Arithmetic Crosscorrelations of Feedback with Carry Shift Register Sequences

Mark Goresky, *Associate Member, IEEE*, and
Andrew Klapper, *Associate Member, IEEE*

***Abstract*— An arithmetic version of the crosscorrelation of two sequences is defined, generalizing Mandelbaum's arithmetic autocorrelations. Large families of sequences are constructed with ideal (vanishing) arithmetic crosscorrelations. These sequences are decimations of the 2-adic expansions of rational numbers p/q such that 2 is a primitive root modulo q.**

***Index Terms*—Crosscorrelations, binary sequences, feedback with carry shift register (FCSR) sequences, 2-adic numbers.**

I. Introduction

In the study of pseudorandom binary sequences, we are often interested in the correlation properties of the sequences. These properties have importance for several practical applications, such as spread-spectrum communication systems, radar systems, signal synchronization, and cryptanalysis, as well as being of theoretical interest as measures of randomness.

The usual notion of crosscorrelation of two binary sequences $\boldsymbol{S}$ and $\boldsymbol{T}$ is the sum $\sum_i (-1)^{S_i+T_i}$, where the addition in the exponent is modulo 2. For many classes of sequences, this sum is quite difficult to evaluate. The purpose of this correspondence is to investigate a different notion of crosscorrelation between sequences, the *arithmetic crosscorrelation*. The usual crosscorrelation can be thought of as the number of ones minus the number of zeros in one period of the sequence formed by adding $\boldsymbol{S}$ and $\boldsymbol{T}$ bit by bit modulo 2. Mandelbaum [5] investigated a notion of what amounts to *arithmetic autocorrelation* in which a sequence was added to a shift of itself *with carry*, rather than bit by bit modulo 2. He showed that certain sequences have ideal arithmetic autocorrelations, in the sense that they are zero for nontrivial shifts.

In this correspondence we extend the notion of arithmetic autocorrelation to the arithmetic crosscorrelation of two sequences. We then show how to construct families of sequences with ideal crosscorrelations. That is, all the nontrivial crosscorrelations are identically zero. The sizes of these families are conjectured to be large on the basis of statistical evidence. This is in stark contrast to the case of ordinary crosscorrelations, where there are well-known lower bounds on the minimal crosscorrelations in families of a given size. For example, the maximum shifted crosscorrelation in a family of M sequences of period N is at least $M^2\sqrt{(M-1)/(MN-1)}$ [6].

Our sequences include the Barrows–Mandelbaum arithmetic codes as a special case. (We have previously referred to these as ℓ-sequences, in order to stress the analogy with m-sequences.) Like the Barrows–Mandelbaum codes, they may be used for synchronization [5]. However, our sequences have an added feature: they may be used for simultaneous synchronization and identification in a multiuser environment.

To give a simple example, suppose a central dispatch (S) wishes to send individual messages to one of a number of possible clients $(R_1, R_2, \cdots, R_k)$. Choose a family of sequences with at least k members; let N denote the period of each sequence in this family. Each client R_i is assigned a signature sequence $\boldsymbol{S_i}$ from the family. The client monitors a common synchronization and identification channel on which S broadcasts a bitstream $\boldsymbol{T}$. At the nth clock tick, the client R_i computes the arithmetic crosscorrelation $\Theta_{\boldsymbol{S_i},\boldsymbol{T}}(n)$ between her signature sequence $\boldsymbol{S_i}$ and a window of size N in the bitstream $\boldsymbol{T}$.

When the dispatcher S sends the signature sequence $\boldsymbol{S_i}$ on the common channel, the receiver R_i will compute a single large arithmetic crosscorrelation Θ exactly N clock ticks later, while the other clients will compute $\Theta_{\boldsymbol{S_j},\boldsymbol{T}}(m) = 0$ for all $j \neq i$ and for all shifts m. So the client R_i has been identified as the intended recipient and her receiver has been synchronized to the message.

If $\boldsymbol{S}$ is any sequence and $d \neq 0$ is an integer, then the sequence $\boldsymbol{T}$ is said to be a *d-fold decimation* of $\boldsymbol{S}$ if for every i, we have $T_i = S_{di}$. Let q be a power of a prime, $q = p^e$, such that 2 is a primitive root modulo q. (That is, $\mathrm{ord}_q(2)$, the least positive power of 2 that is congruent to one modulo 2, is $p^{e-1}(p-1)$. Hence $\mathrm{ord}_q(2)$ is as large as possible). We let $\boldsymbol{S}$ be the 2-adic expansion of a fraction $-a/q$, with $0 < a < q$, an ℓ-sequence (these concepts are explained in the next section). We further let $\mathcal{F}_{\boldsymbol{S}}$ be the family of all d-fold decimations of $\boldsymbol{S}$, where d is relatively prime to the period of $\boldsymbol{S}$. Our main result is the following.

Theorem 1: If $\boldsymbol{R}$ and $\boldsymbol{T}$ are sequences in $\mathcal{F}_{\boldsymbol{S}}$, then the arithmetic crosscorrelation of $\boldsymbol{R}$ and $\boldsymbol{T}$ with shift τ is zero unless $\tau = 0$ and $\boldsymbol{T} = \boldsymbol{R}$.

This theorem is proved in Section III.

II. Definitions and Background

To precisely describe the arithmetic crosscorrelation of two binary sequences, we first need some algebraic framework. Recall that a 2-*adic integer* is an infinite series $\sum_{i=0}^{\infty} s_i 2^i$, where $s_i \in \{0, 1\}$, and the usual variable in a power series is replaced by the integer 2. The set $\boldsymbol{Z}_2$ of 2-adic integers forms a ring under the operations of addition and multiplication *with carry*. The field of fractions of this ring, the 2-adic numbers $\boldsymbol{Q}_2$ can be identified with the completion of the rational numbers at the non-Archimedean valuation associated to the ideal (2). This field contains the rational numbers $\boldsymbol{Q}$, and $\boldsymbol{Q} \cap \boldsymbol{Z}_2$ is the set of 2-adic integers $\sum_i s_i 2^i$ whose associated binary sequence $(s_0, s_1, \cdots)$ is eventually periodic. Equivalently, this intersection is the set of rational numbers of the form a/q with q odd. Such a number corresponds to a strictly periodic sequence if and only if $-q < a \leq 0$. Note that in $\boldsymbol{Z}_2$, we have $-1 = 1 + 2 + 2^2 + \cdots$. See Koblitz [4] for the basics on 2-adic algebra.

The authors have previously studied the sequences that arise from 2-adic integers in terms of the generation by feedback shift register like devices called *feedback with carry shift registers* (*FCSR's*) [2], [3]. These devices output precisely 2-adic expansions of rational numbers, hence all eventually periodic sequences. The theory of linear feedback shift registers (LFSR) can be thought of as arising by associating to a binary sequence a power series in one indeterminate with the terms of the sequence as coefficients [1]. The algebraic

Manuscript received March 25, 1996; revised December 10, 1996. This work was supported by the National Science Foundation under Grant NCR-9400762.

M. Goresky was with Department of Mathematics and College of Computer Science, Northeastern University, Boston, MA 02115 USA. He is now with the School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540 USA.

A. Klapper is with the Department of Computer Science, University of Kentucky, Lexington, KY 40506-0046 USA.

Publisher Item Identifier S 0018-9448(97)03871-6.

theory of 2-adic numbers provides a framework analogous to LFSR theory for the study of FCSR sequences.

An FCSR consists of an n-bit register with contents $x_0, \cdots, x_{n-1}$, plus some additional memory m. The state is updated by computing an integer linear combination

$$\sigma = \sum_{i=1}^{n} q_i x_{n-i} + m$$

with $q_i \in \{0, 1\}$. Then we replace the state $x_0, \cdots, x_{n-1}, m$ by $x_1, \cdots, x_{n-1}, (\sigma \bmod 2)$, $\sigma/2$ (integer division without remainder), and output x_0. The output is always eventually periodic.

Suppose $\boldsymbol{S}$ is the sequence generated by this FCSR, with associated rational number $\alpha = a/q$. Then $q = \sum_{i=1}^{n} q_i 2^i - 1$, and is called the *connection number* of the FCSR. It is the analog in FCSR theory to the connection polynomial in LFSR theory. Conversely, the 2-adic expansion of every rational number with denominator q can be output by an FCSR with connection number q. There is also a representation of FCSR sequences that is analogous to the trace representation of LFSR sequences. Let $\gamma = 2^{-1} \bmod q$. Then, if $\boldsymbol{S}$ is strictly periodic, there is an integer B such that $S_i = (B \cdot \gamma^i \bmod q) \bmod 2$. This double mod notation means first reduce modulo q to a reduced residue (i.e., in the range from 0 to $q - 1$), then take the parity of the result. In particular, the period of $\boldsymbol{S}$ is the order of 2 modulo q. The period of an FCSR can be made large by choosing q so that 2 is a primitive root modulo q. For this to be true, it is necessary that q be a power of a prime $q = p^e$, but no precise condition for 2 to be a primitive root modulo q is known. However, it is known that if 2 is a primitive root modulo p and modulo p^2, then it is a primitive root modulo p^e for every e. The resulting sequences are the 2-adic analogs of m-sequences (maximum period linear feedback shift register sequences). We call them *ℓ-sequences* (for "long sequences"). They are exponentially longer than the smallest FCSR's that generate them, and are nearly deBruijn sequences. It is these sequences which we consider here. See [3] for details of these and other facts about FCSR sequences.

Recall that the ordinary crosscorrelation with shift τ of two sequences $\boldsymbol{S}$ and $\boldsymbol{T}$ of period N can be defined either as the sum $\sum_{i=1}^{N} (-1)^{s_i + t_{i+\tau}}$ or as the number of zeros minus the number of ones in one period of the bitwise exclusive–or of $\boldsymbol{S}$ and the τ shift of $\boldsymbol{T}$ [1]. The arithmetic crosscorrelation is the with-carry analog of the latter definition.

Definition 1: Let $\boldsymbol{S}$ and $\boldsymbol{T}$ be two eventually periodic sequences with period N, and let $0 \le \tau < N$. Let $\boldsymbol{T}^\tau$ be the sequence formed by shifting $\boldsymbol{T}$ by τ positions, $T_i^\tau = T_{i+\tau}$. Let α and β_τ be the 2-adic numbers whose coefficients are given by $\boldsymbol{S}$ and $\boldsymbol{T}^\tau$, respectively. Then the sequence of coefficients associated with $\alpha - \beta_\tau$ is eventually periodic, and its period divides N. The *shifted arithmetic crosscorrelation* $\Theta_{\boldsymbol{S},\boldsymbol{T}}(\tau)$ of $\boldsymbol{S}$ and $\boldsymbol{T}$ is the number of zeros minus the number of ones in a complete period of length N of $\alpha - \beta_\tau$. When $\boldsymbol{S} = \boldsymbol{T}$, the arithmetic crosscorrelation is called the *arithmetic autocorrelation* of $\boldsymbol{S}$.

If for all τ such that $\boldsymbol{S}$ and $\boldsymbol{T}^\tau$ are distinct we have $\Theta_{\boldsymbol{S},\boldsymbol{T}}(\tau) = 0$, then $\boldsymbol{S}$ and $\boldsymbol{T}$ are said to have ideal arithmetic correlations. A family of sequences is said to have ideal arithmetic correlations if every pair of sequences in the family has ideal arithmetic correlations.

III. Main Results

As stated in the introduction, if $\boldsymbol{S}$ is any sequence and $d \ne 0$ is an integer, then the sequence $\boldsymbol{T}$ is said to be a *d-fold decimation* of $\boldsymbol{S}$ if for every i, we have $T_i = S_{di}$. Throughout this section we let q be a power of a prime, $q = p^e$, such that 2 is a primitive root modulo q. We let $\boldsymbol{S}$ be the 2-adic expansion of a fraction $-a/q$, with $0 < a < q$, an ℓ-sequence (henceforth we obscure the distinction between a binary sequence and its associated 2-adic number). We further let $\mathcal{F}_{\boldsymbol{S}}$ be the family of all d-fold decimations of $\boldsymbol{S}$, where d is relatively prime to the period of $\boldsymbol{S}$.

The remainder of this section consists of a proof of Theorem 1. We first need a constraint on the sequences that can occur as ℓ-sequences.

Proposition 1: If $\boldsymbol{S}$ is an ℓ-sequence, then second half of one period of $\boldsymbol{S}$ is the bitwise complement of the first half.

Proof: Let $q = p^e$. We have that $2^{\phi(q)} \equiv 1 \bmod q$. That is,

$$p^e \,\big|\, 2^{p^{e-1}(p-1)} - 1 = \left(2^{p^{e-1}(p-1)/2} - 1\right)\left(2^{p^{e-1}(p-1)/2} + 1\right).$$

These two factors are relatively prime, so p^e divides one of them. By the primitivity of 2 modulo q, p^e cannot divide the first factor, hence it divides the second factor. It follows that $2^{\phi(q)/2} \equiv -1 \bmod q$. Equivalently, $\gamma^{\phi(q)/2} \equiv -1 \bmod q$.

We have $S_i = (B\gamma^i \bmod q) \bmod 2$ for some B. Thus

$$\begin{aligned} S_{i+\phi(q)/2} &= -(B\gamma^i \bmod q) \bmod 2 \\ &= q - (B\gamma^i \bmod q) \bmod 2 \end{aligned}$$

which is the complement of S_i. □

The above property is extended to decimations of ℓ-sequences.

Corollary 1: Let $d > 0$ be relatively prime to $p^e - p^{e-1}$, the period of $\boldsymbol{S}$. Let $\boldsymbol{T}$ be a d-fold decimation of $\boldsymbol{S}$. Then the second half of one period of $\boldsymbol{T}$ is the complement of the first half.

Proof: Note that d must be odd and $S_j = 1 - S_{j+(p^e - p^{e-1})/2}$. Thus we have

$$\begin{aligned} T_i = S_{id} = 1 - S_{id + \frac{p^e - p^{e-1}}{2}} &= 1 - S_{\left(i + \frac{p^e - p^{e-1}}{2}\right)d} \\ &= 1 - T_{i + \frac{p^e - p^{e-1}}{2}}. \end{aligned}$$ □

In general, if α is a 2-adic number, then we let $\bar{\alpha}$ be the complementary 2-adic number. That is, we replace each 1 by 0 and each 0 by 1 in the 2-adic expansion of α. Then $\alpha + \bar{\alpha} = -1$.

Corollary 2: Let $d > 0$ be relatively prime to $p^e - p^{e-1}$, the period of $\boldsymbol{S}$. Let $\boldsymbol{T}$ be a d-fold decimation of $\boldsymbol{S}$. Let $\boldsymbol{T}$ be the 2-adic expansion of a fraction $-b/q'$, with $\gcd(b, q') = 1$. Then q' divides $2^{(p^e - p^{e-1})/2} + 1$.

Proof: Let α be the 2-adic number associated to $\boldsymbol{T}$. By Corollary 1

$$-\alpha - 1 = \bar{\alpha} = x + 2^{(p^e - p^{e-1})/2}\alpha$$

for some ordinary integer x (in fact, $0 \le x < 2^{(p^e - p^{e-1})/2}$). It follows that

$$\left(2^{(p^e - p^{e-1})/2} + 1\right)\alpha = -(x + 1)$$

and, therefore,

$$\left(2^{(p^e - p^{e-1})/2} + 1\right)b = q'(x + 1).$$

The corollary follows since b and q' are relatively prime. □

Theorem 2: Let c and d be relatively prime to $p^e - p^{e-1}$. Let $\boldsymbol{R}$ and $\boldsymbol{T}$ be c- and d-fold decimations of $\boldsymbol{S}$, respectively. If $\boldsymbol{T}$ is a shift of $\boldsymbol{R}$ with shift τ, then the arithmetic crosscorrelation of $\boldsymbol{R}$ and $\boldsymbol{T}$ with shift τ is $p^e - p^{e-1}$. In all other cases, the arithmetic crosscorrelations of $\boldsymbol{R}$ and $\boldsymbol{T}$ are zero.

Proof: Let $N = p^e - p^{e-1}$. Let $\boldsymbol{R}$ and $\boldsymbol{T}$ have associated 2-adic numbers $\alpha' = -a'/q'$ and $\alpha'' = -a''/q''$, respectively, with

$$\gcd(a', q') = \gcd(a'', q'') = 1.$$

The shift of $\boldsymbol{T}$ by τ corresponds to a 2-adic integer $2^{N-\tau}\alpha'' + x$ for some ordinary integer x. The arithmetic crosscorrelation of $\boldsymbol{R}$ and $\boldsymbol{T}$ with shift τ is the number of zeros minus the number of ones in one length $p^e - p^{e-1}$ period of

$$\beta = \alpha' - (2^{N-\tau}\alpha'' + x) = -(a'q'' - 2^{N-\tau}q'a'' + xq'q'')/q'q''.$$

If $\boldsymbol{T}$ is a shift of $\boldsymbol{R}$ with shift τ, then $\beta = 0$ and the result follows.

Suppose $\boldsymbol{T}$ is not a shift of $\boldsymbol{R}$ with shift τ. Let $\boldsymbol{U}$ be the sequence associated to β. It suffices to show that any period of $\boldsymbol{U}$ is balanced. Let $\beta = -b/r$ with $\gcd(b,r) = 1$. Then $r = \mathrm{lcm}(q', q'')$, so by Corollary 2, r divides $2^{N/2}+1$. Moreover b is nonzero. In a single period of $\boldsymbol{U}$, we have

$$U_i = (B \cdot 2^{-i} \bmod r) \bmod 2.$$

for some B. Thus

$$\begin{aligned} U_{i+N/2} &= \left(B \cdot 2^{-(i+N/2)} \bmod r\right) \bmod 2 \\ &= (B \cdot 2^{-i} \cdot 2^{-N/2} \bmod r) \bmod 2 \\ &= (-B \cdot 2^{-i} \bmod r) \bmod 2. \end{aligned}$$

Since r is odd, for any $y \neq 0$ the parity of $-y$ is the complement of the parity of y. Thus $U_{i+N/2}$ is the complement of U_i. These elements occur in pairs in U (since N is a period of U), so U is balanced. □

Theorem 1 follows immediately.

IV. Computing Arithmetic Crosscorrelations

If $\boldsymbol{T}$ and $\boldsymbol{R}$ are two periodic sequences with associated 2-adic numbers α and β, the sequence associated with the difference $\alpha - \beta$ may not be strictly periodic (though it must be eventually periodic). Thus at first glance, computing the arithmetic crosscorrelation of two sequences is problematic. How many bits of the difference must be computed before we reach the periodic part? As it turns out, however, the number of bits needed is well bounded.

Proposition 2: Let $\boldsymbol{T}$ and $\boldsymbol{R}$ be periodic sequences with period N. Let α and β be the 2-adic numbers associated with $\boldsymbol{T}$ and $\boldsymbol{R}$. Let $\boldsymbol{U}$ be the sequence associated with $\alpha - \beta$. Then $\boldsymbol{U}$ is strictly periodic from at least U_N on.

Proof: As noted earlier, the strict periodicity of $\boldsymbol{T}$ and $\boldsymbol{R}$ implies that there are integers a, q, b, and r such that $\alpha = -a/q$, $\beta = -b/r$, $0 \le a \le q$, and $0 \le b \le r$. Thus

$$\alpha - \beta = \frac{bq - ar}{qr}.$$

We have $-qr \le -ar \le bq - ar \le bq \le rq$.

If $bq - ar \le 0$, then $\alpha - \beta$ is strictly periodic and we are done. Otherwise,

$$\alpha - \beta = 1 + \frac{bq - ar - qr}{qr}$$

and $-qr < bq - ar - qr \le 0$. Therefore, $(bq - ar - qr)/qr \triangleq \gamma$ is strictly periodic. Let $\boldsymbol{V}$ be the sequence associated with γ. If every bit of $\boldsymbol{V}$ is 1, then $\gamma = -1$ and $\alpha - \beta = 0$. Otherwise, there is an $i < N$ such that $V_i = 0$. It follows that the bits of $\alpha - \beta = 1 + \gamma$ are identical to those of γ from bit $i+1$ on. This proves the proposition. □

Consequently, the arithmetic crosscorrelation of $\boldsymbol{T}$ and $\boldsymbol{R}$ can be computed by computing the first $2N$ bits of the difference $\alpha - \beta$, and finding the difference between the number of zeros and the number of ones in the last N of these $2N$ bits. This is a linear time computation in N (though not easily parallelizable as is the case with standard crosscorrelations).

Furthermore, since $-\beta = \bar{\beta} + 1$, we can compute $\alpha - \beta$ as $\alpha + \bar{\beta} + 1$. If the carry from computing the first N bits of $\alpha + \bar{\beta} + 1$ is 1, then $\alpha + \bar{\beta} + 1$ is strictly periodic, so the first N bits suffice. Otherwise, the periodic part is exactly the first N bits of $\alpha + \bar{\beta}$. Thus if we want to avoid storing α and β, we can simultaneously compute the arithmetic crosscorrelation based on the first N bits of $\alpha + \bar{\beta} + 1$ and $\alpha + \bar{\beta}$ as the bits arrive, and use the former if there is a carry from the first N bits, and the latter otherwise.

V. Generating the Sequences

If the sequences we have discussed are to find application, it is important that they be easily generated. As discussed in Section II, an ℓ-sequence $\boldsymbol{S}$ can be easily generated by a short FCSR. One way to generate a d-fold decimation $\boldsymbol{T}$ of $\boldsymbol{S}$ is to iterate the FCSR for $\boldsymbol{S}$ d times for each output bit. Unfortunately, this takes dN total iterations, which is nearly N^2 if d is large.

Consider the analogous situation for FCSR's. Here, if we iterate a register d times, the state will still have been updated by a linear function of the original state. The resulting register in turn can be replaced by an LFSR of the same or smaller size. The same thing does not work for FCSR's. Iterating the register results in more complicated state change functions, due to the mix of integer and modulo 2 operations.

An alternative is to use the exponential representation. Recall that if q is the connection number of the FCSR that outputs $\boldsymbol{S}$, and $\gamma = 2^{-1} \bmod q$, then there is an integer B such that

$$S_i = (B \cdot \gamma^i \bmod q) \bmod 2.$$

It follows that if $\delta = \gamma^d$, then

$$T_i = (B \cdot \delta^i \bmod q) \bmod 2.$$

Thus one way to generate $\boldsymbol{T}$ is to initialize a register with the value B, and at each iteration output the value modulo 2 and update the state by multiplying it by δ modulo q. This update can be done in time $O(\log(q)^2)$ (faster using divide-and-conquer techniques). This is much faster than iterating the original FCSR d times for most values of d, though not as fast as using the original FCSR to generate $\boldsymbol{S}$.

VI. Distinctness

Let $\boldsymbol{S}$ be an ℓ-sequence with associated 2-adic number $-p/q$, q prime. Let $t = \lfloor \log_2(q) \rfloor$ and let $N = \phi(q)$ be the period of $\boldsymbol{S}$. In this section we prove the following theorem.

Theorem 3: $\boldsymbol{S}$ is different from every shift of a $d = -1$ decimation of $\boldsymbol{S}$.

Proof: The basis for the proof is an analysis of the numbers of occurences of certain t and $t+1$ bit binary sequences in a single period of $\boldsymbol{S}$. A k bit sequence $\bar{x} = (x_0, \cdots, x_{k-1})$ occurs n times in a period of $\boldsymbol{S}$ if

$$n = |\{j, 0 \le j \le N-1 : \; s_j = x_0, s_{j+1} = x_1, \cdots s_{j+k-1} = x_{k-1}\}|.$$

We start by recalling a result from [3]. Every shift of $\boldsymbol{S}$ corresponds to a 2-adic number of the form $-a/q$, with $0 < a < q$ and $\gcd(a,q) = 1$. Thus each occurrence of a t bit pattern $\bar{x}$ in $\boldsymbol{S}$ corresponds to such an a with

$$x \overset{\text{def}}{=} \sum_{i=0}^{t-1} x_i 2^i \equiv \frac{-a}{q} \bmod 2^t.$$

Denote by $[z]$ the residue of z modulo 2^t in the range 1 to 2^t. Every $\bar{x}$ occurs at least once, since we can take $a = [-qx] = [-\epsilon x]$, where $\epsilon = q - 2^t$. Any $\bar{x}$ occurs twice in $\boldsymbol{S}$ if and only if there is an a with $0 < a < \epsilon$ such that $x \equiv -a/q \bmod 2^t$, or, equivalently, if and only if $[-\epsilon x] < \epsilon$. In this case, the second occurence corresponds to $-(a + 2^t)/q$. Also, it follows from [3] that every $t+1$ bit sequence occurs at most once in $\boldsymbol{S}$. Finally, it was shown in [3] that the second half of any period of $\boldsymbol{S}$ is the complement of the first half. Thus the number of occurences of any t bit sequence in $\boldsymbol{S}$ is the same as the number of occurences of its complement.

Let $\boldsymbol{T}$ be a -1-fold decimation of $\boldsymbol{S}$. That is, for each i, $t_i = s_{-i}$ (where all arithmetic in the indices is modulo N). We use repeatedly the fact that a -1 decimation carries a t bit sequence of consecutive symbols to another t bit sequence of consecutive symbols. It is the failure of this fact for more general decimations that makes us unable to extend the proof.

Suppose $\boldsymbol{T}$ is a shift of $\boldsymbol{S}$. Then there is a j such that $s_i = t_{j+i} = s_{-j-i}$. We now show that we may assume j is 0 or 1. Consider the t bit sequence $\bar{x} = (0, 0, \cdots, 0)$. We have $x = 0$ for this $\bar{x}$, so $[-qx] = 2^t$. It follows that $\bar{x}$ occurs exactly once in $\boldsymbol{S}$. Since this sequence is a palindrome, it follows that

$$i = -i - t + 1 - j. \tag{1}$$

Note that if we replace $\boldsymbol{S}$ by a shift of $\boldsymbol{S}$ and then take a -1 decimation, the result is a shift of $\boldsymbol{S}$ (although j may change). Thus we are free to choose i. It follows from (1) that we can choose i to make $j = 0$ or 1, depending on whether $t-1$ is even or odd. From here on we assume this, that is, either t is odd and $s_i = s_{-i}$ for all i, or t is even and $s_i = s_{-i-1}$ for all i.

It also follows that $(0, \cdots, 0)$ and $(1, \cdots, 1)$ are the only t bit palindromes that occur only once.

Lemma 1: If $\bar{x} = (x_0, \cdots, x_t)$ is a $t+1$ bit palindrome, then $\bar{x}$ cannot occur in $\boldsymbol{S}$. Moreover, $\bar{x}' = (x_0, \cdots, x_{t-1})$ occurs exactly once in $\boldsymbol{S}$.

Proof: For the first statement, suppose first that t is odd and $s_i = s_{-i}$ for all i. Suppose $\bar{x}$ occurs at $s_j, \cdots, s_{j+t}$. Since $\bar{x}$ occurs at most once, we must have $j \equiv -j - t \bmod N$. This is impossible since N is even and t is odd. A similar argument works if t is even and $s_i = s_{-i-1}$.

For the second statement, if $\bar{x}'$ occurs twice, it must be followed by distinct bits in its different occurrences (any $t+1$ bit sequence occurs at most once). Therefore, $\bar{x}$ occurs, which we have just shown is false. □

We now proceed to derive a series of inequalities for ϵ by applying Lemma 1 to various sequences. In each case $\bar{x}$ has length t.

1) The sequence $\bar{x} = (1, 0, 0, \cdots, 0)$ occurs once. Here $x = 1$, so $\epsilon \le [-\epsilon] = 2^t - \epsilon$. Hence $\epsilon < 2^{t-1}$ (we cannot have equality since ϵ is odd).

2) The sequence $\bar{x} = (1, 0, 0, \cdots, 1)$ occurs twice. Here $x = 2^{t-1} + 1$, so $\epsilon > [-\epsilon x] = 2^{t-1} - \epsilon$. Thus $\epsilon > 2^{t-2}$.

3) The sequence $\bar{x} = (0, 0, 1, 1, \cdots, 1, 0)$ occurs once. Here $x = 2^{t-1} - 4$, so $\epsilon \le [-\epsilon(2^{t-1} - 4)] = [4\epsilon + 2^{t-1}]$. By points 1) and 2), $2^t < 4\epsilon + 2^{t-1} < 3 \cdot 2^t$. If $4\epsilon + 2^{t-1} > 2 \cdot 2^t$

$$[4\epsilon + 2^{t-1}] = 4\epsilon + 2^{t-1} - 2 \cdot 2^t > \epsilon$$

so $\epsilon > 2^{t-1}$. This contradicts point 1), so $\epsilon < 3 \cdot 2^{t-3}$.

4) The sequence $\bar{x} = (1, 1, 0, 0, \cdots, 0, 1, 1)$ occurs twice. Here $x = 2^{t-1} + 2^{t-2} + 3$. Thus

$$\begin{aligned}\epsilon > [-\epsilon(2^{t-1} + 2^{t-2} + 3)] &= [\epsilon(2^{t-2} - 3)] \\ &= [2^{t-2} + e_1 2^{t-1} - 3\epsilon]\end{aligned}$$

where $\epsilon \equiv 1 + 2e_1 \bmod 4$. Suppose $e_1 = 1$. Then $-3 \cdot 2^{t-3} < 3(2^{t-2} - \epsilon) < 0$ so

$$[2^{t-2} + e_1 2^{t-1} - 3\epsilon] = 2^t + 2^{t-2} + 2^{t-1} - 3\epsilon < \epsilon.$$

It follows that $7 \cdot 2^{t-4} < \epsilon$, which contradicts point 3). Therefore, $e_1 = 0$. Now

$$[2^{t-2} + e_1 2^{t-1} - 3\epsilon] = 2^t + 2^{t-2} - 3\epsilon < \epsilon.$$

It follows that $5 \cdot 2^{t-4} < \epsilon$.

5) The sequence $\bar{x} = (0, 0, 0, 1, 1, \cdots, 1, 0, 0)$ occurs once. Here $x = 2^{t-2} - 8$. Thus

$$\epsilon \le [-\epsilon(2^{t-2} - 8)] = [2^{t-1} + 2^{t-2} + 8\epsilon]$$

(since $e_1 = 0$). It follows from points 3) and 4) that

$$[2^{t-1} + 2^{t-2} + 8\epsilon] = 2^{t-1} + 2^{t-2} + 8\epsilon - 3 \cdot 2^t.$$

It then follows that $\epsilon > (9/7) \cdot 2^{t-2}$.

6) The sequence $\bar{x} = (1, 0, 0, 1, 1, \cdots, 1, 0, 0)$ occurs once. Here $x = 2^{t-2} - 7$. Thus

$$\epsilon \le [-\epsilon(2^{t-2} - 7)] = [2^{t-1} + 2^{t-2} + 7\epsilon].$$

It follows from points 3) and 5) that

$$[2^{t-1} + 2^{t-2} + 7\epsilon] = 2^{t-1} + 2^{t-2} + 7\epsilon - 3 \cdot 2^t.$$

It then follows that $\epsilon > 3 \cdot 2^{t-3}$, which contradicts point 3).

This completes the proof of Theorem 3. □

VII. Conclusion

To construct a family of sequences with ideal arithmetic correlations, we choose an ℓ-sequence $\boldsymbol{S}$ of period $N = p^{e-1}(p-1)$ and include all d-fold decimations of $\boldsymbol{S}$ with d relatively prime to N. The number of such d is

$$\phi(N) = \begin{cases} p^{e-2}(p-1)\phi(p-1), & \text{if } e \ge 2 \\ \phi(p-1), & \text{if } e = 1. \end{cases}$$

If, for example, $p-1$ is two times a prime number, then the number of such d is about half the period. However, it is possible that distinct decimations give rise to sequences that are cyclic permutations of each other. This in fact happens when $p^e \in \{5, 9, 11, 13\}$. In these cases, the periods are $4, 6, 10,$ and $12,$ respectively, the values of $\phi(N)$ are $2, 2, 4,$ and 4, respectively, while the numbers of cyclically distinct decimations are $1, 1, 1,$ and $2,$ respectively. Surprisingly, a computer search has shown that for every other ℓ-sequence with period up to 4253, the set of these decimations is cyclically distinct.

Conjecture 1: Let $\boldsymbol{S}$ be an ℓ-sequence with connection number p^e and period N. Suppose $p^e \notin \{5, 9, 11, 13\}$. Let d_1 and d_2 be relatively prime to N and incongruent modulo N. If $\boldsymbol{R}$ is a d_1-fold decimation of $\boldsymbol{S}$ and $\boldsymbol{T}$ is a d_2-fold decimation of $\boldsymbol{S}$, then $\boldsymbol{R}$ and $\boldsymbol{T}$ are cyclically distinct.

We have proved this conjecture in the case $d_1 = 1$ and $d_2 = -1$. Should this conjecture be proved in general, we will have produced large families of cyclically distinct sequences with ideal arithmetic correlations. It would also be interesting to know whether it is possible to construct still larger families with ideal arithmetic correlations.

References

[1] S. Golomb, *Shift Register Sequences*. Laguna Hills, CA: Aegean Park, 1982.

[2] A. Klapper and M. Goresky, "2-Adic shift registers," in *Fast Software Encryption: Proc. 1993 Cambridge Security Workshop,* (Springer-Verlag Lecture Notes in Computer Science), vol. 809, R. Anderson, Ed. New York: Springer-Verlag, 1994, pp. 174–178.

[3] ——, "Feedback shift registers, combiners with memory, and 2-adic span," *J. Cryptology*, to be published.

[4] N. Koblitz, *p-Adic Numbers, p-Adic Analysis, and Zeta Functions*. New York: Springer-Verlag, 1984.

[5] D. Mandelbaum, "Arithmetic codes with large distance," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 237–242, 1967.

[6] L. R. Welch, "Lower bounds on the maximum correlation of signals," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 397–399, 1974.