6. Dai, Z.D., Beth, T., Gollmann, D.: Lower bounds for the linear complexity of sequences over residue rings. In: Damgård, I.B. (ed.) EUROCRYPT 1990. LNCS, vol. 473, pp. 189–195. Springer, Heidelberg (1991). https://doi.org/10.1007/3-540-46877-3_16
7. Dai, Z.D.: Binary sequences derived from ML-sequences over rings I: periods and minimal polynomials. J. Cryptol. **5**(3), 193–207 (1992)
8. Fan, S.Q., Han, W.B.: Random properties of the highest level sequences of primitive sequences over $\mathbf{Z}/(2^e)$. IEEE Trans. Inf. Theory **49**(6), 1553–1557 (2003)
9. Hu, Z., Wang, L.: Injectivity of compressing maps on the set of primitive sequences modulo square-free odd integers. Cryptogr. Commun. **7**(4), 347–361 (2015)
10. Huang, M.Q., Dai, Z.D.: Projective maps of linear recurring sequences with maximal p-adic periods. Fibonacci Q. **30**(2), 139–143 (1992)
11. Lidl, R., Niederreiter, H.: Finite Fields. Encyclopedia of Mathematics and Its Applications, vol. 20. Cambridge University Press, Cambridge (1997)
12. Kamlovskii, O.V.: Frequency characteristics of linear recurrences over Galois rings. Matematicheskii Sbornik **200**, 31–52 (2009)
13. Klapper, A., Goresky, M.: Feedback shift registers, 2-adic span, and combiners with memory. J. Crypt. **10**(2), 111–147 (1997)
14. Korobov, N.M.: Exponential Sums and Their Applications. Kluwer, Dordrecht (1992)
15. Kuzmin, A.S., Nechaev, A.A.: Linear recurring sequences over Galois ring. Russ. Math. Surv. **48**(1), 171–172 (1993)
16. Qi, W.F., Yang, J.H., Zhou, J.J.: ML-sequences over rings Z/(2$^{\mathrm{e}}$): I. Constructions of nondegenerative ML-sequences II. Injectivness of compression mappings of new classes. In: Ohta, K., Pei, D. (eds.) ASIACRYPT 1998. LNCS, vol. 1514, pp. 315–326. Springer, Heidelberg (1998). https://doi.org/10.1007/3-540-49649-1_25
17. Tian, T., Qi, W.F.: Injectivity of compressing maps on primitive sequences over $\mathbf{Z}/(p^e)$. IEEE Trans. Inf. Theory **53**(8), 2966–2970 (2007)
18. Ward, M.: The arithmetical theory of linear recurring series. Trans. Am. Math. Soc. **35**(3), 600–628 (1933)
19. Wikipedia, Trigonometric functions, Wikipedia website (2018). https://en.wikipedia.org/wiki/Trigonometric_functions#Calculus
20. Zhu, X.Y., Qi, W.F.: Compression mappings on primitive sequences over $\mathbf{Z}/(p^e)$. IEEE Trans. Inf. Theory **50**(10), 2442–2448 (2004)
21. Zhu, X.Y., Qi, W.F.: Further result of compressing maps on primitive sequences modulo odd prime powers. IEEE Trans. Inf. Theory **53**(8), 2985–2990 (2007)

22. Zhu, X.Y., Qi, W.F.: On the distinctness of modular reduction of maximal length modulo odd prime numbers. Math. Comput. **77**(263), 1623–1637 (2008)
23. Zheng, Q.X., Qi, W.F.: Distribution properties of compressing sequences derived from primitive sequences over $\mathbf{Z}/(p^e)$. IEEE Trans. Inf. Theory **56**(1), 555–563 (2010)
24. Zheng, Q.X., Qi, W.F.: A new result on the distinctness of primitive sequences over $\mathbf{Z}/(pq)$ modulo 2. Finite Fields Appl. **17**(3), 254–274 (2011)
25. Zheng, Q.X., Qi, W.F., Tian, T.: On the distinctness of binary sequences derived from primitive sequences modulo square-free odd integers. IEEE Trans. Inf. Theory **59**(1), 680–690 (2013)
26. Zheng, Q.X., Qi, W.F.: Further results on the distinctness of binary sequences derived from primitive sequences modulo square-free odd integers. IEEE Trans. Inf. Theory **59**(6), 4013–4019 (2013)
27. Zheng, Q.X., Qi, W.F., Tian, T.: On the distinctness of modular reduction of primitive sequences over $\mathbf{Z}/(2^{32}-1)$. Des. Codes Crypt. **70**(3), 359–368 (2014)
28. ETSI/SAGE Specification: Specification of the 3GPP Confidentiality and Integrity Algorithms 128-EEA3 & 128-EIA3. Document 4: Design and Evaluation Report; Version: 2.0; Date: 9th Sep. 2011. Tech. rep., ETSI 2011. http://www.gsmworld.com/our-work/programmes-and-initiatives/fraud-and-security/gsm_security_algorithms.htm