The Annals of Statistics
2017, Vol. 45, No. 6, 2483–2510
https://doi.org/10.1214/16-AOS1531

OPTIMAL DESIGN OF FMRI EXPERIMENTS USING CIRCULANT (ALMOST-)ORTHOGONAL ARRAYS[1]

BY YUAN-LUNG LIN*, FREDERICK KIN HING PHOA* AND MING-HUNG KAO†

Academia Sinica and Arizona State University†*

Functional magnetic resonance imaging (fMRI) is a pioneering technology for studying brain activity in response to mental stimuli. Although efficient designs on these fMRI experiments are important for rendering precise statistical inference on brain functions, they are not systematically constructed. Design with circulant property is crucial for estimating a hemodynamic response function (HRF) and discussing fMRI experimental optimality. In this paper, we develop a theory that not only successfully explains the structure of a circulant design, but also provides a method of constructing efficient fMRI designs systematically. We further provide a class of two-level circulant designs with good performance (statistically optimal), and they can be used to estimate the HRF of a stimulus type and study the comparison of two HRFs. Some efficient three- and four-levels circulant designs are also provided, and we proved the existence of a class of circulant orthogonal arrays.

1. Introduction. Rapid event-related functional Magnetic Resonance Imaging (ER-fMRI) allows the shape estimation of hemodynamic response function (HRF) associated with transient brain activation evoked by various mental stimuli. An ER-fMRI design is a sequence of stimuli to be presented to an experimental subject, and such design is regarded as a circulant design [16, 17]. In the study of a fMRI experiment, a design may contain tens to hundreds of stimuli. Each stimulus evokes cerebral neuronal activity, leading to a rise and fall in the ratio of oxy- to deoxy-blood in the cerebral blood vessels at a brain voxel (3D image unit), and a change in the strength of magnetic field is detected by the MR scanner. This change is described by a function of time called the hemodynamic responses function (HRF). After the onset of a stimulus, the HRF takes several second to completely return to its baseline. Statistical inference is made on the brain activity by an MR scanner that collects data via the repeated scans on a subject's brain.

Received January 2016; revised November 2016.
[1]Supported by (a) Career Development Award of Academia Sinica (Taiwan) Grant Number 103-CDA-M04 and (b) Ministry of Science and Technology (Taiwan) Grant Numbers 102-2628-M-001-002-MY3, 104-2118-M-001-016-MY2 and 105-2118-M-001-007-MY2.
MSC2010 subject classifications. Primary 05B10, 05B15; secondary 62K15.
Key words and phrases. Circulant almost orthogonal arrays, complete difference system, design efficiency, hemodynamic response function.

The inference about the HRF is thus of the main interest in most fMRI studies. See Lazar [20] for more details.

Buračas and Boynton [1] proposed the use of m-sequence to precisely estimate HRF. The good performance of m-sequence is reported in several studies [14, 25, 26]. A good property of a m-sequence $d_1 d_2 \dots d_n$ is every nonzero t-tuple appears exactly once in the set $\{(d_i, \dots, d_{i+t-1}) | i = 1, \dots, n\}$ where $d_{n+j} = d_j$. The length of an m-sequence is often set to $n = (Q+1)^l - 1$ where $Q+1$ is a prime, Q is the total number of stimulus types and l is a positive nonzero integer, for example, 11012202 is an m-sequence of length 8. However, the application is unfortunately limited due to the large gap of run size n, thus an extended m-sequence [16] is recommended. In specific, an additional 0 is inserted to a $(t-1)$-tuple of zero in an m-sequence, so that a zero t-tuple is included. In the previous example, the extended m-sequence can be 110012202 or 110122002. In literature review, m-sequence is widely used since it preserves (nearly) equal frequency of t-tuple across stimulus types. However, only few effects can be estimated. Highly efficient designs with flexible run sizes are thus called for.

Recently, Kao [17] proposed the use of Hadamard sequence (H-sequence), obtained by Paley difference set [29], for ER-fMRI experiments with one stimulus type. For example, 0010111 is an H-sequence. An obvious advantage of using H-sequence is its run size flexibility, but it only fits for specific $n \equiv 3 \pmod 4$. Then Craigen et al. [10] introduced the circulant partial Hadamard matrix (CPHM) for the purpose of solving the problems in stream cypher cryptanalysis. An $n \times n$ matrix $A = (a_{i,j})$ is *circulant* if $a_{i+1,j+1} = a_{i,j}$ where the subscripts are reduced modulo n. An r-row-regular *circulant partial Hadamard matrix* H, denoted by r-$H(k \times n)$, is an $k \times n$ circulant (± 1)-matrix with each row sum r such that $HH^T = n\mathbf{I}_k$. When $n \equiv 0 \pmod 4$, CPHMs with zero row sum are highly efficient designs for fMRI experiments [18]. Although the CPHM is more powerful and efficient than H-sequence, both of them are still important when different run sizes are required. In this work, our goal is to propose a unified method to construct circulant designs for fMRI experiments with any run sizes. Moreover, our method is also adapted for constructing circulant designs of any s-levels for $s \geq 2$.

The optimality of m-sequences, extended m-sequences, H-sequences and CPHMs are roughly reported as follows. The m-sequences are A-optimal by computational results in [1, 25]. Extended m-sequences are universally optimal [16, 18] for studies with two stimuli, and D-optimal for studies with stimulus type more than two. The H-sequences are ϕ_p-optimal for estimating a HRF when $p \in [0, 1]$ [8]. In addition, the H-sequences are universally optimal by inserting a 0 to a run of consecutive 0's, called extended H-sequences, and a CPHM is also universally optimal [8]. The definitions of the optimal criteria please refer to Appendix A. In 2015, Cheng and Kao [8] developed a general theory to guide the selection of fMRI designs for estimating a HRF and for conducting a comparison of two HRFs. Based on Φ_p-optimality criterion, they provided a strategy to the selection of fMRI designs under different parameter p when $n \equiv 0, 1, 3 \pmod 4$.

However, there are many research challenges such as the case $n \equiv 2 \pmod 4$. In this work, we introduce a unified structure that can construct not only the above sequences but also circulant designs with any run sizes.

The present study focuses on a generalized structure of circulant designs for any level setting. We propose a circulant design called *circulant (almost-)orthogonal array* (CAOA) that guarantees the frequency of all t-tuples to be almost equal. In the next section, we introduce some mathematical terminologies and a statistical model for estimating HRFs. In Section 3, the concept and properties of CAOAs are introduced and a class of CAOA is proposed. We then present the study of two-level CAOAs with various run sizes, and the optimality of these designs are discussed in Section 4. Furthermore, lists of three- and four-levels CAOAs are given in Section 5. In addition, we also proved the existence of circulant OAs. Some discussions on the proposed designs and a conclusion are given in the last section. For clarity, all proofs are organized in Appendix B.

2. Notation and background.

2.1. *Statistical model.* In a fMRI experiment, a mental stimulus to be presented to an experiment subject can possibly occur every τ_{ISI} seconds, where τ_{ISI} is a pre-specified time. An event-related fMRI sequence can be represented as an ordered sequence $\mathbf{d} = (d_1, \dots, d_n)$, where $d_i \in \{0, \dots, Q\}$, and Q is the total number of stimulus types. For example, an experiment with q stimulus types ($Q = q$) can be viewed as a $(q + 1)$-ary sequence $\mathbf{d}$. The qth stimulus (e.g., a picture of a familiar face) occurs at $(i - 1)\tau_{\mathrm{ISI}}$ when $d_i = q$, and there is no stimulus onset at $(j - 1)\tau_{\mathrm{ISI}}$ if $d_j = 0$. The study of the HRF helps us to understand the effects of the stimuli to the brain activity [20, 24].

We consider the following model for estimating the HRF (see also [11, 16, 26]):

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{S}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}. \tag{2.1}$$

Here, $\mathbf{y} = (y_1, \dots, y_n)$ where y_i is the measurement of a brain voxel collected by an MR scanner at the ith time point, $\mathbf{h} = (\mathbf{h}_1^T, \dots, \mathbf{h}_K^T)^T$ represents the unknown magnitudes of the HRFs, where $\mathbf{h}_i = (h_{1,i}, \dots, h_{Q,i})^T$, $h_{q,i}$ is the ith magnitude of the HRF from the qth stimulus type; K is determined by the duration of the HRF, counting from the onset of a stimulus to the HRFs complete return to baseline. The matrix $\mathbf{X} = [\mathbf{X}_{(1)}, \dots, \mathbf{X}_{(K)}]$ is a $n \times QK$ zero-one design matrix, where $\mathbf{X}_{(i)} = [\mathbf{x}_{1,i}, \dots, \mathbf{x}_{Q,i}]$ is the design matrix of the ith height of the Q HRFs and the ith element of the vector $\mathbf{x}_{q,i} = 1$ if $d_i = q$ and 0 otherwise. The vector $\mathbf{S}\boldsymbol{\gamma}$ is the nuisance term with a specified $\mathbf{S}$ and an unknown parameter $\boldsymbol{\gamma}$. The vector $\boldsymbol{\varepsilon}$ represents the noise with mean 0 and covariance matrix $\boldsymbol{\Sigma}$.

In this work, we assume that the last $K - 1$ elements of the design $\mathbf{d}$ are presented in the burn-in period before the first valid fMRI measurement. It is necessary to allow the MR scanner to reach a steady state in the burn-in period, and the measurements collected in this period are discarded from the subsequent statistical analysis. Thus, $\mathbf{X}_{(k)} = \mathbf{U}^{k-1}\mathbf{X}_{(1)}$ where $\mathbf{U} = (u_{i,j})_{n\times n}$ is a permutation matrix with $u_{i,i-1} = u_{1,n} = 1$, for $i = 2, \dots, n-1$, $k = 2, \dots, K$, and 0 otherwise. This implies that the design matrix $\mathbf{X}_{(i)}$ of Model (2.1) must be in circulant setting. Here, we adopt the statistical model proposed by Kao [16], which is a special case of Model (2.1), to estimate HRFs. In addition, the circulant property is one of the model assumptions. Please refer to [17, 18] for more details. The model on the estimation of a HRF and the comparison of two HRFs will be discussed in Section 4.

2.2. *Circulant designs.* In literatures, lots of good designs are applied into experiments for rending precise statistical inference such as orthogonal arrays. An *orthogonal array* (OA) of size n, with k constraints, s symbols and having strength t, denoted by $OA(n, k, s, t)$, is a $k \times n$ matrix $\mathbf{A}$ of s symbols such that all the ordered t-tuples of the symbols occur n/s^t times as column vectors of any $t \times n$ submatrix of $\mathbf{A}$; see [13] for more details. The advantages of using an OA as an experimental plan include the orthogonality and projectivity of effect estimates [5, 6, 33]. However, an obvious weakness of OA is its inflexible run size, which must be a multiple of s^t.

In the aspect of fMRI experiments, designs with circulant property are required for estimating HRFs, and such designs have not been studied in literatures. OA is useful and powerful, but it cannot be utilized in fMRI experiments. A Hadamard matrix is known to be an $OA(n, n-1, 2, 2)$ and it is conjectured to exist for any $n \equiv 0 \pmod 4$, but a circulant Hadamard matrix of order $n > 4$ is conjectured to be nonexistence [34]. A 0-$H(k \times n)$ is a two-symbol, n-run, k-factor *circulant orthogonal array*; it could be applied to fMRI experiments [18]. Given n, the study focuses on the maximum value of k such that an $k \times n$ CPHM exists. A computational result was given in [10, 21, 27] for $n \le 76$. A general theory that connects the general difference set and CPHM was proposed by Lin et al. [21]. An algorithm was provided to search for CPHMs, and the lower bounds were successfully improved. Since the CPHMs were first introduced for stream cypher cryptanalysis, two-level designs are the primary consideration. We introduced the *circulant (almost-)orthogonal array* (CAOA), which presents a general framework of circulant designs.

DEFINITION 2.1. A circulant $k \times n$ array $\mathbf{A}$ with entries from $Z_s = \{0, 1, \dots, s-1\}$ is said to be a circulant almost orthogonal array (CAOA) with s levels, strength t and bandwidth b, if each ordered t-tuple α based on Z_s occurs $\lambda(\alpha)$ times as column vectors of any $t \times n$ submatrix of $\mathbf{A}$ such that $|\lambda(\alpha) - \lambda(\beta)| \le b$ for any two t-tuples α and β; such array $\mathbf{A}$ is denoted by $CAOA(n, k, s, t, b)$. For convenience, its first row is called *the generating vector*.

The entries of a CAOA can be also defined on any s-element set by certain mapping if the description is clear. It is obvious that a 0-$H(K \times n)$ is equivalent to an $CAOA(n, K, 2, 2, 0)$ by replacing -1 with 0. When $s = 2$, the transpose of a CAOA can be regarded as the design matrix $\mathbf{X}$ in Model (2.1). When $s \geq 3$, a CAOA can be transformed to be $\mathbf{X}$ by proper mapping. For example, the first row of a $CAOA(n, K, 2, 2, 0)$ is an event-related fMRI sequence d. If the experimenters apply this sequence into a fMRI experiment, then they can estimate the HRF via Model (4.2). In addition, the parameter K is the key of how many time points of a HRF that can be independently estimated. Therefore, the larger the value of K, the greater the power of Model (4.2).

Traditionally, the run size of $OA(n, k, s, t)$ is constrained by $n \equiv 0 \pmod{s^t}$. We instead introduce the bandwidth of CAOA and guarantee that each t-tuple occurs at least $\lfloor n/s^t \rfloor$ number of times when $b \leq 1$, where $\lfloor \cdot \rfloor$ is a floor function. Two questions arise: (1) What is the maximum value of k such that a CAOA exists? (2) How to find a good circulant design when n is not a multiple of s^t? There is a class of generalized OAs called *partially balanced arrays* (BA) introduced by Chakravarti [2]. It tackles the simpler version of our two questions without the requirements of circulant and bandwidth property. BAs are used as multifactorial designs when efficient designs are not easy to find; for a detailed description, please refer to [2, 3]. Our CAOA is, by definition, more flexible than BA, and BA is in fact a special case of CAOA if the frequency of each t-tuple is pre-specified.

For a two-level experiment with 12 runs, one can choose $OA(12, 11, 2, 2)$, but it fails in a fMRI experiment. Since a circulant $OA(12, k, 2, 2)$ does not exist when $6 \leq k \leq 11$, our $CAOA(12, 5, 2, 2, 0)$ becomes the best choice to be applied. Moreover, if 14 runs are allowed to be performed in a fMRI experiment, then $CAOA(14, 7, 2, 2, 1)$ is better than $CAOA(12, 5, 2, 2, 0)$. Even if there are only five factors of interest in the experiment, one can obtain a good design by deleting the last two rows of $CAOA(14, 7, 2, 2, 1)$. Their generating vectors are listed in Table 1. These designs are constructed via *general difference set* (GDS) introduced by Lin et al. [21], and it is the first systematic method to construct CPHMs. We recall the definition of GDS here.

DEFINITION 2.2. A $(n, k; \lambda_1, \dots, \lambda_{n-1})$ GDS is a set $D = \{d_1, \dots, d_k\}$ of distinct elements of Z_n such that the difference l appears λ_l times in the multi-set $\{d_i - d_j \pmod{n} | d_i, d_j \in D, i \neq j\}$ for $l = 1, \dots, n-1$.

For example, let $D = \{1, 2, 6, 8\} \subset Z_8$, then the collection of the differences of any two elements in D is $\{7, 1, 3, 5, 1, 7, 4, 4, 2, 6, 2, 6\}$. Thus D is a $(8, 4; 2, 2, 1, 2, 1, 2, 2)$ GDS.

3. Structure and properties of CAOAs. The GDS method is an efficient tool for searching two-level CAOAs, however, it is not applicable in multi-level cases. We are going to introduce a new system to describe a circulant structure of multi-level designs, which can be considered as an extension of GDS. Suppose X_i and X_j are two subsets of Z_n. A *difference frequency set* (DFS) of an

TABLE 1
The generating vectors of CAOA$(n, k, 2, 2, b)$, *for* $6 \le n \le 50$. T_2-*CAOAs are marked by* n^*

n	***k***	***b***	**Generating vector**
6	2	1	000111
6*	3	1	001011
7	7	1	0111001
8	3	0	00111010
9	3	1	001011110
10	3	1	0001011110
10*	5	1	0001101011
11	11	1	00010110111
12	5	0	001001111010
13	5	1	0010001011111
14	4	1	01001111101000
14*	7	1	00010111001011
15	15	1	001111010110010
16	7	0	0001110111010010
17	6	1	01110100001001111
18	6	1	001110111101000001
18*	8	1	000011010100110111
19	19	1	0010101111001101100
20	7	0	00010101100111101100
21	8	1	010101101111100110000
22	7	1	0010100111111011000010
22*	11	1	0001001011100010110111
23	23	1	00011111010110011001010
24	9	0	011000000110100111011101
25	9	1	0011101011111011000100100
26	9	1	00000010001110101111011011
26*	13	1	00001101010110000110111011
27	12	1	000011011011110101000100111
28	9	0	0000001010110011111001101011
29	11	1	00010001001111001111110100101
30	10	1	000000111001101111101011010001
30*	11	1	010011011000011110111000100101
31	31	1	0100001110101000111101101110010
32	12	0	00011101111100101101010000011001
33	12	1	000100001111011001111101011010001
34	11	1	0000011001010101101101111110001100
34*	17	1	0010111001110100000101110011101001
35	35	1	01001101010000100111011111000111010
36	14	0	011101011111101000011010010001001100
37	13	1	0010000101000100110001111101011011111
38	12	1	00110010110011111110101110000001010010
38*	19	1	01100001010111100100110000101011110011
39	15	1	000111010100110010111001111010110110000
40	17	0	0001101101111110001111011000100101010100010
41	14	1	01011100001011011101110111100101100010000

TABLE 1
(*Continued*)

n	***k***	***b***	**Generating vector**
42	13	1	010100000010111100111011111010011100100100
42*	18	1	001011000010101110110011110000100011011101
43	43	1	0110101100010000001110100011111011100101 0011
44	16	0	00000011100100010010010111110101110011101011
45	16	1	011010110100000010100010011001111001111110100
46	14	1	0110011111111010011110000100010010000110101010
46*	23	1	0000101001100110101111000001010011001101011111
47	47	1	00001000011010100011011001001110101001111011111
48	17	0	000000010011011000111011101011000111110010110101
49	17	1	0000000101100100101111010011100111110101011100110
50	16	1	01011011011101101010001110000000110001000001101111 1
50*	11	1	00111010010001110110101111000100000101110011011010

ordered pair (X_i, X_j) is a multi-set $\{a - b \pmod{n} | a \in X_i, b \in X_j\}$, denoted by $DFS_n(X_i, X_j)$. The notation $\lambda_l^{i,j}$ is the occurrence frequency of the nonzero element $l \in Z_n$ in the $DFS_n(X_i, X_j)$. In general, the difference zero is not considered, and thus it is omitted in the notation of this paper. If $X_i = X_j$, then $DFS_n(X_i, X_i)$ shows the frequency of each difference except the element zero in a X_i. Thus, $DFS_n(X_i, X_i)$ describes the structure of the GDS X_i. A partitioned set $V = \{V_0, V_1, \ldots, V_{s-1}\}$ is an *equitable partition* if $||V_i| - |V_j|| \le 1$ for all $i \ne j$ where $|V_i|$ is the cardinality of the set V_i. In summary, a GDS presents the difference structure of any two elements in a group, and a DFS describes the difference of any two elements in different groups.

We then define *complete difference system* (CDS) that summarizes the information from GDS and DFS, to understand the whole difference structure. Let $V = \{V_0, V_1, \ldots, V_{s-1}\}$ be a partition of Z_n. An r-frequency matrix of V is an $s \times s$ matrix $\mathbf{\Lambda}_r = (\lambda_r^{i,j})$ where $\lambda_r^{i,j}$ is the frequency of the nonzero element $r \in Z_n$ in $DFS_n(V_i, V_j)$. A CDS of V is an ordered $(n-1)$-tuple $(\mathbf{\Lambda}_1, \ldots, \mathbf{\Lambda}_{n-1})$ that describes frequency matrices of V. Let $I_D(\mathbf{\Lambda})$ be the smallest index $k \ge 2$ such that $\mathbf{\Lambda}_1 = \cdots = \mathbf{\Lambda}_{k-1} = \mathbf{\Lambda}$ but $\mathbf{\Lambda}_k \ne \mathbf{\Lambda}$. If $\mathbf{\Lambda}_i = \mathbf{\Lambda}$ for all i, then $I_D(\mathbf{\Lambda}) = \infty$. If $\mathbf{\Lambda}_1 \ne \mathbf{\Lambda}$, then $I_D(\mathbf{\Lambda}) = 1$. Given a frequency matrix $\mathbf{\Lambda}$, we say $V = \{V_0, \ldots, V_{s-1}\}$ is an $(n, k, s, \mathbf{\Lambda})$-CDS if V is a partition of Z_n and $I_D(\mathbf{\Lambda}) = k$. Its *incidence matrix* is defined as follows. Please refer to Example 3.4 for a simple demonstration.

DEFINITION 3.1. Let V be an $(n, k, s, \mathbf{\Lambda})$-CDS. The incidence matrix of V is an $k \times n$ matrix $\mathbf{A} = (a_{i,j})$ defined by

$$a_{i,j} = l \qquad \text{if } j \in V_l + (i-1),$$

where $V_l + (i-1) = \{x + (i-1) | \text{for all } x \in V_l\}$ and all elements are reduced modulo n; $i = 1, \ldots, k$, $j = 1, \ldots, n$ and $l = 0, \ldots, s-1$.

The r-frequency matrix is crucial for understanding the circulant structure. It describes the framework between ith and $(i+r)$th rows. Given a partition V, all difference compositions can be quickly grasped through CDS. We then show the equivalence relation between CDS and CAOAs.

THEOREM 3.2. *Let V be an $(n, k, s, \mathbf{\Lambda})$-CDS with $s, k \geq 2$ and a given frequency matrix $\mathbf{\Lambda}$. Each $2 \times n$ subarray, consisting of the ith and jth rows of the incidence matrix of V, contains each ordered pair exactly $\lambda_{j-i}^{x,y}$ times, where $\lambda_{j-i}^{x,y}$ is the entry of $\mathbf{\Lambda}_{j-i}$, $1 \leq i < j \leq n$, $0 \leq x, y \leq s-1$.*

Theorem 3.2 implies that an $(n, k, s, \mathbf{\Lambda})$-CDS is equivalent to a CAOA of strength two. In addition, the bandwidth of a CAOA is relevant to the frequency matrix $\mathbf{\Lambda}$. We denote the bandwidth of a matrix $\mathbf{M} = (m_{i,j})$ by $B(\mathbf{M}) = \max\{m_{i,j} - m_{i',j'} | \text{for all } m_{i,j}, m_{i',j'}\}$. Then the following corollary follows.

COROLLARY 3.3. *A $CAOA(n, k, s, 2, b)$ exists if and only if there exists an $(n, k, s, \mathbf{\Lambda})$-CDS such that $B(\mathbf{\Lambda}) = b$. In addition, the incidence matrix of $(n, k, s, \mathbf{\Lambda})$-CDS is the required CAOA.*

Instead of searching all combinations completely and counting the frequency of all pairs, the CDS summarizes the information of all differences efficiently. For instance, let $V = \{V_0, V_1\}$ where $V_0 = \{1, 2, 3, 5, 9, 10, 12\}$ and $V_1 = \{4, 6, 7, 8, 11, 13, 14\}$. It is easy to verify that V is a $(14, 7, 2, \mathbf{\Lambda})$-CDS with $\mathbf{\Lambda} = 4\mathbf{J}_2 - \mathbf{I}_2$, where $\mathbf{J}_2$ is a square all-ones matrix of order 2 and $\mathbf{I}_2$ is an identity matrix of order 2. Its incidence matrix is a $CAOA(14, 7, 2, 2, 1)$. Assume $\lambda^{i,j}$ is the (i, j)-entry in $\mathbf{\Lambda}$, the $\lambda^{i,j}$ represents the frequency of (i, j) pair in any 2×14 subarray and describes the frequency of the element $r \in Z_{14}$ in $DFS_{14}(V_i, V_j)$ for $r = 1, \ldots, 6$. As we mentioned before, our method CDS is applicable for constructing circulant OAs. We give an example as follows.

EXAMPLE 3.4. Let $n = 18$, $V_0 = \{1, 2, 3, 9, 14, 17\}$, $V_1 = \{5, 8, 10, 11, 12, 18\}$ and $V_2 = \{4, 6, 7, 13, 15, 16\}$. By simply counting the differences of GDSs V_i and $DFS_n(V_i, V_j)$ for $i \neq j$, it is easy to obtain $\lambda_r^{i,j} = 2$ for $r = 1, 2, 3$. Therefore, $V = \{V_0, V_1, V_2\}$ is a $(18, 4, 3, 2\mathbf{J}_3)$-CDS. By Corollary 3.3, its incidence matrix is a $CAOA(18, 4, 3, 2, 0)$ shown below:

$$\begin{pmatrix} 0 & 0 & 0 & 2 & 1 & 2 & 2 & 1 & 0 & 1 & 1 & 1 & 2 & 0 & 2 & 2 & 0 & 1 \\ 1 & 0 & 0 & 0 & 2 & 1 & 2 & 2 & 1 & 0 & 1 & 1 & 1 & 2 & 0 & 2 & 2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 2 & 1 & 2 & 2 & 1 & 0 & 1 & 1 & 1 & 2 & 0 & 2 & 2 \\ 2 & 0 & 1 & 0 & 0 & 0 & 2 & 1 & 2 & 2 & 1 & 0 & 1 & 1 & 1 & 2 & 0 & 2 \end{pmatrix}.$$

In addition, it is a circulant $OA(18, 4, 3, 2)$.

A fMRI experiment of $n = 18$ time points with three stimuli is considered. Traditionally, an m-sequence with length $3^3 - 1 = 26$ is utilized as the experimental plan. Kao et al. [19] indicated that m-sequences can be suboptimal under A-optimality. However, due to its large sequence length, the truncated m-sequence is used in practice even though it often loses its original efficiency. On the other hand, an extended m-sequence of length 3^2 is considered as another candidate due to its D-optimality [16]. However, its length is too short and only two time points height of each HRF (i.e., $k = 2$) can be analyzed. Instead of using these variants of m-sequences, our $CAOA(18, 4, 3, 2, 0)$ is a better candidate. Our design can disentangle the aggregate HRFs at the first four time points. It is a circulant orthogonal array (i.e., $b = 0$) that allows us to independently estimate four time points height of each HRF. Theorem 2 of [16] suggests that it is a D-optimal design, and it might be universally optimal.

Recall that a $CAOA(n, k, 2, 2, 0)$ is a CPHM for $n \equiv 0 \pmod 4$. Lin et al. [21] proposed an algorithm to search for a specific GDS such that a $CAOA(n, k, 2, 2, 0)$ has maximum value of k. Indeed, a $CAOA(n, k, 2, 2, 0)$ can be constructed by a CDS $V = \{D, \bar{D}\}$, where D is a GDS and $\bar{D}$ is its complement.

In view of foregoing discussion, the existence of a $CAOA(n, k, s, 2, b)$ is equivalent to the existence of a specific $(n, k, s, \mathbf{\Lambda})$-CDS with $B(\mathbf{\Lambda}) = b$. Next, we focus on the existence of $(n, k, s, \mathbf{\Lambda})$-CDS. A $(n, k, s, \mathbf{\Lambda})$-CDS is not guaranteed to exist if we arbitrarily choose a frequency matrix $\mathbf{\Lambda}$. For example, a $(12, 2, 3, \mathbf{\Lambda})$-CDS does not exist if $\mathbf{\Lambda} = (\lambda^{i,j})_{3\times 3}$ with $\lambda^{0,0} = \lambda^{1,1} = \lambda^{0,2} = 2$ and $\lambda^{i,j} = 1$ otherwise. We propose some useful properties, based on the CDS, for selecting a suitable $\mathbf{\Lambda}$.

PROPOSITION 3.5. *Let* $V = \{V_0, V_1, \ldots, V_{s-1}\}$ *be a partition of* Z_n, *and* $(\mathbf{\Lambda}_1, \ldots, \mathbf{\Lambda}_{n-1})$ *be its CDS. For all* $r \in Z_n \setminus \{0\}$, *we have*:

(a) $\lambda_r^{i,j} = \lambda_{n-r}^{j,i}$,
(b) $\sum_{j=0}^{s-1} \lambda_r^{i,j} = |V_i|$ *for any fixed* i *and* $\sum_{i=0}^{s-1} \lambda_r^{i,j} = |V_j|$ *for any fixed* j,
(c) $\sum_{0\le i,j\le s-1} \lambda_r^{i,j} = n$.

Using Proposition 3.5(b), our search becomes efficient by avoiding the search of many nonexistence CDS. The details are discussed in Section 5. Continuing the previous example, since $\sum_{j=0}^{2} \lambda_r^{0,j} \neq \sum_{i=0}^{2} \lambda_r^{i,0}$ for all r, there is no such $(12, 2, 3, \mathbf{\Lambda})$-CDS. According to Proposition 3.5(a), $0 < I_D(\mathbf{\Lambda}) < \frac{n}{2}$ or $I_D(\mathbf{\Lambda}) = \infty$ for any $\mathbf{\Lambda}$, so $k \le \lfloor n/2 \rfloor$ if $k \neq n$. Then a simple upper bound is derived by counting $I_D(\mathbf{\Lambda})$ of a CDS via Corollary 3.3.

PROPOSITION 3.6. *Let* $s \ge 3$, $\mathbf{\Lambda} = (\lambda^{i,j})_{i,j\in Z_s}$ *be the frequency matrix of a* $(n, k, s, \mathbf{\Lambda})$-*CDS and* $B(\mathbf{\Lambda}) = b$. *If a* $CAOA(n, k, s, t, b)$ *with* $\mathbf{\Lambda}$ *exists*, *then*

$$k \le \min\{|V_i|(|V_i| - 1)/2\lambda^{i,i} \mid i = 0, 1, \ldots, s-1\} + 1.$$

The upper bound is general enough that can be treated as a threshold in computer search. The case of $s = 2$ is slightly different and will be discussed in the next section and the choice of the frequency matrix $\mathbf{\Lambda}$ is discussed in Section 5. We have a class of CAOAs that reach the upper bound. Such CAOAs can be constructed by an m-sequence of length $q^m - 1$. A well-known property of m-sequence is that every nonzero t-tuple occurs equal times as we collect all consecutive t elements along the sequence. However, another important property is called two-tuple balance property [12]. In terms of CDS terminology, if we construct a CAOA by the m-sequence, then its frequency matrix equals to $\mathbf{\Lambda} = (\lambda^{i,j})_{q\times q}$, where $\lambda^{0,0} = q^{m-2} - 1$ and $\lambda^{i,j} = q^{m-2}$ otherwise. Hence, we have the following lemma.

LEMMA 3.7. *If q is a prime power and $m \geq 2$, then there exists a CAOA$(q^m - 1, (q^m - 1)/(q - 1), q, 2, 1)$.*

It can be proven by linear algebra [12], however, it can also be proven by CDS. Consider an m-dimensional Euclidean geometry on a finite field with q elements. There are q parallel $(m - 1)$-flats; they form a partition of all points. One flat forms a $(\frac{q^m-1}{q-1}, \frac{q^{m-1}-1}{q-1}, \frac{q^{m-2}-1}{q-1})$ Singer's difference set corresponding to an $(m - 1)$-dimensional projective geometry [35]; the others form a GDS individually [31, 32]. Any two distinct $(m - 1)$-flats also have special difference structures; it can be proven by shifting one of these two flats and discussing their DFS.

It is not easy to understand the matrix structure of an m-sequence despite that it has good properties. In coding theory, a code word which is a column vector of a zero-one matrix. Since the Hamming distance between two code words is relevant to its correcting ability, the relationship of columns is mainly of interest. For instance, for a binary m-sequence of length $2^5 - 1$, each nonzero t-tuple occurs 2^{5-t} times and the zero t-tuple occurs $2^{5-t} - 1$ times for $1 \leq t \leq 5$. If we use such m-sequence to construct a circulant matrix, then every binary code word of length 5 is one-to-one corresponding to each of its column. Therefore, they usually focus on the columns not rows. However, we aim at discussing the relationship between any two rows. On the other hand, the sequence structure of m-sequences has been widely studied, but its matrix structure is unclear. In the above example, its matrix structure is a 31×31 circulant matrix of strength two but not strength three. However, the m-sequence of length 31 corresponds to a *CAOA*$(31, 7, 2, 3, 1)$, *CAOA*$(31, 6, 2, 4, 1)$ and *CAOA*$(31, 5, 2, 5, 1)$, respectively.

In addition, Liu [25] recommended a truncated m-sequence that is obtained by leaving out the last $l - n$ elements of an m-sequence of length $l > n$. Such variant of m-sequence can suffer efficiency loss, and the reason can be easily explained through CDS. Roughly speaking, the q parallel $(m - 1)$-flats guarantee the difference system of m-sequence; however, the truncated m-sequence destroys such system. This implies that the frequency of differences of any two points on the same flat and on different flats are orderless.

The CDS presents circulant matrix structure in a difference method point of view; it helps us to understand the matrix structure of a circulant matrix. The construction of CAOAs with high strength is still under investigation.

We now introduce another simple construction method, called the *doubling method*, that can obtain large CAOAs from the repetition of some small one.

LEMMA 3.8. *For any positive integer, $l \geq 2$. If there is a CAOA(n, k, s, t, b), then there exists a CAOA(ln, k, s, t, lb).*

The above method is an easy and quick way to obtain a CAOA of large size, and it is very useful when $b = 0$. Its application will be discussed in Sections 4 and 5.

4. Two-level CAOA for estimating HRF. In this section, we concentrate on the optimal fMRI designs for estimating a HRF of one stimulus type and comparing the HRFs (or effects) of two stimulus types, and their constructions. Kao [18] studied optimal fMRI designs by considering the following special case of Model (2.1):

$$\mathbf{y} = \boldsymbol{\gamma}\mathbf{j}_n + \mathbf{X}_d\mathbf{h} + \boldsymbol{\varepsilon}, \tag{4.1}$$

where $\mathbf{j}_n$ is an all-ones vector, $\mathbf{X}_d = [d, \mathbf{U}d, \dots, \mathbf{U}^{K-1}d]$, d is a fMRI sequence and $\mathbf{U}$ is a permutation matrix. Model (4.1) estimates a HRF of one stimulus type. The next model compares the HRFs of two stimulus types:

$$y_i = \gamma + \sum_{k=0}^{K-1} \{x_{1,i-k}h_{1,k+1} + x_{2,i-k}h_{2,k+1}\} + \varepsilon_i, \qquad 1 \leq i \leq n, \tag{4.2}$$

where y_i is the fMRI measurement at the ith time point, $h_{q,i}$ is the HRF of the qth stimulus type at the ith time point, $x_{q,i}$ is an indicator for $q = 1, 2$ such that $x_{q,i} = 1$ when $d_i = q$ and 0 otherwise, the second subscript of x is reduced modulo n and the remaining terms are as in Model (2.1).

The height difference between two HFRs, say $\theta_k = h_{1,k} - h_{2,k}$, is of special interest and Model (4.2) can be rewritten as follows:

$$y_i = \gamma + \sum_{k=0}^{K-1} \{a_{i,k}\zeta_{k+1} + b_{i,k}\theta_{k+1}\} + \varepsilon_i, \qquad 1 \leq i \leq n, \tag{4.3}$$

where $a_{i,k} = (x_{1,i-k} + x_{2,i-k})/2$, $b_{i,k} = (x_{1,i-k} - x_{2,i-k})/2$, $\zeta_k = h_{1,k} + h_{2,k}$, and $\theta_k = h_{1,k} - h_{2,k}$. The studies of these models have been discussed in [8, 16–18, 21].

Let $\mathbf{D} = (d_{i,j})_{n \times K}$ be the transpose of a *CAOA*$(n, K, 2, 2, b)$ with symbols $q \in \{1, 2\}$, and $x_{q,i-k} = 1$ when $d_{i,k} = q$. Then a fMRI design $\mathbf{d} = (d_{i,1})$ is represented as the first row of a CAOA, and its optimality is equivalent to the optimality of a CAOA. Recently, Cheng and Kao [8] comprehensively discussed the cases $n \equiv$

$0, 1, 3 \pmod 4$ and developed a theory to guide the selection of optimal fMRI designs. Although the optimality of these designs has already been proven, known results are still missing. Even in a small range from 4 to 50, many of them are unknown. The purpose of our study is to search for CAOAs whose K is maximum, and to fill the gap of the known results in $n \equiv 0 \pmod 4$ (CPHMs) and $n \equiv 3 \pmod 4$ (H-sequences). Our results are discussed in the separate subsections for $n \equiv 0, 1, 2, 3 \pmod 4$, respectively. For the definition of all optimality criteria, please refer to Appendix A.

4.1. $n \equiv 0 \pmod 4$. As we mentioned before, a 0-$H(K \times n)$ is a $CAOA(n, K, 2, 2, 0)$. According to Lin et al. [21], each 0-$H(K \times n)$ possesses maximum value of k for $n \leq 52$ and lower bounds of K are derived for $56 \leq n \leq 76$. Evidently, these results are better than extended m-sequences as its K is usually very small. Another construction of CPHMs proposed by Cheng and Kao [8] inserts a 0 to a run of g 0's in a H-sequence. For example, one can obtain a H-sequence of length $n = 131$ via a Paley difference set, and a 0 is then inserted to obtain a $CAOA(132, 9, 2, 2, 0)$. However, a $CAOA(32, 12, 2, 2, 0)$ in Table 1 can precisely estimate the contrast $h_{1,k} - h_{2,k}$ for $k = 1, \ldots, 12$. The design that we obtain is shorter ($32 \ll 132$), and can accommodate a HRF with a longer duration ($12 > 9$).

The optimality of a fMRI design of length $n \equiv 0 \pmod 4$ has been proved by Kao [18]. Let $\mathcal{D}_n$ be the collection of all fMRI designs with length n. For any design $\mathbf{d} = (d_1, \ldots, d_n) \in \mathcal{D}_n$, let $n_k^{(pq)} = \#\{i \mid (d_{i-k}, d_i) = (p, q), i = 1, \ldots, n\}$ be the number of time points when a p is preceded by a q at a time distance k. Here, $d_{i-k} = d_{n+i-k}$ when $i \leq k$.We then obtain a lemma below.

LEMMA 4.1. *If there exists a* $CAOA(n, K, 2, 2, 0)$ *with generating vector* $\mathbf{d}^* \in \mathcal{D}_n$, *then* $\mathbf{d}^*$ *is universally optimal for estimating* $\mathbf{h}$ *in Model* (4.2) *and inference on* $\theta = (\theta_1, \ldots, \theta_K)^T$ *in Model* (4.3).

Table 2 gives the values of K of all known $CAOA(n, K, 2, 2, 0)$. The first row is the size of n, the second row is the H-sequence with adding one zero

TABLE 2
A list of $CAOA(n, K, 2, 2, 0)$ *when* $n \leq 200$

n	4	8	12	16	20	24	28	32	36	40	44	48	52
H_1	2	2	3	na	5	5	na	5	na	na	6	5	na
CPHM	2	3	5	7	7	9	9	12	14	17	16	17	20
CAOA	2	3	5	7	7	9	9	12	14	17	16	17	20
n	56	60	64	68	72	76	80	84	88		$\cdots$		200
H_1	na	6	na	6	6	na	6	7	na		$\cdots$		6
CPHM	20	7	12	na	na	na	na	na	na		$\cdots$		na
CAOA	23	14	14	14	14	14	17	13	16		$\cdots$		17

to H-sequence, the third row is the CPHMs in [10] and the fourth row is our CAOAs. The value of K is maximum when $4 \le n \le 52$, and it is a lower bound when $n \ge 56$. These designs are universally optimal for estimating the contrast $h_{1,i} - h_{2,i}$. If the symbols of a $CAOA(n, K, 2, 2, 0)$ is 0 and 1, then it is an optimal design for estimating the HRF of one stimulus type. Although known results are limited to small dimensions, they are useful to obtain a design with large n and certain k via Lemma 3.8. For instance, if a $n = 132$ time points experiment is required and each stimulus appears every 4 seconds, then it is a 9-minute fMRI experiment. The extended H-sequence of length $n = 132$ can accommodate a typical 32-second HRF [i.e., $K = (32/4) + 1 = 9$], and it is a $CAOA(132, 9, 2, 2, 0)$. However, we can provide a $CAOA(132, 16, 2, 2, 0)$ with generating vector $\mathbf{d} = (\mathbf{d}', \mathbf{d}', \mathbf{d}')$ by Lemma 3.8, where $\mathbf{d}'$ is the generating vector of the $CAOA(44, 16, 2, 2, 0)$ in Table 1. Instead of using a design with $n = 132$ from the supplement of [18], our design can accommodate a HRF with a longer duration, and thus is suggested to be used.

Furthermore, we prove that there exists a $CAOA(4u, 14, 2, 2, 0)$ [i.e., circulant $OA(4u, 14, 2, 2)$] when $u \ge 9$. This is the first result that guarantees the existence of circulant OAs for all $n \equiv 0 \pmod 4$. For consistency, we will prove it in next section. In our supplementary material [22], we provide a list of universally optimal fMRI designs of length $n \le 600$ that accommodate a typical 32-second (i.e., $K \le 9$) HRF; a nontypical HRF with a long duration is allowed for many n.

4.2. $n \equiv 1, 3 \pmod 4$. Define the information matrices for all the parameters and let $\mathbf{h}$ in Model (4.1) be $\mathbf{M}(\mathbf{X}_d) = \mathbf{X}_d^T \mathbf{X}_d$ and $\mathbf{M}_b(\mathbf{X}_d) = \mathbf{X}_d^T(\mathbf{I}_n - n^{-1}\mathbf{J}_n)\mathbf{X}_d$, respectively. Let $\mathbf{D} = (d_{i,j})_{n \times K}$ be the transpose of a $CAOA(n, K, 2, 2, 1)$ where $n \equiv 1, 3 \pmod 4$ and $\mathbf{D}^* = 2\mathbf{D} - \mathbf{J}_{n \times K}$. By Corollary 3.3, there exists an $(n, K, s, \boldsymbol{\Lambda})$-CDS with $B(\boldsymbol{\Lambda}) = 1$. Suppose that $\boldsymbol{\Lambda} = (\lambda^{i,j})$; it is easy to verify that $\lambda^{0,1} = \lambda^{1,0}$ via Proposition 3.5(b), so $|\lambda^{0,0} - \lambda^{1,1}| = 1$. Without loss of generality, we assume $\lambda^{1,1} = \lambda^{0,0} + 1$. Since $B(\boldsymbol{\Lambda}) = 1$, $\lambda^{0,0} = \lambda^{1,0} = \lambda^{1,0} = \lfloor n/4 \rfloor$ and $\lambda^{1,1} = \lceil n/4 \rceil$ when $n \equiv 1 \pmod 4$; $\lambda^{0,0} = \lfloor n/4 \rfloor$ and $\lambda^{1,1} = \lambda^{1,0} = \lambda^{1,0} = \lceil n/4 \rceil$ when $n \equiv 3 \pmod 4$.

Any two columns of $\mathbf{D}^*$ contains $\lambda^{i,j}$ pairs (i, j) as row vectors, so their dot product is equal to 1. It implies that $\mathbf{M}(\mathbf{D}^*) = (n - 1)\mathbf{I}_K + \mathbf{J}_K$ when $n \equiv 1 \pmod 4$ and $\mathbf{M}(\mathbf{D}^*) = (n + 1)\mathbf{I}_K - \mathbf{J}_K$ when $n \equiv 3 \pmod 4$, respectively. Let $\mathbf{D}^T \mathbf{J}_n \mathbf{D} = (m_{i,j})_{K \times K}$, then $m_{i,j} = (\sum_{k=1}^K d_{i,k})(\sum_{k=1}^K d_{j,k})$ can be derived. We have $(\mathbf{D}^*)^T \mathbf{J}_n \mathbf{D}^* = \mathbf{J}_K$. Then $\mathbf{M}_b(\mathbf{D}^*) = (n - 1)[\mathbf{I}_K + n^{-1}\mathbf{J}_K]$ when $n \equiv 1 \pmod 4$ and $\mathbf{M}_b(\mathbf{D}^*) = (n + 1)[\mathbf{I}_K - n^{-1}\mathbf{J}_K]$ when $n \equiv 3 \pmod 4$.

According to Theorem 2.1 and Lemma 2.5 of Cheng and Kao [8], the optimality of our CAOAs can be rewritten as follows.

LEMMA 4.2. *Let* $\mathbf{d}$ *be the generating vector of a* $CAOA(n, K, 2, 2, 1)$:

(a) *If* $n \equiv 1 \pmod 4$, *then* $\mathbf{d}$ *is optimal for estimating* $\mathbf{h}$ *of Model* (4.1) *for all type* 1 *criteria*.

TABLE 3
A list of $CAOA(4u+1, K, 2, 2, 1)$ *when* $4u+1 < 50$

n	5	9	13	17	21	25	29	33	37	41	45	49
H_2	na	na	na	na	5	5	na	5	na	na	6	5
CAOA	2	3	5	6	8	9	11	12	13	14	16	17

(b) *If* $n \equiv 3 \pmod 4$, *then there exists an* $N_0(K, p_0)$ *such that whenever* $n \geq N_0(K, p_0)$, $\mathbf{d}$ *is* Φ_p*-optimal for estimating* $\mathbf{h}$ *of Model* (4.1) *for any* $p \in [0, p_0]$.

Furthermore, Cheng and Kao [8] developed a theory to guide the selection of $N_0(K, 1)$ (i.e., $p_0 = 1$) such that $\mathbf{d}$ in Lemma 4.2(b) is A- and D-optimal for estimating the HRF when $n \geq N_0(K, 1)$. Moreover, the optimality of CAOAs for comparing two HRFs is rewritten as follows.

LEMMA 4.3. *Let* $\mathbf{D}^* = \mathbf{D} + \mathbf{J}_{n\times K}$ *where* $\mathbf{D}$ *is the transpose of a* $CAOA(n, K, 2, 2, 1)$ *and* $\mathbf{d}$ *be the generating vector of* $\mathbf{D}^*$:

(a) *If* $n \equiv 1 \pmod 4$, *then* $\mathbf{d}$ *is optimal for estimating* θ *of Model* (4.3) *for all type* 1 *criteria*.

(b) *If* $n \equiv 3 \pmod 4$, *then* $\mathbf{d}$ *is* A*-optimal and* Φ_p*-optimal for estimating* θ *of Model* (4.3) *for all* $p \in [0, 1]$ *when* $n \geq N_0(K, 1)$.

Prior to Lin et al. [21], the extended H-sequence is the only systematic way to construct $CAOA(4u+1, K, 2, 2, 1)$, but the value of K is small. Using the DVA algorithm proposed in Lin et al. [21], we successfully found many $CAOA(4u+1, K, 2, 2, 1)$ and the value of K is larger than that of the extended H-sequence. Table 3 is a list of known $CAOA(4u+1, K, 2, 2, 1)$ when $4u+1 < 50$. The second row is the results of the extended H-sequence obtained by adding two 0's to a H-sequence in [8]. The third row is our $CAOA(4u+1, K, 2, 2, 1)$. The value of K is maximum when $4u < 30$ by a complete search. Although the maximum value of K is still uncertain when $4u \geq 30$, it is about $(4u+1)/3$ via our empirical study. Developing systematic constructions for $CAOA(4u+1, K, 2, 2, 1)$ with maximum K is a topic of future research.

According to Lemma 3.7, there exists a square matrix $CAOA(2^m - 1, 2^m - 1, 2, 2, 1)$ for the case of $n = 4u+3$. It is interesting that a $(4u+3, 4u+3, 2, \mathbf{\Lambda})$-CDS with $B(\mathbf{\Lambda}) = 1$ can be obtained by a cyclic $(4u+3, 2u+1, u)$ difference set and its complement. A (n, k', λ) difference set is known to be relevant to a (n, b', r', k', λ) symmetric balanced incomplete block design if $n = b'$ and $r' = k'$. Without loss of generality, assume that 0 appears $\lfloor n/2 \rfloor$ number of times in each row, then $k' = \lfloor n/2 \rfloor$. Since $\lambda(n-1) = r'(k'-1)$, $\lambda = u$ is an integer only if $n = 4u+3$. This implies that a $CAOA(n, n, 2, 2, 1)$ exists only if $n \equiv 3 \pmod 4$, and it can be obtained by a cyclic $(4u+3, 2u+1, u)$ difference set. In fact, such

TABLE 4
A list of $CAOA(4u+3, K, 2, 2, 1)$ *when* $4u+3 < 50$

n	7	11	15	19	23	27	31	35	39	43	47
H-seq	7	11	15	19	23	na	31	35	na	43	47
CAOA	7	11	15	19	23	12	31	35	15	43	47
$N_0(K,1)$	na	5	6	8	9	11	12	13	15	16	18

CAOA can be easily generated by the Paley, Singer or twin prime power difference sets ([29, 35, 36]). They are summarized in the corollary below.

COROLLARY 4.4. *A* $CAOA(n, n, 2, 2, 1)$ *exists if*:

(1) $n \equiv 3 \pmod 4$ *and n is a prime*.
(2) $n = p(p+2)$ *where* p *and* $p+2$ *are both odd prime*.
(3) $n = 2^m - 1$ *where* $m \geq 2$.

Even though Corollary 4.4 is powerful, there are still many CAOAs of $n \equiv 3 \pmod 4$ whose K does not attain n, such as 27 and 39. We find both of them which are all Φ_p-optimal for any $p \in [0, 1]$. Table 4 provides a list of known $CAOA(4u+3, K, 2, 2, 1)$ when $4u+3 < 50$, where the second and third rows are the results of the H-sequence and ours, respectively. The fourth row is the maximal value of K such that $n \leq N_0(K, 1)$.

4.3. $n \equiv 2 \pmod 4$. Comparing with the optimal fMRI designs with $n \equiv 0, 1, 3 \pmod 4$, those with $n \equiv 2 \pmod 4$ are not simple to construct. Based on the discussion in [4, 7, 23], a design $\mathbf{D}$ is optimal if $M(\mathbf{D})$ is a 2 by 2 block matrix with two diagonal submatrices $(n-2)\mathbf{I}_{K/2} + 2\mathbf{J}_{K/2}$ and zero otherwise. Since fMRI designs are circulant, it is impossible to get a circulant design whose information matrix is a block matrix. Recently, Cheng et al. [9] proved that a $CAOA(4u+2, K, 2, 2, 1)$ is Φ_p-optimal if its information matrix is $(n-2)\mathbf{I}_K + 2\mathbf{J}_K$. Such CAOA exists in our empirical study and they outperform other CAOAs when $n \equiv 2 \pmod 4$.

Although such design is known to be optimal when the off-diagonal entries of its information matrix is +2, but the value of K is usually small (see T_1 in Table 5).

TABLE 5
A list of $CAOA(4u+2, K, 2, 2, 1)$ *when* $4u+2 \leq 50$

n	6	10	14	18	22	26	30	34	38	42	46	50
T_1	2	3	4	6	7	9	10	11	12	13	14	16
T_2	3	5	7	8	11	13	11	17	19	18	23	21
$D_{\text{eff}}(\%)$	93	89	89	94	89	90	98	91	91	97	92	97

In the light of the pattern of near-Hadamard matrices [23], we consider another design, the off-diagonal entries of its information matrix is -2. Let $\mathbf{D}$ be the transpose of a $CAOA(n, K, 2, 2, 1)$, $\mathbf{D}$ is called Type$_1$ if $\mathbf{M}(\mathbf{D}) = (n-2)\mathbf{I}_K + 2\mathbf{J}_K$ and Type$_2$ if $\mathbf{M}(\mathbf{D}) = (n+2)\mathbf{I}_K - 2\mathbf{J}_K$; we denote them T_1-$CAOA(n, K, 2, 2, 1)$ and T_2-$CAOA(n, K, 2, 2, 1)$, respectively. T_2-$CAOA(n, K, 2, 2, 1)$ always has a larger value of K than T_1-$CAOA(n, K, 2, 2, 1)$ in our experience. In particular, we find a series of T_2-$CAOA(n, K, 2, 2, 1)$ whose K attains the upper bound $n/2$. The following key lemma helps us to construct T_2-CAOAs.

LEMMA 4.5. *Let $l > 1$ be an integer. If D is a $(n, k; \lambda_1, \dots, \lambda_{n-1})$ GDS, then $\bigcup_{i=0}^{l-1}(D + in)$ is a $(ln, lk; \lambda'_1, \dots, \lambda'_{ln-1})$ GDS where $\lambda'_{r+in} = l\lambda_r$ and $\lambda'_{jn} = lk$ for $i = 0, 1, \dots, l-1$, $j = 1, 2, \dots, l-1$.*

The above lemma is a simple method to get a larger GDS from a small one. The following theorem suggests a general class of T_2-$CAOA$ whose $K = n/2$ for all odd prime n.

THEOREM 4.6. *There exists a T_2-$CAOA(2n, n, 2, 2, 1)$ for all odd prime n.*

To quantify the D-optimality of a design $\mathbf{D}$, we adopt the D-efficiency criterion of [15, 30]:

$$d_e(\mathbf{D}, \mathbf{D}_o) = \left(\frac{|\mathbf{M}(\mathbf{D})|}{|\mathbf{M}(\mathbf{D}_o)|}\right)^{1/K},$$

where the design $\mathbf{D}_o$ is theoretical optimal, $|\mathbf{X}|$ is the determinant of a matrix $\mathbf{X}$ and K is the number of terms in the model that consists of all main effects. We compare the D-optimality between T_2 and T_1, so $\mathbf{D}_o$ is the transpose of T_1-$CAOA(n, K, 2, 2, 1)$. Hence, according to [9], the D-efficiency of T_2 is formulated by

$$\left(\frac{n-2K+2}{n+2K-2}\right)^{1/K}\left(\frac{n+2}{n-2}\right)^{(K-1)/K}.$$

Table 5 shows our first-handed results. The second and third rows correspond to T_1- and T_2-$CAOA(n, K, 2, 2, 1)$, respectively, and the fourth row is the D-efficiency of T_2.

It is noteworthy that given a fixed n, the D-efficiency decreases when K is increasing. Since the upper bound of a T_2-$CAOA(n, K, 2, 2, 1)$ is $K = n/2$ where $n \equiv 2 \pmod 4$, T_2 designs obtained by the above theorem guarantee at least 90% D-efficiency when $n \geq 26$. Furthermore, the D-efficiency is easily enhanced by deleting some rows of T_2. For instance, we consider the 9-minute fMRI experiment discussed in Section 4.2, where a $CAOA(132, 16, 2, 2, 0)$ is suggested. If a nontypical 120-second HRF is required, then a T_2-$CAOA(134, 31, 2, 2, 1)$, whose generating vector $\mathbf{d}$ is the same with T_2-$CAOA(134, 67, 2, 2, 1)$, is suggested. In fact, such design $\mathbf{d}$ can accommodate a HRF with a long duration up to $K = 58$ and have 99% D-efficiency.

TABLE 6
Unfinished $\mathbf{\Lambda}$

1				5
	2			5
		1		5
			1	4
5	5	5	4	

5. CAOAs with three- and four-levels. The m-sequences are traditionally used in ER-fMRI experiments [1], and the efficiency of a fMRI design is always an important issue for researchers. During the last few years, many reports indicated that the m-sequences may be efficient but not optimal [25, 26, 28]. Recently, Kao [16] proved that an extended m-sequence is D-optimal but a binary extended m-sequence is universally optimal [8]. However, these designs always have a large length but accommodate a HRF with a short duration when $Q > 2$. For example, a ternary extended m-sequence of a length 27, 81, 243 and 729 accommodates 3, 4, 5 and 6 duration time points, respectively. If a 24-second HRF is of interest and the stimulus occurs every 4 seconds, then an experimental subject needs to accept a 50-minute fMRI experiment, which is an unacceptably long experiment for a typical subject. Hence, it is an open question on how to construct optimal designs with a length shorter than the extended m-sequences for $Q \geq 3$ [21]. We unmask a possible solution via finding CAOAs for fMRI experiments with $Q = 3$ and 4 in this section.

The existence of CAOAs is always highly interesting and essential. For two-level CAOAs with a frequency matrix $\mathbf{\Lambda} = (\lambda^{i,j})_{i,j \in Z_2}$, it is known that $\lambda^{1,0} = \lambda^{0,1}$. Thus the frequency matrix is unique when $\lambda^{0,0}$ or $\lambda^{1,1}$ is determined. Hence, the GDS method is used to efficiently find CPHMs. When the level is more than two, $\mathbf{\Lambda}$ is usually not unique even if all $\lambda^{i,i}$s are determined. Furthermore, CAOAs usually do not exist for the arbitrary frequency matrix. For example, a $CAOA(19, K, 4, 2, 1)$ with $\mathbf{\Lambda} = (\lambda^{i,j})_{i,j \in Z_4}$ does not exist when $\lambda^{0,0}, \lambda^{0,2}, \lambda^{2,0} = 2$ and 1 otherwise. By Proposition 3.5, $\mathbf{\Lambda}$ is relevant to the cardinality of each part in a partition $V = \{V_i | i \in Z_n\}$. It is obvious that $|V_i|$ equals to the ith column and the ith row sum of $\mathbf{\Lambda}$. Therefore, we propose a *square principle* for the selection of the frequency matrix. The square principle is illustrated as the following example.

EXAMPLE 5.1. We demonstrate the choice of the frequency matrix of a $CAOA(19, K, 4, 2, 1)$. Suppose that each symbol except 3 occurs five times and the symbol 3 occurs four times in each row. Thus, we consider a partition $V = \{V_0, V_1, V_2, V_3\}$ with $|V_3| = 4$ and $|V_i| = 5$ for $i = 0, 1, 2$. If the frequency $\lambda^{1,1} = 2$ and $\lambda^{i,i} = 1$ are of interest, then we first write down $\mathbf{\Lambda}$ (see Table 6). The numbers on the right-hand side and the bottom are the cardinality of V_i, based on the principle that the sum of the ith row and the ith column equals to $|V_i|$ for all i. If

TABLE 7
Finished $\mathbf{\Lambda}$

1	1	2	1	5
1	2	1	1	5
2	1	1	1	5
1	1	1	1	4
5	5	5	4	

$B(\mathbf{\Lambda}) = 1$, then the solution is unique in this example (see Table 7). Moreover, we exploit $(19, 3, 4, \mathbf{\Lambda})$-CDS and find a $CAOA(19, 3, 4, 2, 1)$ that possesses a maximum number of factors among all possible combinations. The generating vector of $CAOA(19, 3, 4, 2, 1)$ is listed in Table 9.

The square principle only fits to find CAOAs of strength two in this paper, but it can be extended when the strength is more than two. Although this principle treats as a simple criterion to determine the frequency matrix of CAOAs, the choices of the frequency matrix are not unique. For instance, suppose that $\lambda^{i,i} = 2$ for $i = 0, 1, 2$ and $\lambda^{i,j} = 1$ otherwise, then $\mathbf{\Lambda}$ is another choice of the frequency matrix of a $CAOA(19, K, 4, 2, 1)$. However, its maximum value of K is 2, not 3. In our experience, an equitable partition is always better than an arbitrary partition. Thus, if n cannot be equally partitioned into all frequencies, we suggest to consider the increase of the pair $\lambda^{i,j}$ and $\lambda^{j,i}$ before the increase of $\lambda^{i,i}$. However, this is just a rule-of-thumb for an efficient search and it is without theoretical justification.

From these empirical criteria, we find all $CAOA(n, K, s, 2, b)$ that possess the maximum values of K when $n \le 32$, $s = 3$ and $n \le 35$, $s = 4$. Due to the criterion constraint, the bandwidth is $b = 2$ when $s = 3$ and $n \equiv 1 \pmod 9$. Furthermore, the lower bounds are also provided when $33 \le n \le 45$, $s = 3$. This implies that K will increase as n increases. The generating vectors of these CAOAs with bandwidth 0, 1 and 2 are listed in Tables 8 and 9.

We then focus on the construction of a D-optimal $CAOA(n, K, 3, 2, 0)$ for estimating $\mathbf{h}$ in Model (2.1). If a $CAOA(n, K, 3, 2, 0)$ exists, then n must be the multiple of 3^2. Table 8 shows the existence of $CAOA(9u, K, 3, 2, 0)$ when $u = 1, \ldots, 4$, and the value of K is confirmed via a comprehensive search. Similar to $CAOA(n, K, 2, 2, 0)$ in Section 4.1, the value of K increases with an increase of n for $CAOA(n, K, 3, 2, 0)$. However, the difficulty of searching CAOAs of large n increases. For $Q = 2$, Kao [18] compiled a table that provided many optimal designs for fMRI experiments when $n \le 600$. The designs only exist whenever $n - 1$ is a prime, because the construction is based on the extended H-sequences. The value of K is usually small even when n is very large. On the other hand, the extended m-sequence can be constructed systematically when $Q = 3$, but the gap of n is too large. To our best knowledge, there is

TABLE 8
The generating vectors of $CAOA(n, k, 3, 2, b)$ for $8 \leq n \leq 45$

n	***k***	***b***	**Generating vector**
8	4	1	10122021
9	2	0	010211220
10	3	2	0020112122
11	4	1	10200121221
12	3	1	022020111210
13	3	1	0122112002021
14	4	1	00212111201022
15	3	1	012210110212002
16	4	1	0221202210112001
17	4	1	11020122202100121
18	4	0	000102202111012212
19	4	2	2100201120010221212
20	4	1	12022121112201021000
21	5	1	020220111012110212200
22	5	1	0221001212112201102020
23	5	1	11112022001020122100212
24	5	1	010202112201101212002210
25	5	1	0200102220211001212201211
26	13	1	10222001012112011100202122
27	5	0	011021200221222010002011121
28	6	2	0122120102002110020011222121
29	6	1	11221011021212002010001220221
30	6	1	001002111210112012110202002222
31	6	1	0002121101122211022020120012102
32	6	1	00012202210102201202111121102120
33	6	1	120211102202201012100122112000102
34	6	1	2201022212110201212200021011001120
35	6	1	10011102220210212111201012000221220
36	6	0	101101210020002021121220222110011220
37	6	2	0001002201012210221211220200112021211
38	7	1	21011120221211022112010002001021222001
39	7	1	121100121022011121211002022022200010201
40	7	1	2001102021111210221021200002201201012221
41	7	1	22012100112020200200221210100122110211112
42	7	1	001011200112212120200221100102022110121022
43	7	1	0011012010200011201102212121002211210202222
44	7	1	11201220110121102000100212210022220212102011
45	7	0	002101121102011012221220211121000222020120001

no existing method in the literature to construct fMRI designs with $Q = 3$ [i.e., $CAOA(n, K, 3, 2, 0)$] for any $n \equiv 0 \pmod 9$. Here, we propose a new method to construct a $CAOA(9u, 6, 3, 2, 0)$ for all $u \geq 4$, which implies that the lower bound of K is 6 when $n \geq 36$.

TABLE 9
The generating vectors of CAOA$(n, k, 4, 2, b)$ *for* $12 \leq n \leq 35$

n	*k*	*b*	**Generating vector**
12	2	1	032312130102
13	2	1	0323312130201
14	2	1	03223312130201
15	5	1	013110323302122
16	2	0	0132022331211003
17	3	2	10133230110221203
18	3	1	310023032202011213
19	3	1	1112100133020220323
20	3	1	20020331011321231302
21	3	1	202210113230020331131
22	3	1	2310120213311030220013
23	3	1	32022200313012311010213
24	3	1	312011332100230102232031
25	3	1	1210113223133203110020230
26	3	1	22010012103113213231202330
27	3	1	321201302103032002310111223
28	4	1	0122031321130232100310123302
29	4	1	03213122302320112013331002103
30	4	1	032130333120231011230222132001
31	4	1	1013303021110223220012033132123
32	4	0	00212113103311220013030223233201
33	5	2	223101301022132001110312303332120
34	5	1	0012010311103213102333230212220130
35	5	1	31323023300103112020220032211101213

LEMMA 5.2. *Let* $\mathbf{x} = (x_1, \ldots, x_n)$ *and* $\mathbf{y} = (y_1, \ldots, y_m)$ *be the generating vectors of circulant matrices* $\mathbf{X}_{K\times n}$ *and* $\mathbf{Y}_{K\times m}$, *respectively. If* $x_{n-r} = y_{m-r}$ *for all* $r = 0, \ldots, K-2$, *then* $\mathbf{D}_{K\times(n+m)} = (\mathbf{X}|\mathbf{Y})$ *is a circulant matrix.*

The above lemma provides a simple way to build up a large circulant matrix from some smaller circulant matrices.

THEOREM 5.3. *If* $n \equiv 0 \pmod 9$ *and* $n \geq 36$, *then there exists a CAOA*$(n, 6, 3, 2, 0)$. *Furthermore, it is D-optimal for estimating* $\mathbf{h}$ *in Model* (4.1).

Similar to the existence of *CAOA*$(n, 6, 3, 2, 0)$, we also prove the existence of two-level CAOAs. According to Lemma 5.2, Table 10 and the construction in Theorem 5.3, we have the following results.

THEOREM 5.4. *Let* $n \equiv 0 \pmod 4$ *and* $n \geq 36$, *there is a CAOA*$(n, 14, 2, 2, 0)$.

TABLE 10

A generating vector pair for constructing CAOA$(n, 14, 2, 2, 0)$

(n_1, n_2)	Generating vector of $CAOA(n_1, 14, 2, 2, 0)$ Generating vector of $CAOA(n_2, 14, 2, 2, 0)$
(36, 40)	1110110011100 0101000000101111001 0110 0001001100 100000111011110101011110010110
(36, 44)	1110001010000 00101111001011011101100 0111111001 0100011000101000001011011011101100
(36, 48)	1011000100000 01111010011101110101100 0111110010 110101000000010011011000111011101011 00
(36, 52)	1010000001010 00111001101110110100111 0110111110 010100000110011000101010000011110110 100111
(36, 56)	1111101000011 01001000100110001110101 0110100000 110010000001111101110111001011010011 0001110101
(36, 60)	1111010000110 10010001001100011101011 0111000011 110111011001010000010100100000110111 01100011101011
(36, 64)	0000101111110 10111000110010001001011 1111101110 011010011110000010111010010101000011 00011001000100 1011
(36, 68)	0111111010111 00011001000100101100001 1110011101 111100101010110010000010111001101111 01000010001001 01100001

6. Conclusion and discussion. Research on fMRI experimental designs that improve the precision of statistical analyses is a new and wide-open study area. The m-sequences and its variations have been popularly used in fMRI experiments nowadays. Under the model assumptions proposed by Kao [17], H-sequences and extended H-sequences have recently been introduced for fMRI experiments. In order to render precise statistical inference on brain functions, the optimality of fMRI experimental designs is diffusely studied in [8, 16, 18], but there is no single and unified method to construct all of them. This paper aims at proposing a unified method to construct various fMRI designs in a systematic way.

We introduce CAOAs for fMRI experiments, and we propose a new difference method CDS to construct CAOAs that are listed in the tables. The maximum value of K is mainly of interest on the estimation of a HRF and the comparison between the HRFs in an ER-fMRI experiment. Hence, we provide properties and the upper bound of K, and this verifies the existence of CAOAs. Our CAOAs are highly efficient such that they attain the upper bound of K, and their size and near-orthogonality are the same as m-sequences. A simple doubling method is introduced to construct CAOAs with large n via some small known CAOAs.

Following the selection guide of optimal experimental design for fMRI in [8], we effectively find a series of CAOAs. When $n \equiv 0 \pmod 4$, our CAOAs are proved to be universally optimal for estimating a HRF and the contrast of two HRFs. We compare our designs with those found in [8, 10, 18], showing that our results are complete and guarantee the value $K \geq 14$ when $n \geq 36$. In addition,

TABLE 11
T_3-$CAOA(n, K, 2, 2, 0)$ *for all* $6 \le n \le 34$

n	K	D_{eff}	Generating vector
6	3	92.83%	001011
10	5	100%	0001011101
14	7	99.38%	01010000110111
18	9	98.96%	001110101110100001
22	11	99.12%	0000011110111011001010
26	13	100%	00000101011001111101010011
30	15	99.94%	000001010101100111111001010011
34	17	99.32%	0000100010100011011110111001001111

we provide in our supplementary materials [22] a list of universally optimal fMRI designs of length $n \le 600$ that accommodates a typical 32-second ($K \le 9$) HRF. These new designs accommodate a typical HRF of at least 32-seconds. We also show that for $n \le 50$, our CAOAs are optimal for all type 1 criteria when $n \equiv 1 \pmod 4$ and Φ_p-optimal when $n \equiv 3 \pmod 4$. In addition, H-sequences and extended H-sequences are special cases of CAOAs, and our designs possess larger K than extended H-sequences in general.

The existence of optimal CAOAs is still under investigation for $n \equiv 2 \pmod 4$, but we suggest two types of CAOAs for fMRI experiments. The T_1-CAOAs are shown to be Φ_p-optimal in [9] but they have small K, and T_2-CAOAs have large K but only nearly-orthogonal. We provide each type of CAOAs for $n \le 50$, and propose a construction for T_2-CAOAs attaining the theoretical upper bound of K, and its D-efficiency is at least 90% when $n \ge 26$. Besides T_1- and T_2-CAOAs, a new class of CAOAs, namely T_3-CAOAs, is also under investigation. T_3-CAOAs are found to have large, if not maximum, K and high D-efficiency. Unlike T_1- and T_2-CAOAs with only $+2$ or -2 in the off-diagonal entries of their information matrix, respectively, T_3-CAOAs possess mixed combinations of ± 2 in the off-diagonal entries. Table 11 provides some T_3-CAOAs for $6 \le n \le 34$, which $K = n/2$ like T_2-CAOAs. However, they are D-optimal in $n = 10$ and $n = 26$ like T_1-CAOAs. Notice that this class is found via computer enumeration. Since the purpose of this paper is to provide a systematic construction for CAOAs, we do not emphasize T_3-CAOAs as a main result.

When the number of stimulus types is more than two, conventional wisdom suggests to use m-sequences and extended m-sequences in fMRI experiments; however, the gap of their length is too large to implement, while their optimality is still unknown. Although the extended m-sequences are proven to be D-optimal, the value of K is usually small. Therefore, we compile a table of CAOAs of three- and four-level, where most of these designs have larger K even though n is small with respect to the extended m-sequences. Moreover, we prove the existence of $CAOA(9u, 6, 3, 2, 0)$ when $u \ge 4$, which leads to the existence of circulant

$OA(9u, 6, 3, 2)$, a class of D-optimal designs for estimating the HRFs. To our best knowledge, there is no construction that can obtain circulant OAs, so our construction is new and simple.

APPENDIX A: THE CRITERION OF OPTIMALITY

The optimality of fMRI experiments were discussed by Cheng and Kao [8]. Here, we briefly introduce some criteria used in this paper; for the details, please refer to [8].

The information matrix of all parameters and $\mathbf{h}$ in Model (4.2) are $\mathbf{M}(\mathbf{X}_d) = \mathbf{X}_d^T\mathbf{X}_d$ and $\mathbf{M}_b(\mathbf{X}_d) = \mathbf{X}_d^T(\mathbf{I}_n - n^{-1}\mathbf{J}_n)\mathbf{X}_d$, respectively.

DEFINITION A.1. A design $\mathbf{d}$ is said to be universally optimal over a design class if it minimizes $\Phi\{\mathbf{M}_b(\mathbf{X}_d)\}$ for all convex functions Φ such that (i) $\Phi(c\mathbf{M})$ is nonincreasing in $c > 0$, and (ii) $\Phi(\mathbf{PMP}^T) = \Phi(\mathbf{M})$ for any $\mathbf{M}$ and any orthogonal matrix $\mathbf{P}$.

DEFINITION A.2. A design $\mathbf{d}$ is said to be optimal over a design class with respect to all the type 1 criteria if it minimizes $\Phi_{(f)}\{\mathbf{M}_b(\mathbf{X}_d)\} = \sum_{i=1}^K f(\lambda_i(\mathbf{M}_b(\mathbf{X}_d)))$ for any real-valued function f defined on $[0, \infty)$ such that (i) f is continuously differentiable in $(0, \infty)$ with $f' < 0$, $f'' > 0$, and $f''' < 0$, and (ii) $\lim_{x\to 0^+} f(x) = f(0) = \infty$. Here, $\lambda_i(\mathbf{M}_b(\mathbf{X}_d))$ is the ith greatest eigenvalue of $\mathbf{M}_b(\mathbf{X}_d)$, $i = 1, \dots, K$.

DEFINITION A.3. A design $\mathbf{d}$ is said to be Φ_p-optimal over a design class for a given $p \ge 0$ if it minimizes

$$\Phi\{\mathbf{M}_b(\mathbf{X}_d)\} = \begin{cases} |\mathbf{M}_b(\mathbf{X}_d)|^{1/K} & \text{for } p = 0; \\ [\operatorname{tr}\{\mathbf{M}_b^{-p}(\mathbf{X}_d)\}/K]^{1/p} & \text{for } p \in (0, \infty); \\ \Lambda_1(\mathbf{M}_b^{-1}(\mathbf{X}_d)) & \text{when } p = \infty, \end{cases}$$

where $\Lambda_1(\mathbf{M}_b^{-1}(\mathbf{X}_d))$ is the largest eigenvalue of $\mathbf{M}_b^{-1}(\mathbf{X}_d)$.

APPENDIX B: PROOFS

PROOF OF THEOREM 3.2. Given V is a $(n, k, s, \mathbf{\Lambda})$-CDS, we assume $V = \{V_0, \dots, V_{s-1}\}$ is a partition of Z_n. Let $\mathbf{A} = (a_{i',j'})_{s\times s}$ be the incidence matrix of V, $\mathbf{A}_{(i,j)}$ be an $2 \times n$ subarray that consists of the ith and jth rows of $\mathbf{A}$ and $1 \le i < j \le s$. Suppose that each pair (x, y) appears exactly $\lambda(x, y)$ times in $\mathbf{A}_{(i,j)}$ as a column, and $\lambda_{j-i}^{x,y}$ is the frequency of the element $(j - i)$ in $DFS_n(V_x, V_y)$.

We claim that $\lambda_{j-i}^{x,y} = \lambda(x, y)$ for all $x, y \in Z_s$. Assume $\lambda(x, y) \ne 0$. Since each pair (x, y) appears exactly $\lambda(x, y)$ times, there exists $1 \le c_1, c_2, \dots, c_{\lambda(x,y)} \le n$

such that $a_{i,c_l} = x$ and $a_{j,c_l} = y$ where $l = 1, 2, \dots, \lambda(x, y)$. From Definition 3.1, $c_l \in (V_x + (i-1)) \cap (V_y + (j-1))$. Thus, $c_l - (i-1) \in V_x$ and $c_l - (j-1) \in V_y$. Since $[c_l - (i-1)] - [c_l - (j-1)] = j - i$ for all $l = 1, 2, \dots, \lambda(x, y)$, the element $(j-i)$ appears totally $\lambda(x, y)$ number of times in $DFS_n(V_x, V_y)$. Hence, we have $\lambda_{j-i}^{x,y} \geq \lambda(x, y)$.

Now, let $\alpha \in V_x$ and $\beta \in V_y$ such that $\alpha - \beta = j - i$. It follows that $\alpha + (i-1) \in V_x + (i-1)$, $\beta + (j-1) \in V_y + (j-1)$, and $\alpha + (i-1) = \beta + (j-1)$. Therefore, $\alpha + (i-1) \in V_x + (i-1)$ and $(V_y + (j-1))$. Since $a_{i,\alpha+(i-1)} = x$ and $a_{j,\alpha+(i-1)} = y$, $\lambda(x, y) \geq \lambda_{j-i}^{x,y}$. This completes the proof that $\lambda_{j-i}^{x,y} = \lambda(x, y)$. Similarly, the equality holds when $\lambda(x, y) = 0$. $\square$

PROOF OF PROPOSITION 3.5. (a) By definition, $\lambda_r^{i,j} = |\{x - y \equiv r \pmod{n} : x \in V_i, y \in V_j\}|$. Then $x - y \equiv r \pmod{n}$ implies $y - x = -(x - y) \equiv -r \equiv n - r \pmod{n}$ for all $r \in Z_n \setminus \{0\}$. Hence, $\lambda_r^{i,j} = \lambda_{n-r}^{j,i}$.

(b) Since V is a partition of Z_n, each element in Z_n contained in exactly one subset $V_i \in V$. For each element $x \in V_i$, there is exactly one element $y \in Z_n \setminus \{x\}$ such that $x - y \equiv r \pmod{n}$ where $r \in Z_n \setminus \{0\}$ and i is fixed. Therefore, $\sum_{j=0}^{s-1} \lambda_r^{i,j} = |V_i|$ for any fixed i. By (a), $\sum_{i=0}^{s-1} \lambda_r^{i,j} = |V_j|$ for any fixed j.

(c) From (b), it is clear that $\sum_{i=0}^{s-1} \sum_{j=0}^{s-1} \lambda_r^{i,j} = \sum_{i=0}^{s-1} |V_i| = n$. $\square$

PROOF OF LEMMA 3.8. Let $\mathbf{A}$ be a $CAOA(n, k, s, t, b)$ and $\mathbf{D} = (\mathbf{A}| \cdots |\mathbf{A})$ be the composite of l $\mathbf{A}$s. Then $\mathbf{D}$ is obviously a $k \times ln$ circulant matrix. Assume that $\Lambda = (\lambda^{i,j})$ is the frequency matrix of $\mathbf{A}$ such that $B(\mathbf{\Lambda}) = b$. Evidently, $l\Lambda = (l\lambda^{i,j})$ is a frequency matrix of $\mathbf{A}$, because each pair (i, j) occurs totally $l\lambda^{i,j}$ times in any $s \times n$ submatrix of $\mathbf{D}$. Trivially, $B(l\mathbf{\Lambda}) = lb$, so $\mathbf{D}$ is a $CAOA(ln, k, s, t, lb)$. $\square$

PROOF OF LEMMA 4.1. Let $\mathbf{D} = (d_{i,j})_{n \times K}$ be the transpose of a $CAOA$ $(n, K, 2, 2, 0)$ with symbols 1 and 2, so $\mathbf{d}^* = (d_{1,1}, \dots, d_{n,1})$. Since $\mathbf{D}$ is a circulant matrix, $d_{i-k,1} = d_{i,k+1}$. It implies $n_k^{p,q} = \#\{i | (d_{i,k+1}, d_{i,1}) = (p, q), i = 1, \dots, n\}$, so it counts the occurrence frequency of the pair (p, q) in an $n \times 2$ submatrix that consists of the 1st and kth columns of $\mathbf{D}$. By definition, $n_k^{pq} = n/4$ for $p, q = 1, 2$, $1 \leq k \leq K$. According to Theorem 1 in [18], $\mathbf{d}^*$ is universally optimal for inference on $\theta = (\theta_1, \dots, \theta_K)^T$. Furthermore, by replacing 2 with -1, $\mathbf{D}$ is a circulant orthogonal array with $\mathbf{D}^T\mathbf{D} = n\mathbf{I}_K$. So $\mathbf{d}^*$ is universally optimal for estimating $\mathbf{h}$ in Model (4.2). $\square$

PROOF OF LEMMA 4.5. Since D is a $(n, k; \lambda_1, \dots, \lambda_{n-1})$ GDS, there are λ_r ordered pairs (x, y) such that $x - y \equiv r \pmod{n}$ for each $1 \leq r \leq n-1$, where $x, y \in D$. Each pair (x, y) implies the following two equations hold:

$$\big(x + (u+i)n\big) - (y + un) = in + x - y \equiv in + r \pmod{ln} \quad \text{and}$$
$$\big(x + (u' - l + i)n\big) - (y + u'n) = (i - l)n + r \equiv in + r \pmod{ln},$$

where $u = 0, 1, \dots, l-i-1$, $u' = l-i, l-i+1, \dots, l-1$ and $i = 0, 1, \dots, l-1$. For each pair (x, y) that $x - y \equiv r \pmod{n}$, there exists l pairs (x', y') such that $x' - y' \equiv in + r \pmod{ln}$. This implies $\lambda'_{in+r} = l\lambda_r$. Moreover, each difference $\pm jn$ is obtained by replacing y with x in the above two equations; thus, each element $x \in D$ provides l pairs such that the difference $\pm jn$ appears l times. The difference $\pm jn$ appears lk times, so $\lambda'_{jn} = lk$ for $j = 1, 2, \dots, l-1$. □

PROOF OF THEOREM 4.6. Let D be a collection of quadratic elements of $Z_n \setminus \{n\}$ and $\bar{D}$ be the nonquadratic elements of $Z_n \setminus \{n\}$. For convenience, we consider $n = 4u - 1$ and $n = 4u + 1$ individually. In combinatorial design, it is well known that D and $\bar{D}$ are cyclic $(4u-1, 2u-1, u-1)$ difference sets when $n \equiv 3 \pmod{4}$ is a prime. In addition, D and $\bar{D}$ are $(4u+1, 2u; \lambda_q, \lambda_{q^c})$ GDS where $\lambda_q = u - 1$, $\lambda_{q^c} = u$, $q \in D$ and $q^c \in \bar{D}$ when $n \equiv 1 \pmod{4}$ is a prime. According to Lemma 4.5, if S is a $(n, k; \lambda_1, \dots, \lambda_{n-1})$ GDS then $S \cup (S + n)$ is a $(2n, 2k; \lambda'_1, \dots, \lambda'_{2n-1})$ GDS where $\lambda'_i = 2\lambda_i$ and $\lambda'_n = 2k$ for $i \neq n$:

(i) When $n = 4u - 1$, $D \cup (D + n)$ is a $(8u-2, 4u-2; \lambda_1, \dots, \lambda_{8u-1})$ GDS where $\lambda_{4u-1} = 4u - 2$ and $\lambda_i = 2u - 2$ for all $i \neq 4u - 1$. Now, consider the set $D \cup (D + n) \cup \{n\}$. Since -1 is nonquadratic when $n \equiv 3 \pmod{4}$, $-q \in \bar{D}$. For each $q \in D$, the difference of n and q is either $n - q \in (\bar{D} + n)$ or $n + q \in (D + n)$. Similarly, for each $q + n \in (D + n)$, we have a difference with $q \in D$ and $-q \in \bar{D}$. This implies that each element except n appears once when we take the difference between n and $D \cup (D + n)$. Thus, $D \cup (D + n) \cup \{n\}$ is a $(8u-2, 4u-2; \lambda_1, \dots, \lambda_{8u-1})$ GDS where $\lambda_{4u-1} = 4u-2$ and $\lambda_i = 2u-1$ for all $i \neq 4u-1$. Analogously, $\bar{D} \cup (\bar{D}+n) \cup \{n\}$ is also a $(8u-2, 4u-2; \lambda_1, \dots, \lambda_{8u-1})$ GDS where $\lambda_{4u-1} = 4u-2$ and $\lambda_i = 2u-1$ for all $i \neq 4u-1$. Let $V_0 = D \cup (D + n) \cup \{n\}$ and $V_1 = \bar{D} \cup (\bar{D}+n) \cup \{2n\}$. We focus on the occurrence frequency of the difference r in $DFS(V_i, V_j)$, denoted by $\lambda_r^{i,j}$. By Proposition 3.5(b) and (c), $\lambda_r^{0,1} = \lambda_r^{1,0}$ and $\lambda_r^{0,0} + \lambda_r^{1,1} + \lambda_r^{0,1} + \lambda_r^{1,0} = 8u - 2$ for all r. Therefore, $\lambda_n^{0,1} = \lambda_n^{1,0} = 1$ and $\lambda_r^{0,1} = \lambda_r^{1,0} = 2u$ for $r \neq n$. It follows that $V = \{V_0, V_1\}$ is a $(2n, n, 2, \mathbf{\Lambda})$-CDS where $\mathbf{\Lambda} = (2u)\mathbf{J}_2 - \mathbf{I}_2$. By Corollary 3.3, there exists a T_2-$CAOA(2n, n, 2, 2, 1)$.

(ii) When $n = 4u+1$, $D \cup (D+n)$ is a $(8u+2, 4u; \lambda_1, \dots, \lambda_{8u+1})$ GDS where $\lambda_i = 2u - 2$ for all $i \in D$ and $\lambda_i = 2u$ for all $i \in \bar{D}$. Since -1 is quadratic when $n \equiv 1 \pmod{4}$, $-q \in D$. Similar to the proof (i), it is easy to show that there exists a T_2-$CAOA(2n, n, 2, 2, 1)$. □

PROOF OF LEMMA 5.2. The matrix $\mathbf{D}$ is represented below:

$$\left(\begin{array}{cccc|cccc}
x_1 & \cdots & x_{n-1} & x_n & y_1 & y_2 & \cdots & y_m \\
x_n & \cdots & x_{n-2} & x_{n-1} & y_m & y_1 & \cdots & y_{m-1} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_{n-K+3} & \cdots & x_{n-K+1} & x_{n-K+2} & y_{m-K+3} & y_{m-K+4} & \cdots & y_{m-K+2} \\
x_{n-K+2} & \cdots & x_{n-K} & x_{n-K+1} & y_{m-K+2} & y_{m-K+3} & \cdots & y_{m-K+1}
\end{array}\right)_{K \times (n+m)}.$$

TABLE 12
A generating vector pair for constructing CAOA$(n, 6, 3, 2, 0)$

(n_1, n_2)	Generating vector of $CAOA(n_1, 6, 3, 2, 0)$ Generating vector of $CAOA(n_2, 6, 3, 2, 0)$
(36, 45)	112020002001 210110102211001122202212 221011020112110120010002102022200012111202212
(36, 54)	220222110011 220101101210020002021121 011221010100220200001112011002200121102022212212021121
(36, 63)	022110011222 022121120200020012101101 110200220 212110120002012221020211221201112222 100012000210101101

Since $x_{n-r} = y_{m-r}$ for all $r = 0, \dots, K-2$, then $\mathbf{D}$ is obviously circulant. □

PROOF OF THEOREM 5.3. Suppose that $\mathbf{X}$ and $\mathbf{Y}$ are the transpose of $CAOA(n_1, K, s, 2, 0)$ and $CAOA(n_2, K, s, 2, 0)$, respectively. Hence, $\mathbf{D} = (\mathbf{X}|\mathbf{Y})$ is a $OA(n_1 + n_2, K, s, 2)$. If $\mathbf{D}$ is circulant, then $\mathbf{D}$ is a $CAOA(n_1 + n_2, K, s, 2, 0)$.

Let $n = 9u \geq 36$ where u is a positive integer. When $n = 36, 45, 54$ and 63, the $CAOA(n, 6, 3, 2, 0)$ are listed in Table 12 that are found via a computer search. Thus, there exists a $CAOA(9u, 6, 3, 2, 0)$ when $u = 4, 5, 6, 7$. Notice that any two of them have at least five consecutively identical digits. Let $\mathbf{D}_n$ be the transpose of a $CAOA(n, 6, 3, 2, 0)$. For any $n = 9(4p + q)$, $p \geq 1$ and $q = 0, 1, 2, 3$, we construct a matrix $\mathbf{D}_n = (\mathbf{D}_{36}|\cdots|\mathbf{D}_{36}|\mathbf{D}_{36+9q})$ by combining $p-1$ copies of $\mathbf{D}_{36}$ and one copy of $\mathbf{D}_{36+9q}$ where $q = 0, \dots, 3$. By Lemma 5.2, $\mathbf{D}_n$ is a $CAOA(n, 6, 3, 2, 0)$. According to Theorem 2 of Kao [16], it is D-optimal for estimating $\mathbf{h}$ in Model (4.1). □

Acknowledgments. The authors would like to thank the Associate Editor and reviewers for their insightful and constructive suggestions to improve the quality of this work.

SUPPLEMENTARY MATERIAL

Supplement to "Optimal design of fMRI experiments using circulant (almost-)orthogonal arrays" (DOI: 10.1214/16-AOS1531SUPP; .pdf). This supplementary material provides the generating vectors of $COA(n, K, 2, 2, 0)$ when $8 \leq n \leq 600$. These designs are obtained by Lemmas 3.8, 5.2 and Theorem 5.4 when $80 \leq n \leq 600$, and others are found by a computer search.

REFERENCES

[1] BURAČAS, G. T. and BOYNTON, G. M. (2002). Efficient design of event-related fmri experiments using m-sequences. *Neuroimage* **16** 801–813.

[2] CHAKRAVARTI, I. M. (1956). Fractional replication in asymmetrical factorial designs and partially balanced arrays. *Sankhyā* **17** 143–164. MR0085677

[3] CHAKRAVARTI, I. M. (1961). On some methods of construction of partially balanced arrays. *Ann. Math. Stat.* **32** 1181–1185. MR0130770

[4] CHÊNG, C. S. (1978). Optimality of certain asymmetrical experimental designs. *Ann. Statist.* **6** 1239–1261. MR0523760

[5] CHÊNG, C. S. (1980). Orthogonal arrays with variable numbers of symbols. *Ann. Statist.* **8** 447–453. MR0560740

[6] CHENG, C.-S. (1995). Some projection properties of orthogonal arrays. *Ann. Statist.* **23** 1223–1233. MR1353503

[7] CHENG, C.-S. (2014). Optimal biased weighing designs and two-level main-effect plans. *J. Stat. Theory Pract.* **8** 83–99. MR3196641

[8] CHENG, C.-S. and KAO, M.-H. (2015). Optimal experimental designs for fMRI via circulant biased weighing designs. *Ann. Statist.* **43** 2565–2587. MR3405604

[9] CHENG, C.-S., KAO, M.-H. and PHOA, F. K. H. (2017). Optimal and efficient designs for functional brain imaging experiments. *J. Statist. Plann. Inference* **181** 71–80. MR3567999

[10] CRAIGEN, R., FAUCHER, G., LOW, R. and WARES, T. (2013). Circulant partial Hadamard matrices. *Linear Algebra Appl.* **439** 3307–3317. MR3119854

[11] DALE, A. M. (1999). Optimal experimental design for event-related fmri. *Human Brain Mapping* **8** 109–114.

[12] GOLOMB, S. W. and GONG, G. (2005). *Signal Design for Good Correlation: For Wireless Communication, Cryptography, and Radar*. Cambridge Univ. Press, Cambridge. MR2156522

[13] HEDAYAT, A. S., SLOANE, N. J. A. and STUFKEN, J. (1999). *Orthogonal Arrays: Theory and Applications*. Springer, New York. MR1693498

[14] JANSMA, J. M., DE ZWART, J. A., VAN GELDEREN, P., DUYN, J. H., DREVETS, W. C. and FUREY, M. L. (2013). In vivo evaluation of the effect of stimulus distribution on fir statistical efficiency in event-related fmri. *Journal of Neuroscience Methods* **215** 190–195.

[15] JONES, B. and NACHTSHEIM, C. J. (2011). A class of three-level designs for definitive screening in the presence of second-order effects. *Journal of Quality Technology* **43** 1–15.

[16] KAO, M.-H. (2013). On the optimality of extended maximal length linear feedback shift register sequences. *Statist. Probab. Lett.* **83** 1479–1483. MR3048312

[17] KAO, M.-H. (2014). A new type of experimental designs for event-related fMRI via Hadamard matrices. *Statist. Probab. Lett.* **84** 108–112. MR3131263

[18] KAO, M.-H. (2015). Universally optimal fMRI designs for comparing hemodynamic response functions. *Statist. Sinica* **25** 499–506. MR3379084

[19] KAO, M.-H., MANDAL, A., LAZAR, N. and STUFKEN, J. (2009). Multi-objective optimal experimental designs for event-related fmri studies. *NeuroImage* **44** 849–856.

[20] LAZAR, N. (2008). *The Statistical Analysis of Functional MRI Data*. Springer, New York.

[21] LIN, Y.-L., PHOA, F. K. H. and KAO, M.-H. (2016+). Circulant partial Hadamard designs: Construction via general difference sets and its application to fmri experiments. Submitted.

[22] LIN, Y.-L., PHOA, F. K. H. and KAO, M.-H. (2017). Supplement to "Optimal design of fMRI experiments using circulant (almost-)orthogonal arrays." DOI:10.1214/16-AOS1531SUPP.

[23] LIN, Y.-L. and PHOA, F. K. H. (2016). Constructing near-Hadamard designs with (almost) D-optimality by general supplementary difference sets. *Statist. Sinica* **26** 413–427. MR3468358

[24] LINDQUIST, M. A. (2008). The statistical analysis of fMRI data. *Statist. Sci.* **23** 439–464. MR2530545

[25] LIU, T. T. (2004). Efficiency, power, and entropy in event-related fmri with multiple trial types: Part II: Design of experiments. *Neuroimage* **21** 401–413.

[26] LIU, T. T. and FRANK, L. R. (2004). Efficiency, power, and entropy in event-related FMRI with multiple trial types. Part I: Theory. *Neuroimage* **21** 387–400.

[27] LOW, R. M., STAMP, M., CRAIGEN, R. and FAUCHER, G. (2005). Unpredictable binary strings. *Congr. Numer.* **177** 65–75. MR2198651

[28] MAUS, B., VAN BREUKELEN, G. J., GOEBEL, R. and BERGER, M. P. (2010). Robustness of optimal design of fmri experiments with application of a genetic algorithm. *NeuroImage* **49** 2433–2443.

[29] PALEY, R. (1933). On orthogonal matrices. *J. Math. Phys.* 311–320.

[30] PHOA, F. K. H. and LIN, D. K. J. (2015). A systematic approach for the construction of definitive screening designs. *Statist. Sinica* **25** 853–861. MR3409727

[31] RADHAKRISHNA RAO, C. (1945). Finite geometries and certain derived results in theory of numbers. *Proc. Nat. Inst. Sci. India* **11** 136–149. MR0023865

[32] RADHAKRISHNA RAO, C. (1946). Hypercubes of strength "*d*" leading to confounded designs in factorial experiments. *Bull. Calcutta Math. Soc.* **38** 67–78. MR0019291

[33] RAKTOE, B. L., HEDAYAT, A. and FEDERER, W. T. (1981). *Factorial Designs*. Wiley, New York. MR0633756

[34] RYSER, H. J. (1963). *Combinatorial Mathematics*. *The Carus Mathematical Monographs* **14**. Wiley, New York. MR0150048

[35] SINGER, J. (1938). A theorem in finite projective geometry and some applications to number theory. *Trans. Amer. Math. Soc.* **43** 377–385. MR1501951

[36] STANTON, R. G. and SPROTT, D. A. (1958). A family of difference sets. *Canad. J. Math.* **10** 73–77. MR0091249

Y.-L. LIN
F. K. H. PHOA
INSTITUTE OF STATISTICAL SCIENCE
ACADEMIA SINICA
128 ACADEMIA ROAD SECTION 2, NANGANG DISTRICT
TAIPEI CITY 115
TAIWAN
E-MAIL: gaussla@stat.sinica.edu.tw
fredphoa@stat.sinica.edu.tw

M.-H. KAO
SCHOOL OF MATHEMATICS
& STATISTICAL SCIENCES
ARIZONA STATE UNIVERSITY
P.O. BOX 871804
TEMPE, ARIZONA 85287-1804
USA
E-MAIL: mkao3@asu.edu