Construction and Applications of CRT Sequences

Kenneth W. Shum, *Member, IEEE*, and Wing Shing Wong, *Fellow, IEEE*

***Abstract*—Protocol sequences are used for channel access in the collision channel without feedback. Each user accesses the channel according to a deterministic zero-one pattern, called the protocol sequence. In order to minimize fluctuation of throughput due to delay offsets, we want to construct protocol sequences whose pairwise Hamming cross-correlation is as close to a constant as possible. In this paper, we present a construction of protocol sequences which is based on the bijective mapping between one-dimensional sequence and two-dimensional array by the Chinese Remainder Theorem (CRT). In the application to the collision channel without feedback, a worst-case lower bound on system throughput is derived.**

Tags: Protocol sequences, collision channel without feedback, cyclically permutable constant-weight codes, optical orthogonal codes.

I. Introduction

A. Background and Motivation

Randomness is commonly used in the design of multiple-access schemes. For example, in slotted ALOHA, each user transmits a packet with probability p independently. Implementations of such random access schemes in practice usually substitute random variables by pseudo-random numbers. However, high-quality pseudo-random number generation may be too complicated for applications, such as wireless sensor networks, where computing power is limited. The objective of this paper is to construct binary pseudo-random sequences, called protocol sequences, that are tailored to the quality-of-service requirements of interest, such as high throughput and bounded delay.

Protocol sequences are used in multiple-access in the collision channel without feedback [1]. In this paper, we consider a time-slotted system, consisting of a number of transmitters and one receiver. A user sends a packet within the boundaries of a time slot. If exactly one user transmits in a time slot, the received packet is received successfully. If two or more users transmit in the same time slot, a collision is incurred, and the received packet is assumed unrecoverable. If no user transmits in a time slot, that time slot is idle. Since there is no feedback from the receiver, collision resolution algorithm such as the stack algorithm is not possible. Each user repeats his assigned binary protocol sequence periodically, and transmits a packet if and only if the value of the protocol sequence at that time slot equals one. We note that the transmission schedule is independent of the data being sent and there is no cooperation among the users.

The capacity of the collision channel without feedback is characterized by Massey and Mathys in [1]. It is shown that the zero-error sum-capacity is e^{-1}. They use protocol sequences having the special property that the Hamming cross-correlation is independent of relative delay offsets. In fact, the Hamming cross-correlation they considered is a generalized notion of Hamming cross-correlation, which is defined for all nonempty subsets of users, not just for pairs of users. Protocol sequences with this property are called *shift-invariant* sequences [2]. Shift-invariant protocol sequences have the advantage that there is no fluctuation in throughput no matter what the delay offsets are, and hence have the largest worst-case system throughput. Constructions of shift-invariant protocol sequences are considered in [1]–[3]. Nevertheless, shift-invariant protocol sequences have the drawback that the period grows exponentially in the number of users [2]. Even if we relax this requirement and consider protocol sequence sets with only pairwise Hamming cross-correlation being constant, it is shown in [4] that the period grows exponentially in the number of users as well. Long period length has the disadvantage that individual or system throughput is invariant only if it is averaged over a long period. Individual users may suffer short-time starvation. In order to achieve short period length, we must seek for protocol sequences with some small variance in Hamming cross-correlation allowed.

After the seminal work of [1], more constructions of protocol sequences are given in [5]–[9], sometime under the name of *cyclically permutable constant-weight codes*, or *optical orthogonal codes*. The main difference between these works in the literature and our construction is that, the protocol sequences in these papers are required to have small Hamming cross-correlation and auto-correlation. In our construction, Hamming auto-correlation may be very large.

Another class of protocol sequences, called *wobbling sequences* [10], has period equal to M^4, where M is the number of users, with worst-case system throughput provably larger than a positive constant that is approximately equal to 0.25 when M is large. The result in this paper improves upon the wobbling sequences by constructing protocol sequences of order $O(M^2)$ or $O(M^3)$, depending on the models of user activity, while the guarantee of the worst-case system throughput remains the same.

The construction proposed in this paper is based on the Chinese remainder theorem (CRT), and thereby the constructed sequences are called *CRT sequences*. We remark that the use of CRT in the construction of protocol sequences is not new. The constructions of optical orthogonal codes in [5]–[7] also employ CRT. However, the specific construction proposed here is novel. Although the analysis of CRT sequences is quite

Kenneth W. Shum and Wing Shing Wong are with the Dept. of Information Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong. Emails: wkshum@inc.cuhk.edu.hk, wswong@ie.cuhk.edu.hk.

This work was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region under Project 417909.

The result in this paper was partially presented in IEEE Int. Symp. on Inform. Theory, Austin, 2010. This paper appears in IEEE Trans. on Inform. Theory, Nov, 2010.

involved, the generation of such sequences requires no more than computing linear operations in modular arithmetics

The proposed CRT sequences have two special features, *user-identification* and *frame-synchronization* capability. The sender of each successfully received packet can be identified by looking at the channel activity only, without looking into the packet contents. Also, we can determine the start time of a protocol sequence, again, based on the channel activity only. One potential application based on this property is tracking targets by detecting energy pulses coming from multiple distributed sources, for example, as in a multistatic radar system or in a ultrasound based sensor network. While it may be difficult to identify the transmitting source of a single pulse, if each source employs a CRT sequence and sends multiple pulses according to the corresponding protocol sequence, then the identifiability property ensures that a detector can identify the transmitting sources of the pulses and make use of the information for more accurate tracking of the targets.

B. Main Results

The main construction is given in Section II. One of the key ingredients in the construction is the one-to-one correspondence between one-dimensional sequence and two-dimensional arrays via CRT. The correlation property of the constructed sequences is analyzed in Section III.

Applications to the collision channel without feedback are addressed in Section IV with two different user activity models [11]. In the first model, an infinite backlog of data is assumed and the users are active throughout the transmission. We show in Theorem 7 that under this user activity model, we can achieve a throughput of 0.25 with $O(M^{2+\epsilon})$ sequence period, where M is the number of users.

In the second model, users become active only if they have data to send, and remain idle otherwise. If a user becomes active, it is required that the user remains active for at least one period of the protocol sequence assigned. We show in the second part of Section IV that if the number of active users is no more than one half of the number of potential users, and if the protocol sequences are sufficiently long, namely $O(M^3)$, the receiver can detect the set of active users and determine their starting time correctly, even without any packet header.

To facilitate comparison between different protocol sequences, we introduce in Section V a notion called ϵ-uniformity, which measures the variation of Hamming cross-correlation. Other applications of CRT are given in Section VI.

C. Definitions and Notations

Let $\mathbb{Z}_n$ be the ring of residues mod n for a positive integer n. We will reserve the letter L for sequence length.

Definition 1. The components in a sequence of length L are indexed from 0 to $L-1$. The time indices $\{0, 1, \ldots, L-1\}$ is identified with $\mathbb{Z}_L$. The *Hamming weight* of a binary sequence $a(t)$ of length L is denoted by w_a. The *duty factor* [1] of $a(t)$ is the Hamming weight divided by the length,

$$f_a := \frac{1}{L}\sum_{t=0}^{L-1} a(t).$$

For two binary sequences $a(t)$ and $b(t)$ of length L, their *Hamming correlation function* is defined as

$$H_{ab}(\tau) := \sum_{t=0}^{L-1} a(t)b(t-\tau),$$

where τ is the delay offset, and $t-\tau$ is the difference in modulo-L arithmetic. When $a(t) = b(t)$, $H_{aa}(\tau)$ is called the *Hamming auto-correlation* of $a(t)$. When $a(t)$ and $b(t)$ are two different sequences, $H_{ab}(\tau)$ is called the *Hamming cross-correlation* of $a(t)$ and $b(t)$.

When the number of ones in a zero-one sequence is small in compare to the length, the sequence can be compactly represented by specifying the locations of ones.

Definition 2. Given a sequence $a(t)$ of length L, let the *characteristic set* of $a(t)$, denoted by $\mathcal{I}_a$, be the subset of $\mathbb{Z}_L$ such that $t \in \mathcal{I}_a$ if and only if $a(t) = 1$.

Shifting a sequence cyclically by τ is equivalent to translating its characteristic set by τ, with addition performed modulo L. Given a subset $\mathcal{I}$ in $\mathbb{Z}_L$, and $\tau \in \mathbb{Z}_L$, we denote the translation of $\mathcal{I}$ by τ as

$$\mathcal{I} + \tau := \{x + \tau \in \mathbb{Z}_L : x \in \mathcal{I}\}.$$

Expressed in terms of the characteristic set, the Hamming correlation of sequences $a(t)$ and $b(t)$ equals

$$H_{ab}(\tau) = |\mathcal{I}_a \cap (\mathcal{I}_b + \tau)|$$

for all $\tau = 0, 1, \ldots, L-1$, where $|\mathcal{S}|$ denote the size of a set $\mathcal{S}$.

II. The CRT Construction

A. The CRT correspondence

We shall construct sequences with length $L = pq$, where p and q are relatively prime integers. In subsequent discussions, we take p to be a prime number and q an integer not divisible by p.

Define a mapping from $\mathbb{Z}_{pq}$ to the direct sum

$$G_{p,q} := \mathbb{Z}_p \oplus \mathbb{Z}_q$$

by

$$\Phi_{p,q}(x) := (x \bmod p, x \bmod q).$$

By Chinese remainder theorem [12, p.34], $\Phi_{p,q}$ is a bijective map. We will call $\Phi_{p,q}$ the *CRT correspondence*. When the values of p and q are clear from the context, we write $\Phi(x)$ instead of $\Phi_{p,q}(x)$.

It can be easily checked that the CRT correspondence is a linear map, meaning that

$$\Phi_{p,q}(x + x') = \Phi_{p,q}(x) + \Phi_{p,q}(x').$$

Here, the addition on the left hand side is the addition in $\mathbb{Z}_{pq}$, and the addition on the right hand side is the addition in $G_{p,q}$.

A sequence $s(t)$ of length L is associated with a $p \times q$ array $\mathbf{S}(t_1, t_2)$, where t_1 and t_2 range from 0 to $p-1$ and 0 to $q-1$ respectively, via the relation

$$\mathbf{S}(t \bmod p, t \bmod q) = s(t).$$

The corresponding characteristic set of $s(t)$ in $\mathbb{Z}_L$ is mapped to a subset of $G_{p,q}$ via $\Phi_{p,q}$ as well.

Under the CRT correspondence, a one-dimensional cyclic shift of $s(t)$ by one time unit is equivalent to a column-wise shift followed by a row-wise shift. One-dimensional correlation properties can be translated to the two-dimensional ones.

For illustration, consider the time indices $(0, 1, 2, \ldots, 14)$ as an integer sequence of length $L = 15$, $p = 3$ and $q = 5$. By the bijection $\Phi_{3,5}$, this integer sequence is mapped to

$$\begin{bmatrix} 0 & 6 & 12 & 3 & 9 \\ 10 & 1 & 7 & 13 & 4 \\ 5 & 11 & 2 & 8 & 14 \end{bmatrix}.$$

When $(0, 1, \ldots, 14)$ is cyclically shifted to $(14, 0, 1, 2, \ldots, 13)$, the corresponding array is

$$\begin{bmatrix} 14 & 5 & 11 & 2 & 8 \\ 9 & 0 & 6 & 12 & 3 \\ 4 & 10 & 1 & 7 & 13 \end{bmatrix}.$$

Note that the second array can be obtained from the first one by cyclically shift downward by one row and then to the right one column.

For $\boldsymbol{\tau} = (\tau_1, \tau_2) \in G_{p,q}$, define the Hamming correlation between two 2-dimensional arrays $\mathbf{A}$ and $\mathbf{B}$ by

$$H_{\mathbf{AB}}(\boldsymbol{\tau}) := \sum_{\boldsymbol{t}} \mathbf{A}(\boldsymbol{t})\mathbf{B}(\boldsymbol{t} - \boldsymbol{\tau}),$$

with the subtraction calculated in $G_{p,q}$, and $\boldsymbol{t}$ running over all elements in $G_{p,q}$. It is easy to check that this definition of Hamming correlation is compatible with the 1-dimensional analog, i.e.,

$$H_{\mathbf{AB}}(\tau_1, \tau_2) = H_{ab}(\tau)$$

with $\tau_1 \equiv \tau \bmod p$ and $\tau_2 \equiv \tau \bmod q$.

The Hamming correlation between two 2-dimensional arrays $\mathbf{A}$ and $\mathbf{B}$ can be expressed in terms of their characteristic set $\mathcal{I}_{\mathbf{A}}$ and $\mathcal{I}_{\mathbf{B}}$ as

$$H_{\mathbf{AB}}(\boldsymbol{\tau}) = |\mathcal{I}_{\mathbf{A}} \cap (\mathcal{I}_{\mathbf{B}} + \boldsymbol{\tau})|,$$

where $\mathcal{I}_{\mathbf{A}}$ denotes the characteristic set of $\mathbf{A}$,

$$\mathcal{I}_{\mathbf{A}} := \{(i, j) \in \mathbb{Z}_p \oplus \mathbb{Z}_q : \mathbf{A}(i, j) = 1\},$$

and the addition in $\mathcal{I}_{\mathbf{B}} + \boldsymbol{\tau}$ is performed in $\mathbb{Z}_p \oplus \mathbb{Z}_q$.

B. The CRT sequences

We will construct sequences by specifying characteristic sets in $G_{p,q}$. Rows and columns of matrices and arrays will be indexed by $\{0, 1, \ldots, p-1\}$ and $\{0, 1, \ldots, q-1\}$ respectively.

Definition 3. Let p be prime and q be an integer not divisible by p. For $g \in \mathbb{Z}_p$, we define

$$\mathcal{I}_{g,p,q} := \{(g, 1)t \in G_{p,q} : 0 \le t < q\}. \tag{1}$$

The notation $(g, 1)t$ simply means the sum of t copies of $(g, 1)$ in $G_{p,q}$,

$$(g, 1)t := \underbrace{(g, 1) + (g, 1) + \ldots + (g, 1)}_{t}.$$

$$\mathcal{I}_{0,3,5}: \begin{array}{|ccccc|}\hline 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

$$\mathcal{I}_{1,3,5}: \begin{array}{|ccccc|}\hline 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ \hline \end{array}$$

$$\mathcal{I}_{2,3,5}: \begin{array}{|ccccc|}\hline 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ \hline \end{array}$$

Fig. 1. The three arrays associated with characteristic sets $\mathcal{I}_{0,3,5}$, $\mathcal{I}_{1,3,5}$, $\mathcal{I}_{2,3,5}$.

We say that $\mathcal{I}_{g,p,q}$ is generated by g, and g is the *generator* of $\mathcal{I}_{g,p,q}$. If p and q are clear from the context, we write $\mathcal{I}_g$ instead of $\mathcal{I}_{g,p,q}$.

The elements in $\mathcal{I}_{g,p,q}$ form an arithmetic progression in $G_{p,q}$ with common difference $(g, 1)$. We can re-write $\mathcal{I}_{g,p,q}$ in the following form

$$\mathcal{I}_{g,p,q} = \{(gt, t) \in G_{p,q} : 0 \le t < q\},$$

with the product gt in the first component reduced mod p.

Remarks: For $g = 0, 1, \ldots, p-1$, the array with $\mathcal{I}_{g,p,q}$ as the characteristic set contains exactly one "1" in each column. For $g \neq 0$, each block of p consecutive columns form a permutation matrix. (Recall that a permutation matrix is a square zero-one matrix with exactly one "1" in each row and each column.)

Definition 4. (CRT sequences) For $g = 0, 1, \ldots, p-1$, define the *CRT sequence generated by* g, denoted by $s_{g,p,q}(t)$, be the binary sequence of length L obtained by setting

$$s_{g,p,q}(t) = \begin{cases} 1 & \text{if } \Phi_{p,q}(t) \in \mathcal{I}_{g,p,q} \\ 0 & \text{otherwise.} \end{cases}$$

We will write $s_g(t)$ if the values of p and q are understood.

Example 1. $p = 3$ and $q = 5$. The three characteristic sets are:

$$\mathcal{I}_{0,3,5} = \{(0, 0), (0, 1), (0, 2), (0, 3), (0, 4)\},$$
$$\mathcal{I}_{1,3,5} = \{(0, 0), (1, 1), (2, 2), (0, 3), (1, 4)\},$$
$$\mathcal{I}_{2,3,5} = \{(0, 0), (2, 1), (1, 2), (0, 3), (2, 4)\}.$$

The three arrays are shown in Fig. 1. The top left corner in each array is the $(0, 0)$-entry. The generated CRT sequences are listed as follows

$$s_0(t) : 10010\,01001\,00100$$
$$s_1(t) : 11111\,00000\,00000$$
$$s_2(t) : 10010\,00100\,01001.$$

CRT sequences satisfy the following properties.

Theorem 1.

1) *The Hamming weight of each CRT sequence is* q.

2) $s_0(t+p) = s_0(t)$, *i.e., the least period of sequence* $s_0(t)$ *is* p.
3) *For each residue* $i \bmod q$, *there is exactly one "1" at* $t = i, i+q, i+2q, \ldots, i+(p-1)q$ *in* $s_g(t)$.

Proof: As the characteristic set $\mathcal{I}_{g,p,q}$ contains q elements and the CRT correspondence Φ preserves Hamming weight, the first statement in the proposition follows immediately.

For $g = 0$, the elements in characteristic set $\mathcal{I}_{0,p,q}$ have the first coordinate identically equal to zero. For $j = 0, 1, \ldots, q-1$, The pre-image of $(0, j)$ under Φ is a multiple of p. We have $s_0(t) = 1$ whenever t is a multiple of p.

For the last statement in the proposition, consider the image of $i + kq$, for $k = 0, 1, \ldots, p-1$, under the mapping Φ,

$$\Phi(i + kq) = (i + kq \bmod p, i \bmod q).$$

The second coordinate is constant. If k goes through 0 to $p-1$, $\Phi(i+kq)$ will go through the ith column in the array with characteristic set $\mathcal{I}_{g,p,q}$. Because there is exactly one "1" in each column, there is exactly one "1" in positions $t = i, i+q, \ldots, i+(p-1)q$. ∎

III. Correlation Properties of CRT Sequences

We continue to use the notation that p is a prime number, q is an integer not divisible by p. Let $L = pq$ denotes the sequence length. In this section, we determine the Hamming correlation of the CRT sequences. In this section, we will use the notation $\bar{x}$ for the remainder of x after division by p. We distinguish the two additions mod p and mod q by $\oplus_p$ and $\oplus_q$, respectively.

For $g = 0, 1, \ldots, p-1$, the $p \times q$ array corresponding to sequence $s_g(t)$ is denoted by $\mathbf{A}_g$. Recall that the characteristic set of $\mathbf{A}_g$ is given in (1), and $\mathbf{A}_g(i,j) = 1$ if and only if $i \equiv \bar{j}g \bmod p$, for $i = 0, \ldots, p-1$ and $j = 0, \ldots, q-1$. For notational convenience, we let

$$H_{gh}(\tau_1, \tau_2) := H_{\mathbf{A}_g \mathbf{A}_h}(\tau_1, \tau_2). \tag{2}$$

A. Hamming Cross-correlation

Let g and h be two distinct elements in $\mathbb{Z}_p$. As argued in the previous section, the cross-correlation of CRT sequences $s_g(t)$ and $s_h(t)$ is equivalent to the cross-correlation of the associated $p \times q$ arrays $\mathbf{A}_g$ and $\mathbf{A}_h$, namely by counting the number of elements in common to $\mathcal{I}_g$ and $\mathcal{I}_h + (\tau_1, \tau_2)$. By definition, the translation $\mathcal{I}_h + (\tau_1, \tau_2)$ of $\mathcal{I}_h$ is

$$\{((\bar{j}h) \oplus_p \tau_1, j \oplus_q \tau_2) : j = 0, 1, \ldots, q-1\}. \tag{3}$$

By a change of variable, (3) can be written as

$$\{(((\overline{j \ominus_q \tau_2})h) \oplus_p \tau_1, j) : j = 0, 1, \ldots, q-1\}.$$

where $\ominus_q$ denotes subtraction mod q. In the calculation of $(\overline{j \ominus_q \tau_2})h$, $j \ominus_q \tau_2$ must be first reduced to an integer between 0 and $q-1$ before we reduce modulo p. The intersection of $\mathcal{I}_g$ and $\mathcal{I}_h + (\tau_1, \tau_2)$ equals to the number of solutions to

$$\bar{x}g \equiv ((\overline{x \ominus_q \tau_2})h) \oplus_p \tau_1 \bmod p. \tag{4}$$

for $x = 0, 1, \ldots, q-1$.

For nonzero h, we can divide both sides of (4) by h and re-write it as

$$\bar{x}(h^{-1}g) \equiv (\overline{x \ominus_q \tau_2}) \oplus_p (h^{-1}\tau_1) \bmod p.$$

For each fixed τ_2, as τ_1 runs through $\mathbb{Z}_p$, $h^{-1}\tau_1$ also runs through $\mathbb{Z}_p$. Therefore, the distribution of Hamming cross-correlation between $s_g(t)$ and $s_h(t)$ is the same as the distribution of Hamming cross-correlation between $s_{g/h}(t)$ and $s_1(t)$. From this observation, it suffices to take $h = 1$ without any loss of generality. We will consider the Hamming cross-correlation between $s_g(t)$ and $s_1(t)$, for $g = 0$ and $g = 2, 3, \ldots, p-1$.

The following simple lemma is used repeatedly in the derivation of Hamming cross-correlation properties.

Lemma 1. *For each* $b \in \mathbb{Z}_p$, *the number of solutions to*

$$\bar{x} \equiv b \bmod p$$

for x *going through* d *consecutive integers* $c, c+1, \ldots c+d-1$, *equals*

$$\begin{cases} d/p & \text{if } p \text{ divides } d, \\ \lfloor d/p \rfloor + \delta & \text{otherwise,} \end{cases}$$

where δ *equals either 0 or 1.*

Proof: Suppose that d is divisible by p. If we reduce the integers c, $c+1, \ldots, c+d-1$ mod p, we have each element in $\mathbb{Z}_p$ repeated d/p times. Hence, for each $b \in \mathbb{Z}_p$, there are exactly d/p integers in $\{c, c+1, \ldots, c+d-1\}$ whose residue mod p equal b

Now suppose that d is not divisible by p, we divide the d consecutive integers into two parts. Among the first $\lfloor d/p \rfloor p$ integers, for each $b \in \mathbb{Z}_p$, exactly $\lfloor d/p \rfloor$ equals b mod p. The residues of the remaining $d - \lfloor d/p \rfloor p$ integers are distinct. The number of integers in $\{c, c+1, \ldots, c+d-1\}$ whose residues equal b is either $\lfloor d/p \rfloor$ or $\lfloor d/p \rfloor + 1$. ∎

We obtain the following theorem immediately from Lemma 1.

Theorem 2. *The Hamming cross-correlation of* $s_1(t)$ *and* $s_0(t)$ *is equal to either* $\lfloor q/p \rfloor$ *or* $\lfloor q/p \rfloor + 1$.

Proof: If we put $g = 1$ and $h = 0$ in (4), we get

$$\bar{x} \equiv \tau_1 \bmod p.$$

The number of integers in $\{0, 1, \ldots, q-1\}$ that equal $\tau_1 \bmod p$ is either $\lfloor q/p \rfloor$ or $\lfloor q/p \rfloor + 1$ by Lemma 1. ∎

From now on, we assume $q > p$, which is the case of practical interest.

Theorem 3. *Suppose* $q > p$. *Let* m *be the quotient of* q *divided by* p, *i.e.,* $m = \lfloor q/p \rfloor$, *and let* $g \in \mathbb{Z}_p$, $0 \neq g \neq 1$. *Let* $\bar{q}$ *be the residue of* q *mod* p, *and*

$$b_g \equiv (g-1)^{-1}\bar{q} \bmod p. \tag{5}$$

The Hamming cross-correlation between $s_g(t)$ *and* $s_1(t)$ *is bounded between*

$$\begin{cases} m-1 \text{ and } m+1 & \text{if } 0 < b_g < p - \bar{q}, \text{ or} \\ m \text{ and } m+2 & \text{if } p - \bar{q} < b_g < p. \end{cases}$$

The proof of Theorem 3 is in Appendix A. We remark that in Theorem 3, b_g is neither 0 nor $p-\bar{q}$. The value of b_g is nonzero mod p because both $g-1$ and $\bar{q}$ are nonzero. On the other hand, b_g is equal to $p-\bar{q}$ only if $g=0$, which is excluded by assumption.

With more careful book-keeping, we can determine exactly the frequencies of occurrence of Hamming cross-correlation.

Definition 5. Let the distribution of $H_{g1}(\tau_1,\tau_2)$ be

$$N_g(j) := |\{(\tau_1,\tau_2) \in G_{pq} : H_{g1}(\tau_1,\tau_2) = j\}|, \quad (6)$$

for $j = 0, 1, \ldots, q$.

Theorem 4. *With notation as in Theorem 3, we have*

1) $N_0(m) = (p-\bar{q})q$, $N_0(m+1) = \bar{q}q$.
2) *If* $0 < b_g < p-\bar{q}$, *then*

$$N_g(m-1) = \eta \quad (7)$$
$$N_g(m) = q(p-\bar{q}) - 2\eta \quad (8)$$
$$N_g(m+1) = q\bar{q} + \eta, \quad (9)$$

where

$$\eta := mb_g(p - b_g - \bar{q}).$$

3) *If* $p-\bar{q} < b_g < p$, *then*

$$N_g(m) = q(p-\bar{q}) + \theta \quad (10)$$
$$N_g(m+1) = q\bar{q} - 2\theta, \quad (11)$$
$$N_g(m+2) = \theta \quad (12)$$

where

$$\theta := (m+1)(p-b_g)(\bar{q} + b_g - p).$$

The proof is relegated to Appendix B.

Example 2. For the CRT sequences in Example 1 with parameters $p = 3$ and $q = 5$, we tabulate the distribution of Hamming cross-correlation between $s_g(t)$, and $s_1(t)$, for $g \neq 1$, as follows.

g	b_g	$N_g(1)$	$N_g(2)$	$N_g(3)$
0	1	5	10	0
2	2	7	6	2

We note that if $g = 2$, then b_g is 2 mod 3. By Theorem 3, $H_{21}(\tau)$ equals 1, 2, or 3. The Hamming cross-correlation is distributed according to the third part of Theorem 4.

When $q \equiv \pm 1 \bmod p$, the Hamming cross-correlations have three distinct values, for all pairs of distinct CRT sequences chosen from $s_1(t)$, $s_2(t), \ldots, s_{p-1}(t)$.

Theorem 5. *Suppose* $q > p$.

1) *Let* q *be of the form* $mp+1$. *For* $g = 2, 3, \ldots, p-1$, $H_{g1}(\tau)$ *is between* $m-1$ *and* $m+1$.
2) *Let* q *be of the form* $mp+(p-1)$. *For* $g = 2, 3, \ldots, p-1$, $H_{g1}(\tau)$ *is between* m *and* $m+2$.

Proof: For the first part of the theorem, we have $\bar{q}$ equal to 1 mod p. So

$$b_g \equiv (g-1)^{-1}\bar{q} \equiv (g-1)^{-1} \bmod p.$$

When g runs from 2 to $p-1$, the value of $g-1$ runs over all elements in $\mathbb{Z}_p$ except 0 and $p-1$. For odd p, the multiplicative inverse of $p-1$ is itself. Hence the range of $(g-1)^{-1} \bmod p$ is $\mathbb{Z}_p \setminus \{0, -1\}$. We thus obtain $0 < b_g < p-1$. The result now follows from Theorem 3.

The second part can be proved similarly. ∎

B. Hamming Auto-correlation

The distribution of the Hamming auto-correlation can also be determined explicitly.

Theorem 6.

$$H_{00}(\tau) = \begin{cases} q & \text{if } \tau \text{ is a multiple of } p, \\ 0 & \text{otherwise}. \end{cases}$$

For $g = 1, 2, \ldots, p-1$, *and* $k = 0, 1, \ldots, q-1$,

$$H_{gg}(\tau) = \begin{cases} q-k & \text{if } \Phi(\tau) = \pm(g,1)k, \\ 0 & \text{otherwise}. \end{cases}$$

Proof: The first statement follows by the second part of Theorem 1 that $s_0(t)$ has least period p.

For the second statement, we note that $(g,1)t$ runs through the whole group $G_{p,q}$ as t runs from 0 to $pq-1$. This follows from CRT, because for any x and y,

$$gt \equiv x \mod p$$
$$t \equiv y \mod q$$

has one and only one solution mod pq. Via a bijective mapping $t \mapsto (g,1)t$, the set $\mathcal{I}_g$ can be regarded as an arithmetic progression $\{0, 1, 2, \ldots, q-1\}$ in $\mathbb{Z}_{pq}$. The intersection $\mathcal{I}_g \cap (\mathcal{I}_g + (i,j))$ is a subset of $\mathcal{I}_g$ that is also an arithmetic progression. The intersection $\mathcal{I}_g \cap (\mathcal{I}_g + (i,j))$ is nonempty if and only $\pm(i,j) \in \mathcal{I}_g$. If $|\mathcal{I}_g \cap (\mathcal{I}_g + (i,j))|$ is nonzero, then (i,j) is equal to $\pm(g,1)k$ for some $k = 0, 1, \ldots, q-1$, and

$$|\mathcal{I}_g \cap (\mathcal{I}_g + k(g,1))| = q - |k|.$$

∎

IV. Application to the Collision Channel without Feedback

Consider a time-slotted collision channel with K transmitters and one receiver as described in the introductory section in this paper, and assume that there is no cooperation among the users, and no feedback from the receiver. We regard a packet as a Q-ary alphabet chosen from the alphabet set $\Omega = \{1, 2, \ldots, Q\}$. The channel output at slot t equals

$$\begin{cases} 0 & \text{if no user transmits at slot } t, \\ * & \text{if two or more users transmit at slot } t, \\ x & \text{if exactly one user transmits a packet with content } x. \end{cases}$$

Each user is assigned a deterministic and periodic zero-one sequence, called protocol sequence [1]. For $i = 1, 2, \ldots, K$, the protocol sequence associated with user i is denoted by $s_i(t)$, which is a periodic binary sequence of period L. As there is no feedback from the receiver and no cooperation among the users, each user has a relative delay offset τ, which is random but remains fixed throughout the communication session. User i sends a packet at slot t if $s_i(t+\tau) = 1$, and remains silent if $s_i(t+\tau) = 0$.

We consider two different factors in the system model: (i) systems with packet header or without packet header, and (ii) systems with permanent user activity or partial user activity. For system without packet header, the receiver has to identify the set of active users (user identification), and determine the sender of each successfully received packet (packet identification). If packet header is present, information such as user identity is obtained readily by looking up the header, and there is no user and packet identification problem.

For system with partial user activity, the number of active users M is smaller than the total number of potential users T. When a user changes from inactive to active, packets are sent according to the protocol sequence assigned. It is assumed that after all the data are sent, the user must remain inactive for at least L slots before becoming active again. On the other hand, if permanent user activity is assumed, the number of active users M is equal to the total number of users, and each user sends packets periodically according to the protocol sequence.

A. Permanent User Activity with Header

For systems with a packet header, the users are simply identified by packet headers. We calculate the achievable throughput when CRT sequences are used. Let the number of users be M. We pick M sequences from the CRT sequences generated with parameters p and q. To take advantage of the three-valued property mentioned in Theorem 5 we pick a q which is -1 mod p.

Construction 1. Let p be a prime number and k a positive integer. We choose $q = kp - 1$. The CRT construction in Definition 4 yields p protocol sequences of length $L = kp^2 - p$ and weight $w = kp - 1$. We pick any M sequences from this set of CRT sequences.

Since the Hamming weight is $kp - 1$ and the pairwise Hamming cross-correlation is at most $k + 1$ by Theorem 5, the total number of successful packets (summed over all M users) is lower bounded by

$$M[kp - 1 - (M - 1)(k + 1)]. \tag{13}$$

By completing square, we can write the above as

$$(k+1)\left[-\Big(M - \frac{k(p+1)}{2(k+1)}\Big)^2 + \Big(\frac{k(p+1)}{2(k+1)}\Big)^2\right]. \tag{14}$$

Consider (14) as a function of M. The maximum value is obtained when

$$M^* = \frac{k(p+1)}{2(k+1)}, \tag{15}$$

and the maximal value is

$$\frac{(p+1)^2}{4} \cdot \frac{k^2}{k+1}. \tag{16}$$

After dividing the above by the period, we obtain the following lower bound on throughput.

Theorem 7. *Let*

$$\tilde{M} = \Big\lfloor \frac{k(p+1)}{2(k+1)} \Big\rfloor$$

in Construction 1. The system throughput is lower bounded by

$$\frac{1}{4} \cdot \frac{(p+1)^2}{kp^2 - p} \cdot \frac{k^2}{k+1} - \frac{k+1}{kp^2 - p} \tag{17}$$

Proof: Consider the expression in (14) as a function of M, and denote it by $f(M)$. The coefficient of M^2 is $-(k+1)$. If we take $M = \tilde{M}$, we deviate from the optimal value of M^* by at most 1. Therefore,

$$f(M^*) - f(\tilde{M}) = (k+1)(\tilde{M} - M^*)^2 \le k + 1.$$

If we put $M = \lfloor \frac{k(p+1)}{2(k+1)} \rfloor$ in (13), we drop from the optimal value in (16) by at most $k + 1$. Hence the total number of successful packets, divided by the period, is lower bounded by (17). ∎

Theorem 7 provides a lower bound on the worst-case throughput. The mean system throughput, averaged over all delay offsets, is however much higher than the lower bound.

Example 3. We consider an example with $M = 19$ users, using CRT sequences with $p = 37$. The throughput is plotted against sequence length with increasing k, while keeping the duty factor fixed at $1/p$. We compare the lower bound in (17) of worst-case throughput with the average throughput obtained by simulation in Fig. 2. For each k, 100000 delay offset combinations are randomly generated. Beside the mean throughput, the maximum and minimum throughput obtained among these 100000 delay offset combinations are also plotted. We can observe that the mean system throughput is about 0.31, in accordance with the theoretical value

$$M \cdot \frac{1}{p}\Big(1 - \frac{1}{p}\Big)^{M-1} = \frac{19}{37}\Big(1 - \frac{1}{37}\Big)^{18} = 0.314.$$

When the length increases, the lower bound approaches $0.25(p+1)^2/p^2 - 1/p^2 = 0.263$.

Asymptotically, if we increase k and p in such a way that k increases much slower than p, we obtain a lower bound on the system throughput of $1/4$.

Theorem 8. *For arbitrarily small $\epsilon, \delta > 0$, there exists an infinite class of protocol sequence sets of length $O(M^{2+\epsilon})$, where M is the number of sequences, with system throughput lower bounded by $0.25 - \delta$ for all sufficiently large M.*

Proof: We choose $k = \log(p)$. The lower bound of system throughput in Theorem 7 tends to $1/4$, with

$$L \sim kp^2 \sim k(2M)^2 \sim 4M^2 \log M.$$

We can find an integer M_0 sufficiently large such that $4M_0^2 \log M_0$ is less than $M_0^{2+\epsilon}$. The theorem then holds for all $M \ge M_0$. ∎

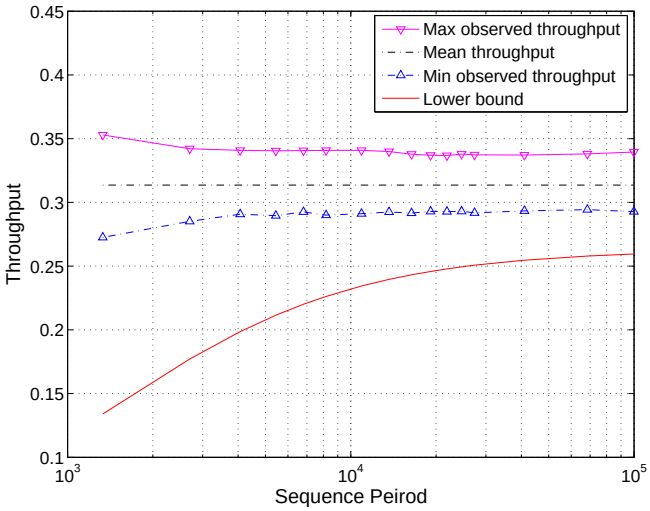

Fig. 2. System Throughput of CRT Sequences for 19 Users, $p = 37$.

B. Partial User Activity without Header

In this application, we use a modified version of the CRT correspondence. Let γ be the multiplicative inverse of p in $\mathbb{Z}_q$, i.e., $\gamma p \equiv 1 \bmod q$. Since p and q are relatively prime, such inverse exists. Define the mapping $\Phi'_{p,q} : \mathbb{Z}_{pq} \to G_{p,q}$ by

$$\Phi'_{p,q}(x) := (x \bmod p, \gamma x \bmod q).$$

It can also be shown that $\Phi'_{p,q}(x)$ is a group isomorphism between $\mathbb{Z}_{pq}$ and $G_{p,q}$.

Because of its unfavorable Hamming auto-correlation property, the sequence generated by $g = 0$ is not used in this application. Given a prime number p, the system supports $p-1$ potential users. We label the users from 1 to $p-1$.

Construction 2. Let p be a prime number, and q be an integer relatively prime to p. Choose γ as described above. For $g = 1, \ldots, p-1$, construct the CRT sequence $s_g(t)$ of length pq by setting

$$s_g(t) = \begin{cases} 1 & \text{if } \Phi'_{p,q}(t) \in \mathcal{I}_{g,p,q} \\ 0 & \text{otherwise.} \end{cases}$$

The sequence generated by g is assigned to user g. We have used $\Phi'_{p,q}(x)$ in the above construction, instead of $\Phi_{p,q}(x)$ as in Construction 1. The cross-correlation properties of the resulting modified CRT sequences are exactly the same as in the previous section, because the proof in the previous section is essentially about two-dimensional Hamming cross-correlation. If Φ is replaced by Φ', all theorems about Hamming cross-correlation and auto-correlation also hold.

Sequence synchronization

For systems without a header, we need to identify the sender of a successful packet. Also, as users may come and go, we also need to determine when a user becomes active. We show that these two tasks can be achieved by merely observing the channel activity, that is whether a time slot is idle, containing a collision or a successful transmission, without looking into the packet contents.

Definition 6. For each time index t, let $c(t)$ be 0 if it is an idle slot, 1 if exactly one user transmits a packet, or * if two or more users transmit. We call $c(t)$ the *channel-activity signal*. We say that $c(t)$ is *matched* to $s_i(t)$ at time t_0 if $\forall t = 0, 1, \ldots, L-1$, $s_i(t) = 1 \Rightarrow c(t_0 + t) = 1$ or $*$.

The receiver stores the channel-activity signal in a first-in-first-out queue.

We want to determine (a) the time when a user becomes active, and (b) the time when an active user change status from active to idle. The receiver keeps track of the active users by maintaining $p-1$ Boolean variables $active(i)$, for $i = 1, 2, \ldots, p-1$. The value of $active(i)$ is set to `FALSE` if the user is idle, and `TRUE` if the user is active. To determine whether user i becomes active at time t_0, the receiver wait until time $t_0 + L$. At that time the channel-activity signal $c(t_0)$, $c(t_0+1), \ldots, c(t_0+L-1)$ are available. The receiver then checks whether $c(t)$ is matched to $s_i(t)$ at time t_0. If there is a match, we declare that user i is active and the starting time of user i is stored in variable $start(i)$. We summarizes the procedure in Algorithm 1 below.

Algorithm 1 Determining when a user becomes active or inactive at time t_0.

```
for i = 1, 2, ..., p − 1 do
    if active(i) = FALSE then
        if c(t) is matched to s_i(t) at t_0 then
            active(i) ← TRUE
            start(i) ← t_0
        end if
    else
        if t_0 − start(i) is a multiple of L and c(t) is not
        matched to s_i(t) at time t_0 then
            active(i) ← FALSE
            start(i) ← 0
        end if
    end if
end for
```

Under some conditions on the number of active users and period length, we can show that the above algorithm is able to identify the starting time of each user.

Theorem 9. *Let p be a prime number, and $q > 2p^2$ be an integer relatively prime to p. Construct $p-1$ CRT sequences by Construction 2 and assign them to users 1 to $p-1$. Suppose that at most $(p+1)/2$ users are active at the same time. Then Algorithm 1 can successfully identify the active users and determine their starting time.*

The proof of Theorem 9 is given in Appendix C.

Example 4. Consider the CRT sequences generated with parameters $p = 7$, $q = 8$ and $\gamma = 7$. The period is equal to 56. Suppose that users 1, 2, 3, 4 and 6 are active. The relative delay offsets of users 1 and 2 are 0, and the relative delay offsets of users 3, 4 and 6 are 1. The CRT sequences $s_1(t)$

$$
\begin{aligned}
s_1(t) &: \underbrace{10001000\ \ 00000001\ \ 00010000\ \ 00000010\ \ 00100000\ \ 00000100\ \ 01000000}\ \ 10\\
s_2(t) &: \underbrace{10010000\ \ 00000100\ \ 00000001\ \ 00100000\ \ 00001000\ \ 00000010\ \ 01000000}\ \ 10\\
s_3(t-1) &: 0\underbrace{1000001\ \ 00000100\ \ 00010000\ \ 00000000\ \ 10000010\ \ 00001000\ \ 00100000\ \ 0}1\\
s_4(t-1) &: 0\underbrace{1000010\ \ 00010000\ \ 00000001\ \ 00001000\ \ 00000000\ \ 10000100\ \ 00100000\ \ 0}1\\
s_6(t-1) &: 0\underbrace{1000000\ \ 00000000\ \ 00000000\ \ 00000000\ \ 00000000\ \ 00000000\ \ 00111111\ \ 1}1\\
\\
c(t) &: **011011\ \ 00010*01\ \ 000*000*\ \ 00101010\ \ 10101010\ \ 10001*10\ \ 0**11111\ \ **\\
\\
s_6(t) &: \underbrace{10000000\ \ 00000000\ \ 00000000\ \ 00000000\ \ 00000000\ \ 00000000\ \ 01111111}
\end{aligned}
$$

Fig. 3. CRT sequences in Example 4. Sequence periods are indicated by underbraces.

to $s_4(t)$ and $s_6(t)$ and the induced channel-activity signal $c(t)$ are shown in Fig. 3

By comparing with $s_6(t)$, the receiver declares that the channel-activity signal is matched to $s_6(t)$ at time 0. The receiver cannot distinguish whether user 6 starts transmitting at time 0 or 1, and erroneously detects that user 6 becomes active at time 0, while the actual delay offsets is 1. Thus Algorithm 1 fails in this case.

The preceding example illustrates that Algorithm 1 does not work if there are too many active users. Theorem 9 asserts that such error in synchronization does not occur if $q \geq 2p^2$ and the number of simultaneously active users does not exceed $(p+1)/2$.

If $q \equiv \pm 1 \bmod p$, the Hamming cross-correlation of the resulting CRT sequences is three-valued (see Theorem 5). For these preferred choices of q, we can improve Theorem 9 by relaxing the requirement $q > 2p^2$ to $q > p^2$.

Theorem 9'. *Let p be a prime number, and $q > p^2$ be an integer such that $q \equiv \pm 1 \bmod p$. Construct $p-1$ CRT sequences by Construction 2 and assign them to users 1 to $p-1$. Then Algorithm 1 can identify the active users and determine their starting time, provided that the number of simultaneously active users is no more than $(p+1)/2$.*

The proof of Theorem 9' is analogous to the proof of Theorem 9 in Appendix C.

Erasure Correction and Throughput

We have shown in Theorem 9 that the receiver is able to figure out the delay offsets of the active users. From a specific user's point of view, the channel reduces to an erasure channel.

In the remainder of this section, we pick q to be an integer of the form $kp+1$, for some integer $k \geq p$, so that q is larger than p^2 and Theorem 9' can be applied.

For each user, the number of successfully received packets in a length is lower bounded by

$$(kp+1) - \Big(\frac{p+1}{2} - 1\Big)(k+1). \tag{18}$$

The first term $kp+1$ is the total number of packets sent by a user in a period. The factor $k+1$ is the maximum Hamming cross-correlation from Theorem 5. After some simplifications, (18) can be written as

$$D(p,k) := \frac{1}{2}[kp + k - p + 3]. \tag{19}$$

An erasure-correcting code can be applied across the packets in a period. We pick Q to be a power of prime such that $Q \geq kp+1$. We encode $D(p,k)$ information packets using a shortened Reed-Solomon (RS) code of length $kp+1$ over the finite field of size Q. The $kp+1$ encoded packets are sent out according to the assigned protocol sequence. Because of the maximal-distance separable property of RS codes, we can recover the information packets. We have thus proved the following

Theorem 10. *Suppose that there are $(p+1)/2$ active users out of $p-1$ potential users in the collision channel without feedback, where p is an odd prime. Each active user can send $D(p,k)$ information packets in a period of $p(kp+1)$ time slots, where $D(p,k)$ is given in* (19) *and k is an integer larger than or equal to p. In particular, if we take $k=p$, the resultant CRT sequences have period p^3+p, and when $(p+1)/2$ users are active, the system throughput is lower bounded by*

$$\frac{p+1}{2} \cdot \frac{0.5(p^2+3)}{p^3+p} \geq 0.25.$$

Example 5. We pick $p = 19$, $k = 19$ and $q = p^2 + 1 = 362$ in Theorem 10. Generate 18 CRT sequences of length $pq = 6878$ by Construction 1. Pick $Q = 512 = 2^9$, which is a prime power larger than the Hamming weight of the CRT sequences under consideration. Using a shortened RS code of length 362 and dimension $D(19,19) = 182$ over the finite field of size 512, we encode 182 information packets in each period for each active user. When 10 users are active, the total number of information packets sent through the system is $182 \times 10 = 1820$, achieving system throughput no less than $1820/6878 = 0.265$.

V. Comparison with Other Protocol Sequences

In order to compare the variation of Hamming cross-correlation due to delay offsets, we introduce in this section a measure of deviation called ϵ-uniformity.

Given any two binary and periodic sequences $a(t)$ and $b(t)$, let $\mathbb{E}_\tau[H_{ab}(\tau)]$ be the expectation of Hamming cross-correlation, with delay offset τ chosen uniformly at random over a period.

Definition 7. We say that the Hamming cross-correlation H_{ab} is *ϵ-uniform* if

$$\frac{|H_{ab}(\upsilon) - \mathbb{E}_\tau[H_{ab}(\tau)]|}{\mathbb{E}_\tau[H_{ab}(\tau)]} \leq \epsilon, \tag{20}$$

for all $\tau = 0, 1, \ldots, L-1$. A sequence set is called *ϵ-uniform* if ϵ is the smallest number such that each pair of distinct sequences is ϵ-uniform. We say that a sequence set is *pairwise shift-invariant* if it is 0-uniform.

In other words, $H_{ab}(\tau)$ is ϵ-uniform if for all delay offsets τ, the percentage difference between $H_{ab}(\tau)$ and the mean is between $-\epsilon$ and ϵ. The notion of ϵ-uniformity is the same as the normalized ℓ_∞ distance between Hamming cross-correlation and the expected value.

The sum of the Hamming cross-correlation over all relative delay offsets in a period equals $w_a w_b$, where w_a and w_b denote the Hamming weight of $a(t)$ and $b(t)$ respectively (cf. Lemma 3 in Appendix B). If we take the average over all delay offsets τ, then

$$\mathbb{E}_\tau[H_{ab}(\tau)] = \frac{w_a w_b}{L}.$$

Hence, the definition of ϵ-uniformity in (20) can be written as

$$\frac{|H_{ab}(\upsilon) - w_a w_b/L|}{w_a w_b/L} \leq \epsilon.$$

A lower bound on the worst-case throughput similar to (17) can be expressed in terms of ϵ-uniformity. Suppose that there are K active users and each of them is assigned a sequence from an sequence set of length L and Hamming weight w. Suppose that the sequence set is ϵ-uniform. By the union bound, a user can successfully send at least

$$w - (K-1)(1+\epsilon)(w^2/L)$$

packets in each period. Individual throughput is lower bounded by

$$f - (K-1)(1+\epsilon)f^2,$$

where f is the duty factor w/L. It can be easily seen that for fixed duty factor f and number of users K, a smaller ϵ yields a larger lower bound on individual throughput.

Example 6. (Constant-weight cyclically permutable codes [5]) In [5, p.948] Example 5, a protocol sequence set of period 156 and Hamming weight 12 is presented. The number of sequences is 169. It is shown that the Hamming cross-correlation is between 0 and 3, and the mean Hamming cross-correlation equals $(12^2)/156 = 12/13$. The maximal deviation from the mean is $3 - 12/13 = 27/13$. This sequence set is thus $(27/12)$-uniform.

Example 7. (Prime sequences [13]) Given a prime p, we can construct a sequence set with period p^2, Hamming weight p and Hamming cross-correlation no more than 2. The mean Hamming cross-correlation is 1. For each pair of distinct prime sequences, the maximal $|H_{ab}(\upsilon) - \mathbb{E}_\tau[H_{ab}(\tau)]|$ over all possible delay offset υ is equal to 1. The prime sequences are therefore 1-uniform.

Example 8. (Extended prime sequences [14]) By padding $p-1$ zeros after every "1" in a prime sequence, we obtain a sequence set with period $p(2p-1)$, Hamming weight p and Hamming cross-correlation either 0 or 1. The mean Hamming cross-correlation is $p/(2p-1)$, which is roughly equal to one half. We can check that the extended prime sequences are 1-uniform.

Example 9. (Wobbling sequences [10]) Based on prime sequences, a class of $(1/p)$-uniform sequence sets is constructed in [10]. The number of sequences is p and the sequence period is p^4.

Example 10. (Shift-invariant and pairwise shift-invariant sequences [1]–[4]) All pairwise shift-invariant is by definition 0-uniform. It is shown in [2] that the period grows exponentially in the number of sequences. For pairwise shift-invariant sequences, the period is also shown to grow exponentially with the number of users [4].

The examples above are presented in an order such that the ϵ-uniformity is decreasing. We can see that the sequence period increases as we go down the list.

Theorem 11. *The sequences obtained by Construction 1 with period $O(kp^2)$ are $(1/k)$-uniform.*

Proof: The Hamming cross-correlation in Construction 1 is obtained by Theorem 5. The mean Hamming cross-correlation is

$$\frac{w^2}{L} = \frac{(kp-1)^2}{p(kp-1)} = \frac{kp-1}{p}.$$

The maximal difference between Hamming cross-correlation and the mean is

$$\begin{aligned}\left|k \pm 1 - \frac{kp-1}{p}\right| &= \frac{1}{p}|\pm p - 1| \\ &\leq \frac{1}{p}(p+1) \\ &= \frac{kp-1}{p}\Big(\frac{p+1}{kp-1}\Big) \\ &= \frac{kp-1}{p}\,O(\frac{1}{k}).\end{aligned}$$

We thus have an $O(1/k)$-uniform sequence set of period $O(kp^2)$. ∎

The trade-off between period length and ϵ-uniformity is summarized in Table I.

To compare with the wobbling sequences, we can take k roughly equal to p. This yields CRT sequences with are $(1/p)$-uniform and period $O(p^3)$. The ϵ-uniformity is roughly the same as wobbling sequences but the period is shorter than the period of the wobbling sequences.

	Period	ϵ
Prime sequences	p^2	1
Extended prime sequences	$p(2p-1)$	1
CRT sequences	$O(kp^2)$	$O(1/k)$
Wobbling sequences	p^4	$1/p$
Shift-invariant sequences	exponential in p	0

TABLE I
TRADEOFF BETWEEN PERIOD LENGTH AND ϵ-UNIFORMITY FOR VARIOUS SEQUENCE SETS FOR p USERS

VI. OTHER APPLICATIONS

A. Extension to Multiple Data Rates

To support service with multiple data rates, we need protocol sequences with different duty factors; a sequence with larger duty factor is assigned to a user with higher data rate requirement. The CRT construction can be extended to cope with multiple data rates. For the sequence $s_{g,p,t}(t)$ generated by g, we replace the characteristic set $\mathcal{I}_{g,p,q}$ by

$$\mathcal{I}_{g,p,q} \cup (\mathcal{I}_{g,p,q} + (1,0)) \cup \ldots \cup (\mathcal{I}_{g,p,q} + (k,0))$$

for some positive integer k. We note that the above is a union of disjoint sets. The resulting sequence has duty factor k/p. The measure of uniformity however does not change; if the original CRT sequence set is ϵ-uniform, the extended sequence set is still ϵ-uniform.

B. Application to Multi-Channel Network

In this network, users can send data to each other, and we have a fully connected system topology. The total bandwidth is divided into p subchannels and each subchannel is assigned to a user. Let the subchannel assigned to user i be denoted by subchannel i, for $i = 1, 2, \ldots, p$. Each user is also assigned a CRT sequence. The half-duplex model is assumed, so that each user cannot transmit and receive at the same time. Users always receive in their assigned subchannel. In a period of L slots, user i can pick one user, say user j, and send packets to user j via subcarrier j using user i's CRT sequence. In one sequence period, user i either: (i) receives packets in subcarrier i for the whole period, or (ii) sends packets to user j in L/p packets using subcarrier j and receives packets in the remaining $L - L/p$ packets in subcarrier i.

The worst-case scenario occurs when user i is sending packets to some other user, and all the remaining users want to send packets to user i in the same period. Using CRT sequences, we can show as in the multiple-access case that the worst-case throughput between each pair of users is lower bounded by a positive constant.

VII. CONCLUSION

A class of protocol sequences, called CRT sequences, whose Hamming cross-correlation is highly concentrated around the mean value is given. When CRT sequences are applied to the collision channel without feedback, we obtain a tradeoff between worst-case throughput and the sequence period. The generation of CRT sequences involves only simple modular arithmetics, and provides a low-complexity solution to multiple-accessing in wireless sensor network.

APPENDIX A
PROOF OF THEOREM 3

In this appendix, g denotes an element in $\mathbb{Z}_p \setminus \{0, 1\}$, and

$$a_g(\tau_1, \tau_2) := (g-1)^{-1}(\tau_1 - \bar{\tau}_2) \bmod p. \tag{21}$$

As defined in (5), b_g denotes $(g-1)^{-1}\bar{q} \bmod p$. We will also use the indicator function $\mathbf{I}$ defined as

$$\mathbf{I}(P) := \begin{cases} 1 & \text{if } P \text{ is true,} \\ 0 & \text{if } P \text{ is false.} \end{cases}$$

Recall that $\bar{x}$ stands for the remainder of x after division by p.

Lemma 2. *The Hamming cross-correlation $H_{g1}(\tau_1, \tau_2)$ as defined in* (2) *satisfies the following properties:*

1) *$H_{g1}(\tau_1, \tau_2)$ equals the number of solutions to*

$$\bar{x} \equiv a_g(\tau_1, \tau_2) \oplus_p b_g \mathbf{I}(0 \le x < \tau_2) \bmod p, \tag{22}$$

   for $x = 0, 1, \ldots, q-1$.
2) *Let (τ_1, τ_2) and (τ_1', τ_2') denote two (2-dimensional) relative delay offsets. If $\tau_1 = \tau_1'$ and $\tau_2 = \tau_2' + kp$ for some integer k, then*

$$H_{g1}(\tau_1, \tau_2) = H_{g1}(\tau_1', \tau_2').$$

Proof: After setting the value of h in (4) to 1, we obtain

$$\bar{x}g \equiv (\overline{x \ominus_q \tau_2}) \oplus_p \tau_1 \bmod p. \tag{23}$$

We want to show that the number of solutions to (23), for $x = 0, 1, \ldots, q-1$, is the same as the number of solutions to (22).

Consider x in two disjoint ranges: (i) $0 \le x < \tau_2$, and (ii) $\tau_2 \le x < q$. In the first case, $x \ominus_q \tau_2$ is congruent to $x + q - \tau_2 \bmod q$. So, for $0 \le x < \tau_2$, (23) is equivalent to

$$\bar{x}g \equiv \bar{x} \oplus_p \bar{q} \ominus_p \bar{\tau}_2 \oplus_p \tau_1 \bmod p \tag{24}$$

where $\bar{q}$ and $\bar{\tau}_2$ are residues of q and τ_2 in $\mathbb{Z}_p$, respectively.

In the second case, for $x = \tau_2, \tau_2 + 1, \ldots, q-1$, (23) is equivalent to

$$\bar{x}g \equiv \bar{x} \ominus_p \bar{\tau}_2 \oplus_p \tau_1 \bmod p. \tag{25}$$

We combine (24) and (25) in one line as

$$\bar{x}(g-1) \equiv -\bar{\tau}_2 \oplus_p \tau_1 \oplus_p \bar{q}\mathbf{I}(0 \le x < \tau_2) \bmod p.$$

Since g is not equal to 1 by assumption, we can divide by $(g-1)$ and obtain (22). This proves the first part of the lemma.

The second part of the lemma is vacuous if $q < p$. So we assume $q > p$. (The case $q = p$ is excluded because it is assumed that q is relatively prime with p.) It is sufficient to prove the statement for $\tau_1 = \tau_1'$ and $\tau_2' = \tau_2 + p$, namely, the number of solutions to (22) and the number of solutions to

$$\bar{x} \equiv a_g(\tau_1, \tau_2 + p) + b_g \mathbf{I}(0 \le x < \tau_2 + p) \bmod p \tag{26}$$

for $x = 0, 1, \ldots, q-1$, are the same. We note that $a_g(\tau_1, \tau_2)$ is equal to $a_g(\tau_1, \tau_2 + p)$. However, the arguments inside the

indicator function are different. We divide the range of x into three disjoint parts:

$$\mathcal{X}_1 := \{0, 1, \ldots, \tau_2 - 1\},$$
$$\mathcal{X}_2 := \{\tau_2, \tau_2 + 1, \ldots, \tau_2 + p - 1\},$$
$$\mathcal{X}_3 := \{\tau_2 + p, \tau_2 + p + 1, \ldots, q - 1\}.$$

For $x \in \mathcal{X}_1$,

$$\mathbf{I}(0 \le x < \tau_2) = \mathbf{I}(0 \le x < \tau_2 + p) = 1.$$

Therefore (22) and (26) have the same number of solutions for x in $\mathcal{X}_1$. For $x \in \mathcal{X}_2$, both (22) and (26) have exactly one solution by Lemma 1. For $x \in \mathcal{X}_3$, we have

$$\mathbf{I}(0 \le x < \tau_2) = \mathbf{I}(0 \le x < \tau_2 + p) = 0,$$

and hence (22) and (26) have the same number of solutions for $x \in \mathcal{X}_3$. In conclusion, the number of solutions to (22) and (26) for $x \in \mathcal{X}_1 \cup \mathcal{X}_2 \cup \mathcal{X}_3$ are the same. This finishes the proof of the second part of the lemma. ∎

Proof of Theorem 3: By the second part of the previous lemma, we only need to consider $\tau_2 = 0, 1, \ldots, p-1$. We first consider the case when b_g is between 1 and $p - \bar{q} - 1$. We further consider two subcases.

(i) $0 \le \tau_2 < \bar{q}$:

Suppose that (22) has no solution for $0 \le x < \tau_2$. As the indicator function in (22) is zero for $x = \tau_2, \tau_2 + 1, \ldots, q-1$, (22) is reduced to

$$\bar{x} \equiv a_g(\tau_1, \tau_2) \bmod p.$$

The number of integers in $\{\tau_2, \tau_2 + 1, \ldots, q - 1\}$, say d, satisfies $\lfloor d/p \rfloor = m$. By Lemma 1, we have either m or $m+1$ solutions to (22) for $x \ge \tau_2$.

Secondly, suppose that (22) has exactly one solution for $0 \le x < \tau_2$. The indicator function in (22) is equal to 1 for $0 \le x < \tau_2$. Hence,

$$0 \le a_g(\tau_1, \tau_2) + b_g < \tau_2. \tag{27}$$

We claim that (22) has no solution for $x = \tau_2, \tau_2 + 1, \ldots, \bar{q} - 1$. Otherwise, we have

$$\tau_2 \le a_g(\tau_1, \tau_2) < \bar{q},$$

which, after combining with the assumption that $1 \le b_g \le p - \bar{q} - 1$, yields

$$\tau_2 < a_g(\tau_1, \tau_2) + b_g < p - 1.$$

This contradicts with (27) and proves the claim. For

$$\bar{q} \le x < q,$$

there are exactly m solutions by Lemma 1. The total number of solutions to (22) for $x = 0, 1, \ldots, q - 1$, is thus $m + 1$. Hence $H_{g1}(\tau_1, \tau_2) = m + 1$.

(ii) $\bar{q} \le \tau_2 < p$:

By Lemma 1, (22) has either 0 or 1 solution for $0 \le x \le \tau_2$, and either $m - 1$ or m solutions for $\tau_2 \le x < q$. Hence, $H_{g1}(\tau_1, \tau_2)$ is within the range of $\{m - 1, m, m + 1\}$.

For $b_g = p - \bar{q} + 1, \ldots, p - 1$, we again consider two subcases.

(i) $0 \le \tau_2 < \bar{q}$:

By Lemma 1, (22) has either 0 or 1 solution for $0 \le x < \tau_2$, and either m or $m + 1$ solutions for $\tau_2 \le x < q$. Therefore, $H_{g1}(\tau_1, \tau_2) \in \{m, m + 1, m + 2\}$.

(ii) $\bar{q} \le \tau_2 < p$:

Suppose that (22) has no solution for $0 \le x < \tau_2$, i.e.,

$$\tau_2 \le a_g(\tau_1, \tau_2) + b_g < p. \tag{28}$$

We claim that (22) must have one solution for x in the following range

$$\tau_2 \le x < p + \bar{q}. \tag{29}$$

From the assumption of $\bar{q} \le \tau_2 < p$, we deduce that

$$\bar{q} < p + \bar{q} - \tau_2 \le p,$$

so that the range in (29) is non-empty and consists of no more than p integers. If the claim were false, we would have no solution to (22) for $\tau_2 \le x < p + \bar{q}$, implying that

$$\bar{q} \le a_g(\tau_1, \tau_2) < \tau_2. \tag{30}$$

Here, we have used the fact that the indicator function in (22) is equal to zero for x in the range in (29). By adding (30) to

$$p - \bar{q} + 1 \le b_g \le p - 1$$

and reducing mod p, we obtain

$$1 \le a_g(\tau_1, \tau_2) + b_g < \tau_2,$$

which is a contradiction to (28). Thus, the claim is proved. For $x = p + \bar{q}, p + \bar{q} + 1, \ldots, q - 1$, there are exactly $m - 1$ solutions to (22) by Lemma 1. Totally there are m solutions, and thus $H_{g1}(\tau_1, \tau_2) = m$.

Finally suppose that (22) has exactly one solution for $0 \le x < \tau_2$. As the number of solutions to (22) for $x = \tau_2, \tau_2 + 1, \ldots, q-1$ is either $m-1$ or m by Lemma 1, the total number of solutions to (22) is either m or $m + 1$.

In any case, we see that $H_{g1}(\tau_1, \tau_2)$ is either m, $m + 1$ or $m + 2$. ∎

Appendix B
Proof of Theorem 4

We use the following property of Hamming cross-correlation which holds for any binary sequence set in general [15].

Lemma 3. *Let $a(t)$ and $b(t)$ be binary sequences with period L and Hamming weight q. Then*

$$\sum_{\tau=0}^{L-1} H_{ab}(\tau) = q^2.$$

We include the short proof for completeness.

Proof:

$$\sum_{\tau=0}^{L-1} H_{ab}(\tau) = \sum_{\tau=0}^{L-1}\sum_{t=0}^{L-1} a(t)b(t-\tau)$$
$$= \sum_{t=0}^{L-1} a(t) \sum_{\tau=0}^{L-1} b(t-\tau)$$
$$= \sum_{t=0}^{L-1} a(t) \sum_{\tau=0}^{L-1} b(\tau) = q^2.$$

The last equality follows from the assumption that the Hamming weights of $a(t)$ and $b(t)$ are both q. ∎

1) By Theorem 2, $N_0(j)$ is nonzero only when $j = m$ or $j = m+1$, whence,

$$N_0(m) + N_0(m+1) = pq. \tag{31}$$

On the other hand, we have

$$mN_0(m) + (m+1)N_0(m+1) = q^2 \tag{32}$$

from Lemma 3. We can eliminate $N_0(m)$ from (31) and (32) and obtain

$$mpq + N_0(m+1) = q^2,$$

which implies $N_0(m+1) = q(q-mp) = q\bar{q}$. Substituting this into (31), we get $N_0(m) = pq - N_0(m+1) = (p-\bar{q})q$. This proves the first part of Theorem 4.

2) Suppose that $b_g = 1, 2, \ldots, p-\bar{q}-1$. We can set up a system of two linear equations in three variables $N_g(m-1)$, $N_g(m)$ and $N_g(m+1)$:

$$\sum_{k=m-1}^{m+1} N_g(k) = pq$$
$$\sum_{k=m-1}^{m+1} kN_g(k) = q^2.$$

The second equality is due to Lemma 3. Solving for $N_g(m)$ and $N_g(m+1)$ in terms of $N_g(m-1)$, we obtain (7) to (9). It remains to evaluate $N_g(m-1)$. The proof is completed by showing the following two claims:

(i) For each $k = 0, 1, \ldots, m-1$, there are exactly

$$b_g(p - b_g - \bar{q}) \tag{33}$$

order pairs (τ_1, τ_2), with $0 \le \tau_1 < p$ and $\tau_2 = kp, kp+1, \ldots, (k+1)p-1$, such that $H_{g1}(\tau_1, \tau_2) = m-1$.

(ii) For $\tau_2 = mp, mp+1, \ldots, q-1$, $H_{g1}(\tau_1, \tau_2)$ does not equal $m-1$ for all $0 \le \tau_1 < p$.

Multiplying (33) by m, we obtain

$$N_g(m-1) = mb_g(p - b_g - \bar{q}).$$

In the following, we complete the proof of part 2 in the theorem by showing (i) and (ii).

By Lemma 2 part 2, we notice that $H_{g1}(\tau_1, \tau_2)$ depends on τ_2 only through the residue of $\tau_2 \bmod p$. Hence it is sufficient prove claim (i) for $k = 0$.

Consider τ_2 from 0 to $p-1$. We want to count the number of times that $H_{g1}(\tau_1, \tau_2) = m-1$, for $0 \le \tau_1, \tau_2 < p$. We have shown in Lemma 2 that $H_{g1}(\tau_1, \tau_2)$ can be computed by counting the number of solutions to (22) for x between 0 and $q-1$. Partition the range of x into two disjoint subsets

$$\mathcal{X}_1 := \{0, 1, \ldots, p+\bar{q}-1\}, \text{ and}$$
$$\mathcal{X}_2 := \{p+\bar{q}, p+\bar{q}+1, \ldots, q-1\},$$

and consider the number of solutions to (22) for x in $\mathcal{X}_1$ and $\mathcal{X}_2$ separately. Because $0 \le \tau_2 < p$, the indicator function $\mathbf{I}(0 \le x < \tau_2)$ in (22) is identically equal to 0 for $x \in \mathcal{X}_2$. The number of solutions to (22) for $x \in \mathcal{X}_2$ is exactly $m-1$ by Lemma 1. The problem reduces to counting the number of pairs $(\tau_1, \tau_2) \in \mathbb{Z}_p^2$ such that (22) has no solution for $x \in \mathcal{X}_1$. Observe that a_g depends on τ_1 and τ_2 through their difference $\tau_1 - \tau_2 \bmod p$. We make a change of variables $(\tau_1, \tau_2) \to (u, \tau_2)$ by defining u as

$$u := (g-1)^{-1}(\tau_1 - \tau_2) \bmod p. \tag{34}$$

Our objective now is to count the number of ordered pairs $(u, \tau_2) \in \mathbb{Z}_p^2$ such that the equation

$$x \equiv u + b_g \mathbf{I}(0 \le x < \tau_2) \bmod p \tag{35}$$

has no solution for $x \in \mathcal{X}_1$. We note that b_g does not depend on τ_2 and u.

If

$$0 \le u < \bar{q}, \tag{36}$$

then (35) has at least one solution over $x \in \mathcal{X}_1$, namely $x = (u \bmod p) + p$. Indeed, as $\tau_2 < p \le x$, the indicator function $\mathbf{I}(0 \le x < \tau_2)$ is evaluated to 0, and thus if we put $x = (u \bmod p) + p$ in (35), we have $(u \bmod p) + p \equiv u \bmod p$, which obviously holds.

On the other hand, if

$$p - b_g \le u < p, \tag{37}$$

then (35) also has at least one solution over $x \in \mathcal{X}_1$ no matter what τ_2 is. Indeed, for $0 \le \tau_2 \le u$, we can set $x = u$. Then $\mathbf{I}(0 \le x < \tau_2)$ equals 0, and we see that $x = u$ is a solution to (35). For $u < \tau_2 < p$, we can set $x = u + b_g - p$. Then we have $x < u \le \tau_2$ and whence $\mathbf{I}(0 \le x < \tau_2) = 1$. When $u < \tau_2 < p$ and $x = u + b_g - p$, (35) becomes

$$u + b_g - p \equiv u + b_g \bmod p.$$

From (36) and (37), $H_{g1}(\tau_1, \tau_2)$ is equal to $m-1$ only when $u = \bar{q}, \bar{q}+1, \ldots, p-b_g-1$. For each such u, we now count the number of $\tau_2 \in \mathbb{Z}_p$ for which (35) has no solution over $x \in \mathcal{X}_1$. If x is a solution of (35), then x can take only two values, namely u or $u + b_g$. When $x = u$ is a solution, x must satisfy $x \ge \tau_2$; when $x = u + b_g$ is a solution, x must satisfy $0 \le x < \tau_2$. So, (35) has no solution over $x \in \mathcal{X}_1$ if and only if for all $x \in \mathcal{X}_1$, we have $u < \tau_2$ and $u + b_g \ge \tau_2$. Putting these two inequalities together, we obtain $u < \tau_2 \le u + b_g$. Consequently, for each $u = \bar{q}, \bar{q}+1, \ldots, p-b_g-1$, there are exactly b_g values of τ_2 such that $H_{g1}(\tau_1, \tau_2) = m-1$. This proves claim (i).

For claim (ii), we count the solutions to (35) for $0 \le x\tau_2$. Since $\mathbf{I}(0 \le x < \tau_2)$ is identically equal to 1, by Lemma 1, the number of solutions to (35) over $0 \le x < \tau_2$ is at least m. So $H_{g1}(\tau_1, \tau_2)$ cannot be $m-1$.

This ends the proof of the second part of Theorem 4.

3) We set up the following system of linear equations in variables $N_g(m)$, $N_g(m+1)$ and $N_g(m+2)$:

$$\sum_{k=m}^{m+2} N_g(k) = pq$$
$$\sum_{k=m}^{m+2} kN_g(k) = q^2.$$

After solving for $N_g(m)$ and $N_g(m+1)$ in terms of $N_g(m+2)$, we obtain (10) to (12). So, we just need to evaluate $N_g(m+2)$.

The evaluation of $N_g(m+2)$ relies on the following two claims:

(i) For each $k = 0, 1, 2, \ldots, m$, there are exactly

$$(p - b_g)(\bar{q} + b_g - p) \tag{38}$$

ordered pairs (τ_1, τ_2), with $0 \le \tau_1 < p$ and $kp \le \tau_2 < kp + \bar{q}$, such that $H_{g1}(\tau_1, \tau_2) = m + 2$.

(ii) For $k = 0, 1, \ldots, m-1$, none of the ordered pair (τ_1, τ_2) with $0 \le \tau_1 < p$ and $kp + \bar{q} \le \tau_2 < (k+1)p$, satisfies $H_{g1}(\tau_1, \tau_2) = m + 2$.

Since (i) and (ii) exhaust all possible $(\tau_1, \tau_2) \in G_{p,q}$, we multiply (38) by $(m+1)$ and obtain

$$N_g(m+2) = (m+1)(p - b_g)(\bar{q} + b_g - p).$$

We prove case (i) and (ii) in the rest of this appendix.

By the second part of Lemma 2, it is sufficient to prove case (i) for $k = 0$. Consider τ_2 from 0 to $\bar{q} - 1$. Divide the range of x into two disjoint parts:

$$\mathcal{X}_3 := \{0, 1, \ldots, \bar{q} - 1\}, \text{ and}$$
$$\mathcal{X}_4 := \{\bar{q}, \bar{q} + 1, \ldots, q - 1\}.$$

As in the proof of the second part of this theorem, we make a change of variable $(\tau_1, \tau_2) \to (u, \tau_2)$ by defining u as in (34). It is noted that $\mathcal{X}_4$ consists of mp consecutive integers, and the indicator function $\mathbf{I}(0 \le x < \tau_2)$ in (35) equals 0 for all $x \in \mathcal{X}_4$. By Lemma 1, there are exactly m solutions to (35) for $x \in \mathcal{X}_4$. So, $H_{g1}(\tau_1, \tau_2)$ equals $m + 2$ if and only if (35) has exactly two solutions over $x \in \mathcal{X}_3$. It reduces the problem to counting the number of pairs (u, τ_2), with $0 \le u < p$ and $0 \le \tau_2 < \bar{q}$, such that (35) has exactly two solutions over $x \in \mathcal{X}_3$.

The only two candidate solutions for (35) are $x = u$ and $x = u + b_g$. We investigate under what condition both u and $u + b_g$ are indeed solutions to (35). $x = u$ is a solution only if $\mathbf{I}(0 \le x < \tau_2) = 0$. This implies that $u < \tau_2$. $x = u + b_g$ is a solution only if $\mathbf{I}(0 \le x < \tau_2) = 1$. Hence $u + b_g < \tau_2$. Combining these two conditions, we obtain

$$u + b_g \le \tau_2 < u.$$

This is possible only if $u + b_g \ge p$, and thus after reduction mod p, we have $(u + b_g \bmod p) < u$. The first necessary condition for both u and $u + b_g$ are solutions to (35) is

$$p - b_g \le u < p. \tag{39}$$

Secondly, as it is required that $x \in \mathcal{X}_3$, we must have $u \in \mathcal{X}_3$, i.e.,

$$0 \le u < \bar{q} \tag{40}$$

Putting (39) and (40) together, we have the following necessary condition on u.

$$p - b_g \le u < \bar{q}. \tag{41}$$

We note that the range of u in (41) is nonempty, because $p - b_g < \bar{q}$ by assumption. u can take on any of the $(b_g + \bar{q} - p)$ value in (41), and for each such u, τ_2 can assume values in $u - \{0, 1, \ldots, p - b_g - 1\}$. Hence, the total number of pairs (τ_1, τ_2) such that $0 \le \tau_2 < \bar{q}$ and $H_{g1}(\tau_1, \tau_2) = m + 2$ is $(p - b_g)(b_g - (p - \bar{q}))$.

For case (ii), we again use the fact that $H_{g1}(\tau_1, \tau_2)$ depends on τ_2 only through the residue of τ_2 mod p, and establish case (ii) only for $k = 0$. Let τ_2 be in the range $\bar{q} \le \tau_2 < p$. Consider the solutions to (35) separately in $0 \le x < \tau_2$ and $\tau_2 \le x < q$. For $0 \le x < \tau_2$, $\mathbf{I}(0 \le x < \tau_2)$ is identically equal to 1. By Lemma 1, there are at most 2 solutions to (35) for $x = 0, 1, \ldots, \tau_2 - 1$. For $\tau_2 \le x < q$, $\mathbf{I}(0 \le x < \tau_2)$ is identically equal to 0, and by Lemma 1, there are at most $m - 1$ solutions to (35) for $\tau_2 \le x < q$. There are totally at most $m + 1$ solutions to (35).

This completes the proof of Theorem 4.

Appendix C
Proof of Theorem 9

In this proof, p is a prime number and q is an integer relatively prime to p and strictly larger than $2p^2$, $L = pq$ is the sequence period, and γ is the multiplicative inverse of $p \bmod q$, i.e., $\gamma p \equiv 1 \bmod q$. The unique integer between 0 and $q - 1$ which is equal to x mod q is denoted by$(x \bmod q)$. The translate of a subset $\mathcal{S}$ in $G_{p,q}$ by (τ_1, τ_2) is defined as

$$\mathcal{S} + (\tau_1, \tau_2) := \{(x, y) + (\tau_1, \tau_2) : (x, y) \in \mathcal{S}\}.$$

Consider user g, where $i = 1, 2, \ldots, p - 1$. If user g starts transmitting at time t_0, then the channel-activity signal is matched to $s_g(t)$ at time t_0. The receiver will never fail to detect the presence of user g, meaning that if user g does start transmitting, the receiver can always detect this change of status from idle to active. The only sources of error are (a) detecting a user but in fact that user is not active, and (b) miscalculation of the start time. We refer to the error in (a) as *false alarm* and (b) as *synchronization error*.

We now show that false alarm cannot occur. Suppose on the contrary that the channel-activity signal is matched to $s_g(t)$ at time t_0, but user g is idle from time t_0 to $t_0 + L - 1$. If this happened, the q time indices in $\mathcal{I}_g + t_0$ would be covered by the protocol sequences of other active users. However, the cross-correlation between user g and each other active user is upper bounded by $\lfloor q/p \rfloor + 2$, by Theorem 3. Because user g is assumed to be inactive in this period, the number of simultaneously active users does not exceed the maximum $(p+1)/2$, and hence the number of time slots in $\mathcal{I}_g + t_0$ with

0	**3**	**6**	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
19	22	25	28	31	34	37	40	43	46	49	52	55	**1**	**4**	**7**	10	13	16
38	41	44	47	50	53	56	**2**	**5**	**8**	11	14	17	20	23	26	29	32	35

Fig. 4. Mapping from $\mathbb{Z}_{57}$ to a 3×19 array $\mathbf{M}_{3,19}$. The numbers 0 to 8 are highlighted

a packet or collision observed is no larger than

$$\begin{aligned}\big(\lfloor q/p \rfloor + 2\big)\Big(\frac{p+1}{2}\Big) &< \Big(\frac{q}{p} + 2\Big)\Big(\frac{p+1}{2}\Big) \\ &= q\Big(\frac{1}{p} + \frac{2}{q}\Big)\frac{p+1}{2} \\ &< q\Big(\frac{1}{p} + \frac{2}{2p^2}\Big)\frac{p+1}{2} \qquad (42) \\ &= \frac{q}{2}\Big(\frac{p+1}{p}\Big)^2\end{aligned}$$

In (42), we have used the assumption that $q > 2p^2$. We note that for all $p \geq 3$, the factor $(p+1)^2/p^2$ is strictly less than two. We obtain

$$\big(\lfloor q/p \rfloor + 2\big)\Big(\frac{p+1}{2}\Big) < q.$$

But q is precisely the total number of ones in a period of $s_g(t)$. The q time slots indices by $\mathcal{I}_g + t + 0$ cannot be covered by any other $(p+1)/2$ CRT protocol sequences. The channel-activity signal $c(t)$ cannot be matched to $s_g(t)$, and therefore false alarm cannot occur.

For synchronization error, assume that user g is idle from time $t_0 - L + 1$ to $t_0 - 1$, and becomes active at time t_0. Our objective is to show that the channel-activity signal is not matched to $s_g(t)$ at time $t_0 - \tau$, for any integer τ between 1 and $L-1$. The idea of showing that synchronization error cannot occur is the following. If the channel-activity signal were matched incorrectly to $s_g(t)$ at $t_0 - \tau$, then the receiver would observe q "1" or "*" at time slots indexed by $t_0 - \tau + \Phi'^{-1}_{p,q}(\mathcal{I}_{g,p,q})$. Among these q time slots, say b of them come from a shifted version of $s_g(t)$, starting at time t_0. We then show that the remaining $q - b$ slots cannot be covered by the other active users. We divide the proof into several propositions below.

In the modified CRT correspondence Φ', an element $t \in \mathbb{Z}_{pq}$ is mapped to $(t \bmod p, \gamma t \bmod q)$. In order to visualize the mapping, we introduce a matrix $\mathbf{M}_{p,q}$.

Definition 8. Given relatively prime integers p and q, let $\mathbf{M}_{p,q}$ be a $p \times q$ matrix whose (i,j)-entry equals t if $i \equiv t \bmod p$ and $j \equiv \gamma t \bmod q$, for $t = 0, 1, \ldots, pq-1$. The rows and columns of $\mathbf{M}_{p,q}$ are indexed by $\{0, 1, \ldots, p-1\}$ and $\{0, 1, \ldots, q-1\}$, respectively.

Each integer from 0 to $pq-1$ appears exactly once in $\mathbf{M}_{p,q}$. An example for $p = 3$ and $q = 19$ is shown in Fig. 4. We pay special attention to the integers from 0 to $p^2 - 1$, and want to get a handle on where they are located in $\mathbf{M}_{p,q}$. In Fig. 4, we can see that 0, 3 and 6 are on the upper left corner of $\mathbf{M}_{3,19}$. The numbers 1, 4 and 7 occupy three consecutive entries in row 1. The numbers 2, 5 and 8 occupy three consecutive entries in row 2.

Proposition 1. *(i) For any i and j, the $(i, j+1)$-entry is equal to p plus the (i,j)-entry mod pq.*

(ii) Under the modified CRT correspondence $\Phi'_{p,q}$, the integers kp, for $k = 0, 1, 2, \ldots, q-1$, are mapped to $(0, k)$. They appear in the first row of $\mathbf{M}_{p,q}$.

(iii) The numbers from 1 to $p^2 - 1$, except the multiples of p, are located between column $2p+1$ and column $q-p-1$ inclusively in $\mathbf{M}_{p,q}$.

Proof: (i) The $(0,1)$-entry in $\mathbf{M}_{p,q}$ is labeled by p, because

$$p \mapsto (p \bmod p, \gamma p \bmod q) = (0, 1),$$

and $\gamma p \equiv 1 \bmod q$ by the defining property of γ.

(ii) For $k = 0, 1, 2, \ldots, q-1$,

$$\Phi'_{p,q}(kp) = (kp, kp\gamma) = (0, k(1)) = (0, k).$$

(iii) We have the following claim:

$$2p < (k\gamma \bmod q) \leq q - 2p \qquad (43)$$

for $k = 1, 2, \ldots, p-1$.

We prove the claim by contradiction. Suppose that $k\gamma$, after reduction mod q, is between 1 and $2p$. Then, $k\gamma p \bmod q$ is equal to p, $2p$, $3p, \ldots,$ or $2p^2$. Since $q > 2p^2$, these p numbers remain unchanged after reduction mod q. However, $k\gamma p \equiv k(\gamma p) \equiv k \bmod q$, and this contradicts the assumption that k is between 1 and $p-1$. Now suppose that $(k\gamma \bmod q)$ is equal to $q - 2p + 1$, $q - 2p + 2, \ldots,$ or $q - 1$. Then, the value of $k\gamma p$, after reduction mod q, is equal to

$$q - 2p^2 + p,\ q - 2p^2 + 2p, \ldots,\ \text{or } q - p.$$

Since $k\gamma p \equiv k \bmod q$, this also contradicts that k is between 1 and $p-1$. This finishes the proof of the claim.

Let ℓ be an integer between 1 and $p^2 - 1$ which is not a multiple of p. We can write ℓ as $mp + k$ for some m and k between 1 and $p-1$. By part (i), the location of ℓ in $\mathbf{M}_{p,q}$ is m steps to the right of the location of k in $\mathbf{M}_{p,q}$. But k cannot be located to the right of column $q-2p$. The right-most column in $\mathbf{M}_{p,q}$ which may contain ℓ is thus $q-p-1$. This finishes the proof of part (ii). ∎

Proposition 2. *(i) Let $\mathcal{S} = \{0, 1, 2, \ldots, p^2 - 1\}$, and $\mathcal{S}'$ be the image of $\mathcal{S}$ under $\Phi'_{p,q}$, i.e., $\mathcal{S}' = \Phi'_{p,q}(\mathcal{S})$. We have*

$$|\mathcal{I}_g \cap (\mathcal{I}_h + (\tau_1, \tau_2)) \cap \mathcal{S}'| \leq 2 \qquad (44)$$

for any given (τ_1, τ_2) and $g \neq h$.

(ii) For $g = 1, 2, \ldots, p-1$, there are exactly p ones in the first p^2 bits of $s_g(t)$, i.e., there are exactly p ones among $s_g(0)$, $s_g(1), \ldots, s_g(p^2-1)$.

The quantity on the left hand side of (44) can be interpreted as the *partial Hamming cross-correlation*, defined as

$$\sum_{t=0}^{p^2-1} s_g(t)s_h(t+\tau).$$

We only consider the number of overlaps in the first p^2 time indices. The proposition asserts that the partial Hamming cross-correlation of two CRT sequences cannot exceed two.

Proof: (i) By part (i) of Prop. 1, for each k between 0 and $p-1$, the following p integers,

$$k, k+p, k+2p, \ldots, k+(p-1)p,$$

occupy p consecutive horizontal entries in the kth row of $\mathbf{M}_{p,q}$. Wrapping around the right boundary of $\mathbf{M}_{p,q}$ is precluded by Prop. 1.

A common element of $\mathcal{I}_g$ and $\mathcal{I}_h + (\tau_1, \tau_2)$ is in the form

$$(gt_1, t_1) = (ht_2 + \tau_1, t_2 + \tau_2) \tag{45}$$

for some t_1 and t_2 between 0 and $q-1$. If $t_2+\tau_2$ is between 0 and $q-1$, then the first coordinates of the two order pairs in (45) are equal to $ht_2+\tau_2$, with t_2 satisfying

$$g(t_2+\tau_2) \equiv ht_2 + \tau_1 \bmod p. \tag{46}$$

If $t_2+\tau_2$ is larger than q, then $t_1 = t_2+\tau_2-q$, and the first coordinates of the two order pairs in (45) are equal to $h_2+\tau_2$, with t_2 satisfying

$$g(t_2+\tau_2-q) \equiv ht_2 + \tau_1 \bmod p. \tag{47}$$

Hence, $ht_2+\tau_1$ may assume only two values mod p, one from (46) and the other one from (47). Let $\mathbf{A}_g$ be the $p \times q$ array with characteristic set $\mathcal{I}_g$. We see that the elements in $\mathcal{I}_g \cap (\mathcal{I}_h + (\tau_1, \tau_2))$ are located in at most two rows in $\mathbf{A}_g$. Since p consecutive entries in a row of $\mathbf{A}_g$ contain exactly one "1", at most two elements in $\mathcal{I}_g \cap (\mathcal{I}_h + (\tau_1, \tau_2))$ are covered by $\mathcal{S}'$.

(ii) Let $\mathbf{A}_g$ be the $p \times q$ array with $\mathcal{I}_g$ as the characteristic set. From the remark before Definition 4, any p consecutive columns in $\mathbf{A}_g$ form a permutation matrix. For each $b = 0, 1, \ldots, p-1$, the time indices

$$b,\ b+p, \ldots, b+(p-1)p$$

are p consecutive entries in a row in $\mathbf{M}_{p,q}$. Hence there is exactly one "1" among $s_g(b)$, $s_g(b+p), \ldots, s_g(b+(p-1)p)$. Since this is true for $b = 0, 1, \ldots, p-1$, we conclude that there are exactly p "1" in $s_g(0)$, $s_g(1), \ldots, s_g(p^2-1)$. ■

Proposition 3. *For $y = 0, 1, \ldots, q-p$, let $\mathcal{P}'_y$ be the index set*

$$\mathcal{P}'_y := \{(i,j) \in G_{p,q} : 0 \le i \le p-1, y \le j \le y+p-1\}.$$

Let $\mathcal{P}_y$ be the corresponding set of time indices in $\mathbb{Z}_{pq}$ under the modified CRT correspondence,

$$\mathcal{P}_y := \Phi'^{-1}_{p,q}(\mathcal{P}').$$

Then

(i) For $g \neq h$, and delay offset (τ_1, τ_2),

$$|\mathcal{I}_g \cap (\mathcal{I}_h + (\tau_1, \tau_2)) \cap \mathcal{P}'_y| \le 2.$$

(ii) For $g = 1, 2, \ldots, p$, there are exactly p ones in $s_g(t)$ for $t \in \mathcal{P}_y$.

Proof: The elements in $\mathcal{P}'_y$ can be regarded as a square submatrix in a $p \times q$ matrix. The proof of part (i) of Prop. 3 is similar to that of Prop. 2 and is omitted. The second part follows from the fact that every p consecutive columns in $\mathbf{A}_g$, which stands for the $p \times q$ matrix with characteristic set $\mathcal{I}_g$, form a permutation matrix, and hence contains exactly p ones. ■

Proposition 4. *Let $\mathcal{S}'$, $\mathcal{P}'_y$ be defined as in Prop. 2 and 3. For $g = 1, 2, \ldots, p-1$, and τ between 1 and $L-1$, the subset $\mathcal{B}_{g,\tau}$ of time indices t in $\{0, 1, 2, \ldots, L-1\}$, which satisfy*

$$s_g(t) = 1 \text{ for } t = 0, 1, \ldots, \tau-1,$$
$$s_g(t) = 1 \text{ and } s_g(t-\tau) = 0 \text{ for } t = \tau, \tau+1, \ldots, L-1,$$

contains $\Phi'^{-1}_{p,q}(\mathcal{S}' \cap \mathcal{I}_g)$, or $\Phi'^{-1}_{p,q}(\mathcal{P}'_y \cap \mathcal{I}_g)$ for some y.

Proof: Let $\hat{s}_g(t)$ be the acyclic shift of $s_g(t)$ to the right by delay offset τ. The time indices in $\mathcal{B}_{g,\tau}$ correspond to the "1" in $s_g(t)$ which is not covered by $\hat{s}_g(t)$. We consider two cases: (i) $p^2 \le \tau < L$, and (ii) $1 \le \tau < p^2$.

Case (i). When $p^2 \le \tau < L$, the first p^2 bits in $s_g(t)$ are not covered by $\hat{s}_g(t)$, and therefore $\mathcal{B}_{g,\tau}$ contains $\Phi'^{-1}_{p,q}(\mathcal{S}' \cap \mathcal{I}_g)$.

Case (ii). Recall that in the proof of Theorem 6, it is mentioned that the intersection of $\mathcal{I}_g$ and $\mathcal{I}_g + (\tau_1, \tau_2)$ is either empty, or an arithmetic progressions in $G_{p,q}$ with common difference $(g, 1)$. The intersection is non-empty if and only if (τ_1, τ_2) equals $(g, 1)k$ for some $k = \pm 1, \pm 2, \ldots, \pm(q-1)$.

Suppose that τ is a nonzero multiple of p between 1 and p^2-1. By part (i) of Prop. 1, $\Phi'_{p,q}(\tau) = (0, \tau/p)$. ($\tau$ is one of the p left-most entries in the first row of $\mathcal{M}_{p,q}$.) In this case, $\Phi'(\tau)$ does not equal $(g, 1)k$ for any $k \in \{1, 2, \ldots, p-1\}$. (We have used the assumption that g is non-zero in this step.) This implies that the ones in $\mathbf{A}_g$ with indices in $\mathcal{P}'_{\tau/p}$ are not covered by $\hat{s}_g(t)$. Thus,

$$\mathcal{B}_{g,\tau} \supseteq \Phi'^{-1}_{p,q}(\mathcal{P}'_{\tau/p} \cap \mathcal{I}_g).$$

Now suppose that τ is between 1 and p^2-1 but is not a multiple of p. Let $(\tau_1, \tau_2) = \Phi'_{p,q}(\tau)$. Using similar argument as in the previous paragraph, if $(\tau_1, \tau_2) \notin \mathcal{I}_g$, then

$$\mathcal{B}_{g,\tau} \supseteq \Phi'^{-1}_{p,q}(\mathcal{P}'_{\tau_2} \cap \mathcal{I}_g).$$

Otherwise if $(\tau_1, \tau_2) \notin \mathcal{I}_g$, then the intersection of $\mathcal{I}_g$ and $\mathcal{I}_g + (\tau_1, \tau_2)$ equals

$$\{(g, 1)k :\ k = \tau_2, \tau_2+1, \ldots, q-1\}.$$

Consider the p columns in $\mathbf{M}_{p,q}$ to the left of τ, namely the time indices associated with $\mathcal{P}'_{\tau_2-p}$ in $\mathbf{M}_{p,q}$. The corresponding time slots are not covered by $\hat{s}_g(t)$. Therefore,

$$\mathcal{B}_{g,\tau} \supseteq \Phi'^{-1}_{p,q}(\mathcal{P}'_{\tau_2-p} \cap \mathcal{I}_g).$$

■

Suppose that $\hat{s}_g(t)$ is actually transmitted at time t_0 and the receiver tries to match the channel-activity signal $c(t)$ with $s_g(t)$ at time $t_0 - \tau$ (See Fig. 5). If the channel-activity signal were mistakenly matched to $s_g(t)$ at $t_0 - \tau$, then by the

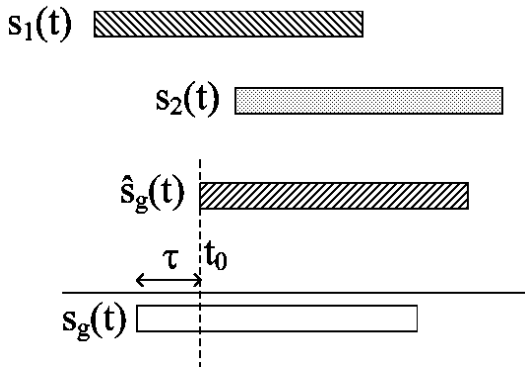

Fig. 5. The protocol sequence $\hat{s}_g(t)$ is transmitted at time t_0. Sequences $s_1(t)$ and $s_2(t)$ are transmitted at some other time. The receiver tries to match $c(t)$ with $s_g(t)$ at $t_0 - \tau$.

previous proposition, either (i) the p "1" in the first p^2 bits of $s_g(t)$, which occur at time indices

$$(\Phi'^{-1}_{p,q}(\mathcal{I}_g) \cap \{0, 1, \ldots, p^2 - 1\}) + t - \tau,$$

or (ii) the p "1" in $s_g(t)$, which occur at time indices in

$$\Phi'^{-1}_{p,q}(\mathcal{I}_g \cap \mathcal{P}'_y) + t - \tau,$$

for some y between 0 and $q - p$, are covered by the other active users. By Prop. 2 and (3), each of the other active users can contribute at most two overlapping slots. As there are no more than $(p-1)/2$ other active users, the total number of ones that can be covered by other active users are at most $2 \cdot (p-1)/2$, which is strictly smaller than p. This proves that the channel-activity signal cannot be matched to $s_g(t)$ at $t_0 - \tau$ for any $\tau = 1, 2, \ldots, L-1$. This completes the proof of Theorem 9.

REFERENCES

[1] J. L. Massey and P. Mathys, "The collision channel without feedback," *IEEE Trans. Inform. Theory*, vol. 31, no. 2, pp. 192–204, Mar. 1985.

[2] K. W. Shum, C. S. Chen, C. W. Sung, and W. S. Wong, "Shift-invariant protocol sequences for the collision channel without feedback," *IEEE Trans. Inform. Theory*, vol. 55, pp. 3312–3322, Jul. 2009.

[3] V. C. da Rocha, Jr., "Protocol sequences for collision channel without feedback," *IEE Electron. Lett.*, vol. 36, no. 24, pp. 2010–2012, Nov. 2000.

[4] Y. Zhang, K. W. Shum, and W. S. Wong, "On pairwise shift-invariant protocol sequences," *IEEE Commun. Lett.*, vol. 13, no. 6, pp. 453–455, 2009.

[5] N. Q. A, L. Györfi, and J. L. Massey, "Constructions of binary constant-weight cyclic codes and cyclically permutable codes," *IEEE Trans. Inform. Theory*, vol. 38, no. 3, pp. 940–949, May 1992.

[6] L. Gyöfi and I. Vajda, "Construction of protocol sequences for multiple-access collision channel without feedback," *IEEE Trans. Inform. Theory*, vol. 39, no. 5, pp. 1762–1765, Sep. 1993.

[7] O. Moreno, Z. Zhang, P. V. Kumar, and V. A. Zinoviev, "New constructions of optimal cyclically permutable constant weight codes," *IEEE Trans. Inform. Theory*, vol. 41, no. 2, pp. 448–455, Mar. 1995.

[8] S. Bitan and T. Etzion, "Constructions for optimal constant weight cyclically permutable codes and difference families," *IEEE Trans. Inform. Theory*, vol. 41, no. 1, pp. 77–87, Jan. 1995.

[9] F. R. K. Chung, J. A. Salehi, and V. K. Wei, "Optical orthogonal codes: design, analysis and applications," *IEEE Trans. Inform. Theory*, vol. 35, no. 3, pp. 595–604, May 1989.

[10] W. S. Wong, "New protocol sequences for random access channels without feedback," *IEEE Trans. Inform. Theory*, vol. 53, no. 6, pp. 2060–2071, Jun. 2007.

[11] L. Györfi and S. Győri, *Multiple Access Channel – Theory and practice*. Amsterdam: IOS press, 2007, ch. Coding for multiple-acccess channel without feedback, pp. 299–326.

[12] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*. New York: Springer-Verlag, 1990.

[13] A. A. Shaar and P. A. Davies, "Prime sequences: quasi-optimal sequences for OR channel code division multiplexing," *IEE Electron. Lett.*, vol. 19, no. 21, pp. 888–890, Oct. 1983.

[14] G.-C. Yang and W. C. Kwong, "Performance analysis of optical CDMA with prime codes," *IEE Electron. Lett.*, vol. 31, no. 7, pp. 569–570, Mar. 1995.

[15] D. V. Sarwate and M. B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proc. IEEE*, vol. 68, no. 5, pp. 593–619, 1980.