Geometric Orthogonal Codes Better than Optical Orthogonal Codes

Yeow Meng Chee, Han Mao Kiah, San Ling, and Hengjia Wei
School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore
Emails: {ymchee, hmkiah, lingsan, hjwei}@ntu.edu.sg

***Abstract*—The class of geometric orthogonal codes (GOCs) were introduced by Doty and Winslow (2016) for more robust macrobonding in DNA origami. They observed that GOCs are closely related to optical orthogonal codes (OOCs). It is possible for GOCs to have size greater than OOCs of corresponding parameters due to slightly more relaxed constraints on correlations. However, the existence of GOCs exceeding the size of optimal OOCs of corresponding parameters have never been demonstrated. This paper gives the first infinite family of GOCs of size greater than optimal OOCs.**

I. Introduction

Nucleic acids play an important role in the self assembly of nanostructures owing the specificity of the Watson-Crick base pairing. Rothemund [1] showed how a long strand of (scaffold) DNA can be folded into a specific shape (DNA origami) with the help of a carefully designed set of short "staple" DNAs that bind to intended sites on the scaffold DNA, forcing the scaffold DNA to fold in desired ways. Beyond base pairing, *base stacking* between base pairs is another dominant cause of DNA binding. Woo and Rothemund [2] showed that by careful placement of blunt ends in the DNA origami of Rothemund, we can force a set of DNA origamis to bind through base stacking to form intended arrangements. This geometric placement of blunt ends within a DNA origami forms a *macrobond*. Extending the work of Woo and Rothemund [2], Doty and Winslow [3] introduce the class of *geometric orthogonal codes* (GOCs) for the design of more robust macrobonds that weed out undesirable bonding arising from flexibility of DNA and misalignment. GOCs of large size are desirable because they give rise to large number of binding interactions, thereby increasing the number of structures that can potentially be formed.

Doty and Winslow [3] observed that GOCs are closely related to *optical orthogonal codes* (OOCs) introduced by Chung et al. [4]. Although it is possible for GOCs to have size larger than OOCs of corresponding parameters, this has never been demonstrated. The main contribution of this paper is the first construction of GOCs that is better than optimal OOCs. We also improve an upper bound of Doty and WInslow [3] on the size of GOCs.

II. Preliminaries

For an integer $n \geq 2$, let I_n denote the set of integers $\{0, 1, \ldots, n-1\}$. Given $M \subseteq I_n^2$ and $\mathbf{v} \in \mathbb{Z}^2$, the *translation of M by* $\mathbf{v}$ is defined to be $M+\mathbf{v} = \{\mathbf{m}+\mathbf{v} : \mathbf{m} \in M\}$. The *auto-correlation* of M is defined as $\max_{\mathbf{v}\in\mathbb{Z}^2\setminus\{(0,0)\}} |M \cap (M+\mathbf{v})|$. For two subsets $M, M' \subseteq I_n^2$, the *cross-correlation* of M and M' is defined as $\max_{\mathbf{v}\in\mathbb{Z}^2} |M \cap (M'+\mathbf{v})|$.

Let $w \in \{2, 3, \ldots, n^2\}$ and let $\lambda \in \{1, 2 \ldots, w-1\}$. A family $\mathcal{M} = \{M_1, M_2, \ldots, M_m\}$ of size-w subsets (or w-subsets for short) of I_n^2 is an (n, w, λ)-*geometric orthogonal code* (GOC) if

(i) the auto-correlation of M is $\leq \lambda$, for all $M \in \mathcal{M}$, and
(ii) the cross-correlation of M and M' is $\leq \lambda$, for all $M, M' \in \mathcal{M}$ with $M \neq M'$.

The parameter w is called the *macrobond strength* (or *weight*) of $\mathcal{M}$, while λ is its *mismatch strength limit*. We note further that every $M \subseteq I_n^2$ may be identified with an $n \times n$ $(0,1)$-matrix $(m_{i,j})_{0\leq i,j\leq n-1}$, where $m_{i,j} = 1$ if $(i,j) \in M$ and $m_{i,j} = 0$ otherwise.

Let $M(n, w, \lambda)$ denote the largest possible size of an (n, w, λ)-GOC. A code with the largest size is said to be *optimal*. Doty and Winslow derived the following upper bound for $M(n, w, \lambda)$.

Theorem 1 ([3])**.** *Let*

$$\begin{aligned} &U_{\mathrm{GOC}}(n, w, \lambda) \\ &\triangleq \frac{1}{\binom{w}{\lambda+1}} \left[\binom{n^2-1}{\lambda} + \sum_{x_0=1}^{n-1} \sum_{y_0=1}^{n-1} \binom{n^2-x_0-y_0-1}{\lambda-1} \right] \\ &= (1+o(1)) \frac{(\lambda+1)^2 n^{2\lambda}}{w(w-1)(w-2)\cdots(w-\lambda)}. \end{aligned} \tag{1}$$

Then $M(n, w, \lambda) \leq U_{\mathrm{GOC}}(n, w, \lambda)$.

Let N be a positive integer, let $1 \leq \lambda \leq w \leq N$ and let $\mathbb{Z}_N$ denote the integers modulo N. A collection $\mathcal{C} = \{C_1, C_2, \ldots, C_m\}$ of w-subsets of $\mathbb{Z}_N$ is an (N, w, λ)-*optical orthogonal code* (OOC) if the following conditions are satisfied:

(i) $\max_{v\in\mathbb{Z}_N\setminus\{0\}} |C \cap (C+v)| \leq \lambda$, for all $C \in \mathcal{C}$, and
(ii) $\max_{v\in\mathbb{Z}_N} |C \cap (C'+v)| \leq \lambda$, for all $C, C' \in \mathcal{C}$ with $C \neq C'$.

Note that, for translations in the definition of OOCs, addition is performed over the *cyclic group* $\mathbb{Z}_N$, instead of over the integers as in the definition of GOCs.

Chung et al. [4] showed that the size of an (N, w, λ)-OOC is bounded above by $U_{\mathrm{OOC}}(N, w, \lambda)$, where

$$U_{\mathrm{OOC}}(N, w, \lambda) \triangleq \frac{(N-1)(N-2)\cdots(N-\lambda)}{w(w-1)(w-2)\cdots(w-\lambda)}. \tag{2}$$

978-1-5090-4096-4/17/$31.00 ©2017 IEEE

Observe that an (n^2, w, λ)-OOC is an (n, w, λ)-GOC, by regarding each one-dimensional (1D) codeword of length n^2 as the concatenation of the n rows of a two-dimensional (2D) codeword. Comparing (1) and (2), with $N = n^2$, we note that the size of an (n, w, λ)-GOC could possibly exceed the upper bound $U_{\mathrm{OOC}}(n^2, w, \lambda)$. However, no such classes of GOCs are known. While Doty and Winslow [3] constructed a class of (p, p, λ)-GOCs of size $p^{\lambda-1} - p^{\lambda-2}$, for all primes p, and have compared this code size with some known lower bounds for OOCs, this code size does not beat the bound $U_{\mathrm{OOC}}(p^2, p, \lambda) = p^{\lambda-1} + O(p^{\lambda-2})$.

A. Our Contributions

The main contributions of this paper are:

- For suitably large w, an upper bound for $M(n, w, \lambda)$, which is asymptotically equal to (2), with $N = n^2$ (Section III).
- For $t \leq p$ and $p - \lfloor p/t \rfloor \leq \lambda \leq p$, a class of (p, p, λ)-GOCs of size $tp^{\lambda-1} - t$, which exceeds the OOC upper bound $p^{\lambda-1} + O(p^{\lambda-2})$ (Section V). The construction makes use of a class of GOCs, called cyclic orthogonal geometric codes, discussed in Section IV.
- A recursive construction for GOCs, which can increase n while keeping w and λ fixed (Section VI). We also show that, if the input codes are close to optimal, so are the output codes. Examples of GOCs with size exceeding (2) are obtained.
- We construct optimal GOCs for $\lambda = 1$ (Section VII).

The techniques used are from combinatorial design theory. Due to space constraints, many proofs have been omitted.

III. An Asymptotic Upper Bound

In this section, we use a method of Erdős et al. [5] to obtain an asymptotic upper bound on the size of (n, w, λ)-GOC when w is large. For $\mathbf{a} = (a_1, a_2) \in \mathbb{Z}^2$ and positive integer R, let $W_{\mathbf{a},R}$ be an $R \times R$ *window starting at* $\mathbf{a}$, that is, $W_{\mathbf{a},R} = \{a_1, a_1 + 1, \ldots, a_1 + R - 1\} \times \{a_2, a_2 + 1, \ldots, a_2 + R - 1\}$. For $S \subseteq I_n^2$, the *observation of S through the window $W_{\mathbf{a},R}$* is

$$W_{\mathbf{a},R}(S) = \{\mathbf{v} - \mathbf{a} : \mathbf{v} \in S \cap W_{\mathbf{a},R}\}.$$

Note that every observation, by definition, lies within I_R^2.

Theorem 2. *Let w and λ be functions in n. If $w = \Omega(\lambda^4 n^c)$ for some positive constant c, then*

$$M(n, w, \lambda) \leq (1 + o(1)) \frac{n^{2\lambda}}{w^{\lambda+1}}.$$

Therefore, $\lim_{n\to\infty} M(n, w(n), \lambda(n))/U_{\mathrm{OOC}}(n^2, w(n), \lambda(n)) \leq 1$.

Proof. Let $\mathcal{M} = \{M_1, M_2, \ldots, M_m\}$ be an (n, w, λ)-GOC. The number of $R \times R$ windows with nonempty intersection with I_n^2 is $(n + R - 1)^2$, so the number of observations of the elements of $\mathcal{M}$ through these windows is $N = m(n + R - 1)^2$. As each element of I_n^2 is observed through R^2 windows, the average number of elements per observation, over these N observations, is $A = R^2 mw/N$.

On the other hand, suppose the ith observation is of size w_i. Then it has precisely $\binom{w_i}{\lambda+1}$ subsets of size $\lambda+1$. Therefore, there are in total $\sum_{i=1}^{N} \binom{w_i}{\lambda+1}$ subsets of size exactly $\lambda + 1$, induced by these N observations.

Now, since $\mathcal{M}$ is an (n, w, λ)-GOC, we have $|W_{\mathbf{a},R}(M_i) \cap W_{\mathbf{b},R}(M_j)| \leq \lambda$ for any two distinct observations $W_{\mathbf{a},R}(M_i)$ and $W_{\mathbf{b},R}(M_j)$, with $\mathbf{a} \neq \mathbf{b}$ or $i \neq j$. Therefore, all the $(\lambda+1)$-subsets, induced from the observations, are pairwise distinct. The number of possible $(\lambda+1)$-subsets in an $R \times R$ window is $\binom{R^2}{\lambda+1}$. Since all observations lie within the $R \times R$ window I_R^2, we have

$$\sum_{i=1}^{N} \binom{w_i}{\lambda+1} \leq \binom{R^2}{\lambda+1}.$$

Note that $A = \frac{\sum_{i=1}^{N} w_i}{N}$. It follows from the convexity of the function $\binom{x}{\lambda+1}$ in variable x and Jensen's inequality that

$$\binom{A}{\lambda+1} \leq \frac{1}{N} \sum_{i=1}^{N} \binom{w_i}{\lambda+1} \leq \frac{1}{N} \binom{R^2}{\lambda+1}.$$

In other words, $NA(A-1)\cdots(A-\lambda) \leq R^{2\lambda+2}$, or, $mw(A-1)\cdots(A-\lambda) \leq R^{2\lambda}$.

Choose $R = n^{1-c/4}$. Since $w = \Omega(\lambda^4 n^c)$, we have $A = R^2 w/(n + R - 1)^2 = \Omega(\lambda^4 n^{c/2})$. It follows that for n large enough, we have

$$(A-1)(A-2)\cdots(A-\lambda) \geq A^\lambda - \lambda^2 A^{\lambda-1}.$$

Hence, $mw(A^\lambda - \lambda^2 A^{\lambda-1}) \leq R^{2\lambda}$, and

$$\begin{aligned} m &\leq \frac{R^{2\lambda}}{wA^\lambda} + \frac{m\lambda^2}{A} = \frac{(n+R-1)^{2\lambda}}{w^{\lambda+1}} + \frac{m\lambda^2}{A}, \\ &= \frac{n^{2\lambda}}{w^{\lambda+1}} + o\left(\frac{n^{2\lambda}}{w^{\lambda+1}}\right) + \frac{m\lambda^2}{A} \\ &= \frac{n^{2\lambda}}{w^{\lambda+1}} + o\left(\frac{n^{2\lambda}}{w^{\lambda+1}}\right). \end{aligned}$$

The last equation holds as $m = O(\lambda^2 n^{2\lambda}/w^{\lambda+1})$ and $m\lambda^2/A = o(n^{2\lambda}/w^{\lambda+1})$. □

It follows that both OOCs and GOCs share the same asymptotic upper bound when $w = \Omega(\lambda^4 n^c)$, for any constant $c > 0$.

IV. Cyclic Geometric Orthogonal Codes

Let $\mathcal{M}$ be a family of w-subsets of $\mathbb{Z}_m \times \mathbb{Z}_n$. For $M \in \mathcal{M}$, the *auto-correlation* of M is defined as $\max_{\mathbf{v} \in (\mathbb{Z}_m \times \mathbb{Z}_n) \setminus \{(0,0)\}} |M \cap (M + \mathbf{v})|$. For distinct $M, M' \in \mathcal{M}$, the *cross-correlation* of M and M' is defined as $\max_{\mathbf{v} \in \mathbb{Z}_m \times \mathbb{Z}_n} |M \cap (M' + \mathbf{v})|$. (Note that the translations are computed over the group $\mathbb{Z}_m \times \mathbb{Z}_n$.) Such a family $\mathcal{M}$ is called an (m, n, w, λ)-*optical orthogonal signature pattern code* (OOSPC) if, for $M, M' \in \mathcal{M}$ with $M \neq M'$, we have:

(i) the auto-correlation of M is $\leq \lambda$, and
(ii) the cross-correlation of M and M' is $\leq \lambda$.

Optical orthogonal signature pattern codes were studied in the context of optical CDMA networks [6], [7], [8].

We restrict ourselves to the case where $m = n$. Each codeword may then be visualized as a square array. When a translation is applied to such a codeword, entries in the codeword that move off one edge of the array reappear in the array from the opposite edge (due to the modulo n operation), unlike in the case of a GOC, where the symbols simply move out of the array. Therefore, identifying $\mathbb{Z}_n^2$ with I_n^2 as sets in the obvious way, it is easy to see the auto-correlation (resp. cross-correlation) in the OOSPC definition is always no less than the auto-correlation (resp. cross-correlation) in the GOC definition. It follows that an (n, n, w, λ)-OOSPC is also an (n, w, λ)-GOC. For these reasons, we shall refer to an (n, n, w, λ)-OOSPC as an (n, w, λ)-*cyclic geometric orthogonal code* (CGOC). These codes are used in Section V to construct GOCs whose size exceeds (2), with $N = n^2$.

Although an (n^2, w, λ)-OOC is an (n, w, λ)-GOC and the translations in both OOCs and CGOCs are done modulo n, there are differences in their properties. Consider, for example, the codeword (011100000). As a 1D codeword of length nine, it has auto-correlation two. However, when interpreted as the corresponding 2D codeword $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, its auto-correlation, in both the GOC and OOSPC definitions, is one.

Let $C(n, w, \lambda)$ denote the largest possible size of an (n, w, λ)-CGOC. Since CGOCs may be regarded as binary constant weight codes (by identifying subsets of $\mathbb{Z}_n^2$ with $n \times n$ $(0, 1)$-matrices), by using the Johnson bound [9] for constant weight codes, we have the following upper bound on $C(n, w, \lambda)$.

Theorem 3 (Johnson-type bound)**.** *Let*

$$U_{\mathrm{CGOC}}(n, w, \lambda) \triangleq \frac{(n^2-1)(n^2-2)\cdots(n^2-\lambda)}{w(w-1)(w-2)\cdots(w-\lambda)}.$$

Then $C(n, w, \lambda) \le U_{\mathrm{CGOC}}(n, w, \lambda)$.

Although $U_{\mathrm{CGOC}}(n, w, \lambda) = U_{\mathrm{OOC}}(n^2, w, \lambda)$, CGOCs have the potential to yield GOCs whose size exceeds (2), since the correlation in the CGOC definition may be larger than that in GOC definition and there is room to add more codewords.

A. Constructions for CGOCs

Some direct constructions of CGOCs have recently been given by Ji et al.[10].

Theorem 4 ([10])**.** *Let $p \ge 3$ be a prime and λ an integer with $2 \le \lambda \le p$. Then there is a (p, p, λ)-CGOC of size $p^{\lambda-1} - 1$.*

Let G be an abelian group and let r be a positive integer. An $s \times t$ matrix $A = (a_{ij})$ over G is called r-*simple*, if the list of differences of any two column vectors of A contains any element of G at most $r - 1$ times. Chu and Golomb [11] gave a powerful recursive construction for OOCs, based on r-simple matrices over $\mathbb{Z}_v$. Ji et al. [10] modified their approach and gave a recursive construction for CGOCs.

Theorem 5 ([10])**.** *Suppose there exist an (n, w, λ)-CGOC $\mathcal{M}$ and a $w \times N$ $(\lambda + 1)$-simple matrix $D = (d_{ij})$ over $\mathbb{Z}_g^2$. Then there is an (ng, w, λ)-CGOC of size $N|\mathcal{M}|$.*

The following result on r-simple matrices is useful.

Theorem 6 ([10])**.** *Let n be a positive integer with prime power factorization $n = \prod_i p_i^{\alpha_i}$. Let $p_{\min}$ be the smallest p_i and let r be a positive integer with $r \le p_{\min}$. Then there exists a $p_{\min} \times n^{2r}$ matrix which is $(r + 1)$-simple over $\mathbb{Z}_n^2$.*

V. A Direct Construction of GOCs from CGOCs

Recall that an (n, w, λ)-CGOC is also an (n, w, λ)-GOC. The following result therefore follows immediately from Theorem 4.

Corollary 7. *Let $p \ge 3$ be a prime and λ an integer with $2 \le \lambda \le p$. Then there is a (p, p, λ)-GOC of size $p^{\lambda-1} - 1$.*

When $\lambda = O\left(p^{1/4-\epsilon}\right)$ for some $\epsilon > 0$, the condition in Theorem 2 is satisfied. In other words, $M(p, p, \lambda) \le p^{\lambda-1} + o(p^{\lambda-1})$, so the codes in Corollary 7 are asymptotically optimal. However, when $\lambda = \Omega(p^{1/4})$, the condition in Theorem 2 does not hold. Indeed, for some values of λ satisfying $\lambda = \Theta(p)$, we construct, in this section, some GOCs with sizes $tp^{\lambda-1} - O(1)$, where t may be chosen to be greater than one.

In what follows, we canonically identify the elements in I_n with those in $\mathbb{Z}_n$. Given $M \subseteq I_n^2$ and $\mathbf{v} = (v_a, v_b) \in I_n^2$, let the *translation of M by $\mathbf{v}$ modulo n* be $M + \mathbf{v} \pmod n \triangleq \{(m_a + v_a \bmod n, m_b + v_b \bmod n) : (m_a, m_b) \in M\}$.

Let M be a w-subset of I_n^2 such that $|M \cap (\{i\} \times \mathbb{Z}_n)| \le 1$ for each $i \in \mathbb{Z}_n$. In other words, regarding M as an $n \times n$ $(0, 1)$-matrix, there is at most one 1 in each row of M. Clearly, for every $1 \le \delta \le w$, we can find a vector $\mathbf{v}(M, \delta) = (v(M, \delta), 0)$, with $v(M, \delta) \in I_n$, such that $|I_n^2 \cap (M + \mathbf{v}(M, \delta))| = \delta$. Let $\mathrm{tr}(M, \delta)$ denote $M + \mathbf{v}(M, \delta) \pmod n$.

Construction 1. *Suppose that there exists an (n, w, λ)-CGOC $\mathcal{M}$ such that, for each $M \in \mathcal{M}$ and each $i \in \mathbb{Z}_n$, we have $|M \cap (\{i\} \times \mathbb{Z}_n)| \le 1$. For any positive integer t with $t \le w$, let*

$$\mathrm{tr}(\mathcal{M}, i\gamma) = \{\mathrm{tr}(M, i\gamma) : M \in \mathcal{M}\},$$

where $i = 1, 2, \ldots, t$ and $\gamma = \lfloor w/t \rfloor$. Let

$$\mathcal{F} = \mathrm{tr}(\mathcal{M}, \gamma) \cup \mathrm{tr}(\mathcal{M}, 2\gamma) \cup \cdots \cup \mathrm{tr}(\mathcal{M}, t\gamma).$$

If $w - \lfloor w/t \rfloor \le \lambda$, then $\mathcal{F}$ is an (n, w, λ)-GOC of size $t|\mathcal{M}|$.

The key steps in the proof that Construction 1 works are as follows. We first check that each $\mathrm{tr}(\mathcal{M}, i\gamma)$ is an (n, w, λ)-CGOC. It then remains to show that the cross-correlation of $\mathrm{tr}(M, i\gamma)$ and $\mathrm{tr}(M', j\gamma)$ is $\le \lambda$, where $(M, i) \ne (M', j)$. When $M \ne M'$, the case of $i = j$ may be checked directly, while the case of $i \ne j$ follows from the auto- and cross-correlation of $\mathcal{M}$ as CGOC. The remaining case where $M = M'$ and $i \ne j$ may again be checked directly.

It can be verified that the CGOCs in Theorem 4 satisfy the condition of Construction 1. We may therefore apply Construction 1 to obtain the following class of (p, p, λ)-GOCs, whose size exceeds the OOC upper bound $p^{\lambda-1} + O(p^{\lambda-2})$.

Corollary 8. *Let $p \ge 3$ be a prime. Let λ and t be two positive integers with $t \le p$ and $p - \lfloor p/t \rfloor \le \lambda \le p$. Then there is a (p, p, λ)-GOC of size $tp^{\lambda-1} - t$.*

VI. A Recursive Construction for (n, w, λ)-GOCs

In this section, we introduce a recursive approach to construct (n, w, λ)-GOCs of large size. In addition to CGOCs, permutation codes constitute another key ingredient in our method.

Let S_n be the set of permutations on the set $\{1, 2 \ldots, n\}$. Write a permutation $\boldsymbol{\pi} \in S_n$ in the form $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_n)$. The *Hamming distance* between two permutations $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ and $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_n)$ in S_n is defined to be $d_H(\boldsymbol{\sigma}, \boldsymbol{\pi}) = |\{i : \sigma_i \neq \pi_i\}|$.

For $1 \leq d \leq n$, we say that $\emptyset \neq \mathcal{C} \subseteq S_n$ is an (n, d)-*permutation code* if $d_H(\boldsymbol{\sigma}, \boldsymbol{\pi}) \geq d$ for every two distinct permutations $\boldsymbol{\sigma}, \boldsymbol{\pi} \in \mathcal{C}$. Let the largest possible size of an (n, d)-permutation code be denoted by $P(n, d)$. Bounds on $P(n, d)$ and the exact values of $P(n, d)$ under some specific parameters have been studied in [12]. In particular, we have $P(n, d) \leq n!/(d-1)!$.

We now present a recursive construction for GOCs.

Construction 2. *Let $\mathcal{A} = \{A_1, A_2, \ldots, A_{m_1}\}$ be an (n_1, w, λ)-CGOC, let $\mathcal{C} = \{C_1, C_2, \ldots, C_{m_2}\}$ be an (n_2, w, λ)-GOC, and let $\mathcal{P} = \{\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots, \boldsymbol{\pi}_{m_0}\}$ be a $(w, w-\lambda)$-permutation code.*

For each $A_i = \{(a_{i1}, b_{i1}), (a_{i2}, b_{i2}), \ldots, (a_{iw}, b_{iw})\} \in \mathcal{A}$, $C_j = \{(c_{j1}, d_{j1}), (c_{j2}, d_{j2}), \ldots, (c_{jw}, d_{jw})\} \in \mathcal{C}$ and $\boldsymbol{\pi}_k \in \mathcal{P}$, construct a new codeword F_{ijk} as follows:

$$F_{ijk} = \{(a_{i\ell} + n_1 c_{j\boldsymbol{\pi}_k(\ell)}, b_{i\ell} + n_1 d_{j\boldsymbol{\pi}_k(\ell)}) : 1 \leq \ell \leq w\}.$$

Let

$$\mathcal{F} = \{F_{ijk} : 1 \leq i \leq m_1, 1 \leq j \leq m_2, 1 \leq k \leq m_0\}.$$

Then $\mathcal{F}$ is an $(n_1 n_2, w, \lambda)$-GOC of size $m_0 m_1 m_2$.

To show that the auto-correlation of any F_{ijk} is $\leq \lambda$, we use the auto-correlation property of $\mathcal{A}$. For the cross-correlation between F_{ijk} and $F_{i'j'k'}$: (i) When $i \neq i'$, we use the cross-correlation property of $\mathcal{A}$; (ii) When $i = i'$: one subcase uses the auto- and cross-correlation properties of $\mathcal{C}$, while the remaining subcase uses the auto-correlation property of $\mathcal{A}$.

In Construction 2, suppose that

$$\begin{aligned} m_1 &= \alpha \frac{n_1^{2\lambda}}{w(w-1)\cdots(w-\lambda)}, \\ m_2 &= \beta \frac{n_2^{2\lambda}}{w(w-1)\cdots(w-\lambda)}, \text{ and} \\ m_0 &= \gamma w(w-1)\cdots(w-\lambda), \end{aligned}$$

where $\alpha, \gamma \leq 1$ and $\beta \leq (\lambda+1)^2$. Then we can obtain an $(n_1 n_2, w, \lambda)$-GOC of size $\alpha\beta\gamma(n_1 n_2)^{2\lambda}/w(w-1)\cdots(w-\lambda)$. Recall that

$$U_{\mathrm{GOC}}(n_1 n_2, w, \lambda) = (\lambda+1)^2 \frac{(n_1 n_2)^{2\lambda}}{w(w-1)\cdots(w-\lambda)} + o\left((n_1 n_2)^{2\lambda}\right).$$

Hence, if α and γ are close to 1 and β is close to $(\lambda+1)^2$, the size of the resultant code is close to this upper bound. In other words, if the ingredients $\mathcal{A}, \mathcal{C}$, and $\mathcal{P}$ in Construction 2 are "close to optimal", then the new GOC $\mathcal{F}$ obtained is also "close to optimal".

TABLE I
Comparison of the coefficient c with U^*_{GOC} and U^*_{OOC}

n_1	n_2	w	λ	U^*_{OOC}	U^*_{GOC}	c
4	3	4	2	864	7,776	2,112
5	4	4	2	6,666	60,000	23,688
7	6	4	2	129,654	1,166,886	644,160
4	4	5	2	1,092	9,830	2,520
5	5	5	2	6,510	58,593	17,640
7	6	5	2	51,861	466,754	155,520
7	5	6	2	12,505	112,546	16,200
7	6	6	2	25,930	233,377	36,720
3	4	5	3	24,883	398,133	37,440
5	5	5	3	2,034,505	32,552,083	6,748,800
7	6	5	3	45,741,931	731,870,899	198,374,400
7	4	6	3	1,338,584	21,417,346	1,651,680
7	6	6	3	15,247,310	243,956,966	30,636,000
4	6	6	4	155,882,380	3,822,059,520	322,560,000

When $w \leq 6$ and $\lambda < w$, Chu *et al.* [12] showed that a $(w, w-\lambda)$-permutation code of size $w(w-1)\cdots(w-\lambda)$ exists. Therefore, we have the following result on the size of codes resulting from Construction 2 for $w \leq 6$.

Theorem 9. *Let $w \in \{3, 4, 5, 6\}$ and let N be a positive integer whose prime factors are at least w. Suppose that there exists an (n_1, w, λ)-CGOC of size m_1 and an (n_2, w, λ)-GOC of size m_2. Then there exists an $(n_1 n_2 N, w, \lambda)$-GOC of size $cN^{2\lambda}$, where $c = m_1 m_2 w(w-1)\cdots(w-\lambda)$.*

Proof. From Theorem 6, there is a $w \times N^{2\lambda}$ matrix which is $(\lambda+1)$-simple over $\mathbb{Z}_N^2$. By Theorem 5, this matrix and the (n_1, w, λ)-CGOC yield an $(n_1 N, w, \lambda)$-CGOC of size $m_1 N^{2\lambda}$. Applying Construction 2 with this CGOC, together with the (n_2, w, λ)-GOC and the $(w, w-\lambda)$-permutation code from [12], then yields an $(n_1 n_2 N, w, \lambda)$-GOC with the desired size. □

We obtain some lower bounds on the sizes of (n_1, w, λ)-CGOCs and (n_2, w, λ)-GOCs for $w \leq 6$ by computer search. Then, by applying Theorem 9, we obtain some $(n_1 n_2 N, w, \lambda)$-GOCs of size $cN^{2\lambda}$, with c listed in Table I. Recall that

$$\begin{aligned} U_{\mathrm{GOC}}(n_1 n_2 N, w, \lambda) &= \frac{(\lambda+1)^2 (n_1 n_2)^{2\lambda}}{w(w-1)\cdots(w-\lambda)} N^{2\lambda} + o(N^{2\lambda}), \\ U_{\mathrm{OOC}}((n_1 n_2 N)^2, w, \lambda) &= \frac{(n_1 n_2)^{2\lambda}}{w(w-1)\cdots(w-\lambda)} N^{2\lambda} + o(N^{2\lambda}). \end{aligned}$$

In Table I, the coefficients (of $N^{2\lambda}$)

$$\begin{aligned} U^*_{\mathrm{GOC}}(n_1 n_2, w, \lambda) &= \frac{(\lambda+1)^2 (n_1 n_2)^{2\lambda}}{w(w-1)\cdots(w-\lambda)} \text{ and} \\ U^*_{\mathrm{OOC}}((n_1 n_2)^2, w, \lambda) &= \frac{(n_1 n_2)^{2\lambda}}{w(w-1)\cdots(w-\lambda)}, \end{aligned}$$

are also listed (abbreviated as simply U^*_{OOC} and U^*_{GOC}) for comparison with c. We note that c is significantly greater than the corresponding U^*_{OOC} in all these cases. These are again examples of GOCs with size exceeding the OOC upper bound.

VII. Constructions for Optimal GOCs with $\lambda = 1$

We focus now on the case $\lambda = 1$ and investigate the existence of $(n, w, 1)$-GOCs attaining the upper bound $\frac{4n(n-1)}{w(w-1)}$. Our

constructions are based on some combinatorial structures, which we now introduce.

Let v be a positive integer. Let $\mathcal{B} = \{B_1, B_2, \ldots, B_m\}$, where $B_i = (b_{i1}, b_{i2}, \ldots, b_{ik})$, be a family of (ordered) k-tuple of I_v called *blocks*. The *list of directed differences from* B_i is defined to be the multiset $\Delta B_i = \{b_{ik} - b_{ij} : 1 \leq j < k \leq w\}$ for $1 \leq i \leq m$, while the *list of directed differences from* $\mathcal{B}$ is defined to be the multiset union $\Delta\mathcal{B} = \cup_{i=1}^{m} \Delta B_i$. If $\Delta\mathcal{B} = \{1, 2, \ldots, (v-1)/2\}$, then $\mathcal{B}$ is called a *perfect difference family*, or briefly, a $(v, k, 1)$-PDF. Note that, if $\mathcal{B} = \{B_1, B_2, \ldots, B_m\}$ is a $(v, k, 1)$-PDF, we must have $m = (v-1)/(k(k-1))$.

For each block $B_i = (b_{i1}, b_{i2}, \ldots, b_{ik})$ of a $(v, k, 1)$-PDF, let $B'_i = (0, b_{i2} - b_{i1}, \ldots, b_{ik} - b_{i1})$. Then $\Delta B'_i = \Delta B_i$, so $\mathcal{B}' = \{B'_1, B'_2, \ldots, B'_m\}$ is a family of k-tuple of $I_{(v+1)/2}$, with $\Delta\mathcal{B}' = \cup_{i=1}^{m} \Delta B'_i = \{1, 2, \ldots, (v-1)/2\}$. (It is also a $(v, k, 1)$-PDF.) We call such a structure a *strictly perfect difference family* (SPDF). Specifically, an $(n, k, 1)$-SPDF $\mathcal{B} = \{B_1, B_2, \ldots, B_m\}$ is a family of k-tuples of I_n such that $\Delta\mathcal{B} = \{1, 2, \ldots, n-1\}$.

The argument above proves the following fact.

Lemma 10. *A $(v, k, 1)$-PDF exists if and only if a $((v+1)/2, k, 1)$-SPDF exists.*

Another ingredient needed for our construction is the class of *strictly perfect difference matrices* (SPDMs). An SPDM(k, n) is a $k \times (2n-1)$ matrix with entries from I_n such that, for all $1 \leq s < t \leq k$, the *list of differences* $D_{st} = \{d_{sj} - d_{tj} : 1 \leq j \leq 2n-1\} = \{-(n-1), \ldots, -1, 0, 1, \ldots, n-1\}$.

Construction 3. *Suppose that there exist an $(n, w, 1)$-SPDF and an SPDM(w, n). Then an $(n, w, 1)$-GOC of size $\frac{4n(n-1)}{w(w-1)}$ exists.*

This construction shows that SPDFs and SPDMs can be used to construct optimal GOCs. We show further that SPDMs can also be constructed from SPDFs.

An *orthogonal array* $\mathrm{OA}(m, n)$ is an $m \times n^2$ array A, with entries from a set X of n elements, such that, when restricted to any two rows of A, every ordered pair of elements from X occurs in exactly one column of the restricted array. An orthogonal array A is *idempotent* if it contains the n distinct $m \times 1$ vectors $\{(x, x, \ldots, x)^T : x \in X\}$ as columns.

Example 11. *For a prime power q, let $\mathbb{F}_q$ be the field of order q. Let A be a $q \times q^2$ array, with rows labeled by $x \in \mathbb{F}_q$ and columns by $(i, j) \in \mathbb{F}_q^2$, whose entry in row x and column (i, j) is $ix + j$. It is easy to check that A is an idempotent* $\mathrm{OA}(q, q)$.

Construction 4. *Suppose that there exist both an $(n, k, 1)$-SPDF and an idempotent* $\mathrm{OA}(w, k)$. *Then an SPDM(w, n) exists.*

The existence of an $(n, w, 1)$-SPDF implies that $w \leq 5$ [13]. By Example 11, an idempotent $\mathrm{OA}(w, w)$ exists for $2 \leq w \leq 5$. Construction 4 yields an SPDM(w, n). Then, by applying Construction 3, an optimal $(n, w, 1)$-GOC is obtained.

Proposition 12. *Suppose that there exists an $(n, w, 1)$-SPDF. Then an optimal $(n, w, 1)$-GOC of size $\frac{4n(n-1)}{w(w-1)}$ exists.*

As a consequence of Lemma 10, Proposition 12, and existence results on PDFs in [14], [15], we have the following result.

Corollary 13. $M(n, w, 1) = \frac{4n(n-1)}{w(w-1)}$ *when*

(i) $w = 3$, $n \equiv 1, 4 \pmod{12}$*; or*
(ii) $w = 4$, $n \equiv 1 \pmod{6}$, $n \leq 6001$ *and* $n \neq 13, 19$*; or*
(iii) $w = 5$, $n = 61, 81$ *or* 101.

VIII. Conclusion

In this paper, we provid two families of GOCs whose sizes are greater than OOCs of corresponding parameters by a factor greater than one. In addition, we provide a tighter asymptotic upper bound for GOCs in certain regime and show that it is asymptotically equal to the Johnson bound for OOCs in the same regime. Finally, we obtain some optimal GOCs codes which achieve the upper bound U_{GOC} for $\lambda = 1$.

Acknowledgement

The research of Y. M. Chee, H. M. Kiah and S. Ling is supported in part by the Singapore Ministry of Education under Research Grant MOE2015-T2-2-086.

References

[1] P. W. K. Rothemund, "Folding DNA to create nanoscale shapes and patterns," *Nature*, vol. 440, pp. 297–302, 2006.
[2] S. Woo and P. W. K. Rothemund, "Programmable molecular recognition based on the geometry of DNA nanostructures," *Nature Chemistry*, vol. 3, no. 8, pp. 620–627, 2011.
[3] D. Doty and A. Winslow, "Design of geometric molecular bonds," *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, vol. 3, no. 1, pp. 13–23, 2017.
[4] 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, 1989.
[5] P. Erdős, R. Graham, I. Z. Ruzsa, and H. Taylor, "Bounds for arrays of dots with distinct slopes on lengths," *Combinatorica*, vol. 12, no. 1, pp. 39–44, 1992.
[6] A. A. Hassan, J. E. Hershey, and G. J. Saulnier, "Spatial optical CDMA," in *Perspectives in Spread Spectrum*. Springer, 1998, pp. 107–125.
[7] K. Kitayama, "Novel spatial spread spectrum based fiber optic CDMA networks for image transmission," *IEEE J. Sel. Areas Commun.*, vol. 12, no. 4, pp. 762–772, 1994.
[8] G. C. Yang and W. C. Kwong, "Two-dimensional spatial signature patterns," *IEEE Trans. Commun.*, vol. 44, no. 2, pp. 184–191, 1996.
[9] S. M. Johnson, "A new upper bound for error-correcting codes," *IEEE Trans. Inform. Theory*, vol. 8, pp. 203–207, 1962.
[10] L. Ji, B. Ding, X. Wang, and G. Ge, "Asymptotically optimal optical orthogonal signature pattern codes," *preprint*.
[11] W. Chu and S. W. Golomb, "A new recursive construction for optical orthogonal codes," *IEEE Trans. Inform. Theory*, vol. 49, no. 11, pp. 3072–3076, 2003.
[12] W. Chu, C. J. Colbourn, and P. Dukes, "Constructions for permutation codes in powerline communications," *Des. Codes Cryptogr.*, vol. 32, no. 1-3, pp. 51–64, 2004.
[13] J.-C. Bermond, A. Kotzig, and J. Turgeon, "On a combinatorial problem of antennas in radioastronomy," in *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. I*, ser. Colloq. Math. Soc. János Bolyai. North-Holland, Amsterdam-New York, 1978, vol. 18, pp. 135–149.
[14] C. J. Colbourn and J. H. Dinitz, Eds., *The CRC handbook of combinatorial designs*, ser. CRC Press Series on Discrete Mathematics and its Applications. CRC Press, Boca Raton, FL, 1996.
[15] G. Ge, Y. Miao, and X. Sun, "Perfect difference families, perfect difference matrices, and related combinatorial structures," *J. Combin. Des.*, vol. 18, no. 6, pp. 415–449, 2010.