The Annals of Statistics
2015, Vol. 43, No. 6, 2565–2587
DOI: 10.1214/15-AOS1352

OPTIMAL EXPERIMENTAL DESIGNS FOR FMRI VIA CIRCULANT BIASED WEIGHING DESIGNS

BY CHING-SHUI CHENG[*,†] AND MING-HUNG KAO[1,‡]

Academia Sinica[*], *University of California, Berkeley*[†]
and Arizona State University[‡]

Functional magnetic resonance imaging (fMRI) technology is popularly used in many fields for studying how the brain reacts to mental stimuli. The identification of optimal fMRI experimental designs is crucial for rendering precise statistical inference on brain functions, but research on this topic is very lacking. We develop a general theory to guide the selection of fMRI designs for estimating a hemodynamic response function (HRF) that models the effect over time of the mental stimulus, and for studying the comparison of two HRFs. We provide a useful connection between fMRI designs and circulant biased weighing designs, establish the statistical optimality of some well-known fMRI designs and identify several new classes of fMRI designs. Construction methods of high-quality fMRI designs are also given.

1. Introduction. The present study concerns important issues on the design of neuroimaging experiments where the pioneering functional magnetic resonance imaging (fMRI) technology is employed to gain better knowledge on how our brain reacts to mental stimuli. In such an fMRI study, a sequence of tens or hundreds stimuli (e.g., images of 1.5-second flickering checkerboard) is presented to a human subject while an fMRI scanner repeatedly scans the subject's brain to collect data for making statistical inference about brain activity; see Lazar (2008) for an overview of fMRI. The quality of such an inference largely hinges on the amount of useful information contained in the data, which in turn depends on the selected stimulus sequence (i.e., fMRI design). The importance of identifying high-quality experimental designs for fMRI and gaining insights into these designs cannot be overemphasized.

In a seminal work, Buračas and Boynton (2002) advocated the use of the maximal length linear feedback shift register sequences (or m-sequences) as fMRI designs for a precise estimate of the hemodynamic response function (HRF); the HRF models the effect over time of a brief stimulus on the relative concentration of oxy- to deoxy-blood in the cerebral blood vessels at a brain voxel (3D im-

Received December 2014; revised June 2015.
[1]Supported by NSF Grant DMS-13-52213.
MSC2010 subject classifications. 62K05.
Key words and phrases. Circulant orthogonal array, design efficiency, Hadamard matrix, hemodynamic response function, m-sequence.

age unit), and is often studied for gaining information about the underlying brain activity evoked by the stimulus. The m-sequences have since then become very popular in practice. They are also included as part of the "good" initial designs in the computer algorithm of Kao et al. (2009) to facilitate the search of multi-objective fMRI designs. The computational results of Buračas and Boynton (2002) and Liu (2004) suggested that the m-sequences can yield high statistical efficiencies in terms of the A-optimality criterion; the A-criterion, which measures the average variance of parameter estimates, is a design selection criterion widely used in many fields including fMRI [e.g., Dale (1999), Friston et al. (1999)]. By focusing on the D-optimality criterion, Kao (2014) proved the statistical optimality of the binary m-sequences in estimating the HRF. As indicated there, the binary m-sequences are special cases of the *Hadamard sequences* that can be generated from a certain type of Hadamard matrices or difference sets (Section 2.3); all these designs are D-optimal in the sense of minimizing the generalized variance of the HRF parameter estimates. While these designs are expected to be A-optimal, there unfortunately is no theoretical proof of this. One of our contributions here is to address this void. We also identify some new classes of optimal fMRI designs for the estimation of the HRF.

Another common study objective is on comparing HRFs of two stimulus types (e.g., pictures of familiar vs. unfamiliar faces). Some computational results on optimal fMRI designs for this study objective have been reported in Wager and Nichols (2003), Liu (2004), Kao, Mandal and Stufken (2008), Kao et al. (2009) and Maus et al. (2010). However, theoretical work on providing insightful knowledge to guide the selection of designs is scarce. In their pioneering papers, Liu and Frank (2004) and Maus et al. (2010) approximated the frequency of each stimulus type that an A- or D-optimal fMRI design should possess. However, designs attaining the optimal stimulus frequency can still be sub-optimal since the onset times and presentation order of the stimuli play a vital role. Working on this research line, Kao (2015) provided a sufficient condition for fMRI designs to be universally optimal in the sense of Kiefer (1975), and proposed to construct optimal designs for comparing two HRFs via an extended m-sequence (or de Bruijn sequence), a Paley difference set or a circulant partial Hadamard matrix. A major limitation of this recent contribution is that the proposed designs exist only when the design length N is a multiple of 4. New developments on identifying optimal fMRI designs for other practical N are called for.

We consider the two previously described design problems to target optimal fMRI designs for the estimation of the HRF of a stimulus type and for the comparison of two HRFs. Our main idea for tackling these design issues is by formulating them into *circulant biased weighing design* problems. With this approach, we are able to prove that the Hadamard sequences are optimal in terms of a large class of optimality criteria that include both A- and D-criteria. This holds as long as the design length $N(= 4t - 1$ for a positive integer $t)$ of such a design is sufficiently greater than the number of the HRF parameters, K. For given K, a lower bound of

N for the design to be both A- and D-optimal is also derived. This bound is easily satisfied in typical fMRI experiments. In addition, we adapt and extend previous results on (biased) weighing designs to identify some optimal fMRI designs for estimating the HRF when $N = 4t$ and $N = 4t + 1$. These results are further extended to cases where the study objective is on the comparison of two HRFs. We note that the designs that we present here exist in many design lengths for which optimal fMRI are hitherto unidentified. These designs can be applied in practice or serve as benchmarks to evaluate other designs; they help to enlarge the library of high-quality fMRI designs.

The remainder of the paper is organized as follows. In Section 2, we provide relevant background information and present our main results on optimal fMRI designs for estimating the HRF. Our results on optimal fMRI designs for the comparison of two HRFs are in Section 3, and a conclusion is in Section 4. Some proofs of our results are presented in the Appendix.

2. Designs for estimating the HRF.

2.1. *Statistical model and design selection criteria.* Consider an fMRI study where a mental stimulus such as a 1.5-second flickering checkerboard image [Boynton et al. (1996), Miezin et al. (2000)] or a painful heat stimulus [Worsley et al. (2002)] is presented/applied to a subject at some of the N time points in the experiment. Let y_n be the measurement of a brain voxel collected by an fMRI scanner at the nth time point, $n = 1, \dots, N$. We consider the following statistical model:

$$y_n = \gamma + x_n h_1 + x_{n-1} h_2 + \cdots + x_{n-K+1} h_K + \varepsilon_n, \qquad n = 1, \dots, N. \tag{2.1}$$

Here, γ is a nuisance parameter, h_1 represents the unknown height of the hemodynamic response function, HRF, at the stimulus onset time point and h_k is the HRF height at the $(k-1)$th time point following the stimulus onset. The pre-specified integer K is such that the HRF becomes negligible after K time points. The value of x_{n-k+1} in model (2.1) is set to 1 if h_k contributes to y_n and $x_{n-k+1} = 0$ otherwise, and ε_n is noise.

Our first design goal is to find an fMRI design, $\mathbf{d} = (d_1, \dots, d_N)^T$, that allows the most precise least-squares estimate of the HRF parameter vector, $\mathbf{h} = (h_1, \dots, h_K)^T$; here, $d_n = 1$ when a stimulus appears at the nth time point and $d_n = 0$ indicates no stimulus presentation at that time point, $n = 1, \dots, N$. For simplicity, we adopt the following assumptions from previous studies [Kao (2013) and references therein]; see also Kao (2014, 2015) for discussions on these assumptions. First, the last $K - 1$ elements of $\mathbf{d}$ are also presented in the pre-scanning period, that is, before the collection of y_1. With this assumption, the value of x_n in model (2.1) is d_n for $n = 1, \dots, N$, and $x_n = d_{N+n}$ for $n \leq 0$. In addition, while additional nuisance terms may be included in the model at the analysis stage to, say, allow for a trend/drift of $\mathbf{y} = (y_1, \dots, y_N)^T$, we do not assume this extra complication when deriving our analytical results on identifying optimal designs. We also consider independent noise, but our results remain true when $\mathrm{cov}(\boldsymbol{\varepsilon}) = \alpha \mathbf{I}_N + \boldsymbol{\beta}\mathbf{j}_N^T + \mathbf{j}_N \boldsymbol{\beta}^T$, where $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_N)^T$, α is a constant, $\boldsymbol{\beta}$ is a constant vector, $\mathbf{I}_N$ is the identity matrix of order N, and $\mathbf{j}_N$ is the vector of N ones; see also Kushner (1997). Other correlation structures of $\boldsymbol{\varepsilon}$ such as an autoregressive process may be considered, and is a focus of our future study. We now rewrite model (2.1) in the following matrix form:

$$\mathbf{y} = \gamma \mathbf{j}_N + \mathbf{X}_d \mathbf{h} + \boldsymbol{\varepsilon}, \tag{2.2}$$

where $\mathbf{X}_d = [\mathbf{d}, \mathbf{U}\mathbf{d}, \dots, \mathbf{U}^{K-1}\mathbf{d}]$, and

$$\mathbf{U} = \begin{bmatrix} \mathbf{0}_{N-1}^T & 1 \\ \mathbf{I}_{N-1} & \mathbf{0}_{N-1} \end{bmatrix}. \tag{2.3}$$

The information matrix for $\mathbf{h}$ is $\mathbf{M}_b(\mathbf{X}_d) = \mathbf{X}_d^T(\mathbf{I}_N - N^{-1}\mathbf{J}_N)\mathbf{X}_d$, where $\mathbf{J}_N = \mathbf{j}_N\mathbf{j}_N^T$. We also let $\mathbf{M}(\mathbf{X}_d) = \mathbf{X}_d^T\mathbf{X}_d$. Our target is at a $\mathbf{d} \in \mathcal{D} = \{0, 1\}^N$ that minimizes some real function $\Phi\{\mathbf{M}_b(\mathbf{X}_d)\}$ of $\mathbf{M}_b(\mathbf{X}_d)$. We consider the A-optimality criterion, $\Phi_A\{\mathbf{M}\} = \mathrm{tr}\{\mathbf{M}^{-1}\}/K$ for a positive definite $\mathbf{M}$, and D-optimality criterion, $\Phi_D\{\mathbf{M}\} = |\mathbf{M}|^{-1/K}$. In addition, we adopt below some other notions of optimality of designs and information matrices. Specifically, the universal optimality described in Definition 2.1 is due to Kiefer (1975). The type 1 criteria of Cheng (1978) with the version of Cheng (2014), the Φ_p-optimality criteria of Kiefer (1974) for $p \ge 0$, and the (M, S)-optimality [Eccleston and Hedayat (1974)] are also considered. Throughout this work, we set the criterion value to $+\infty$ for designs with a singular information matrix.

DEFINITION 2.1. 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) $\Phi(\mathbf{P}\mathbf{M}\mathbf{P}^T) = \phi(\mathbf{M})$ for any $\mathbf{M}$ and any orthogonal matrix $\mathbf{P}$.

DEFINITION 2.2. 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 2.3. A design $\mathbf{d}$ is said to be Φ_p-optimal over a design class for a given $p \ge 0$ if it minimizes

$$\Phi_p\{\mathbf{M}_b(\mathbf{X}_d)\} = \begin{cases} |\mathbf{M}_b(\mathbf{X}_d)|^{-1/K}, & \text{for } p = 0; \\ [\mathrm{tr}\{\mathbf{M}_b^{-p}(\mathbf{X}_d)\}/K]^{1/p}, & \text{for } p \in (0, \infty); \\ \lambda_1(\mathbf{M}_b^{-1}(\mathbf{X}_d)), & \text{when } p = \infty, \end{cases}$$

where $\lambda_i(\mathbf{M}_b(\mathbf{X}_d))$ is defined as in Definition 2.2.

DEFINITION 2.4. A matrix $\mathbf{M}^*$ is said to be (M, S)-optimal over a class $\mathcal{M}$ of nonnegative definite matrices if (i) $\mathrm{tr}\{\mathbf{M}^*\} = \max_{M\in\mathcal{M}} \mathrm{tr}\{\mathbf{M}\}$, and (ii) $\mathrm{tr}\{(\mathbf{M}^*)^2\} = \min_{M\in\mathcal{M}_m} \mathrm{tr}\{\mathbf{M}^2\}$, where $\mathcal{M}_m \subset \mathcal{M}$ consists of all the matrices having the same trace as $\mathbf{M}^*$.

Furthermore, we only consider optimality criteria Φ such that

$$\text{(2.4) if} \quad \Phi(\mathbf{M}_1) \le \Phi(\mathbf{M}_2), \qquad \text{then } \Phi(c\mathbf{M}_1) \le \Phi(c\mathbf{M}_2) \qquad \text{for all } c > 0.$$

2.2. *Circulant biased weighing designs.* Our strategy for finding optimal fMRI designs is by taking advantage of the link between these designs and circulant biased weighing designs. A biased weighing design problem concerns the selection of a design for efficient estimation of the weights of K objects in N weighings on a spring/chemical balance that has an unknown systematic bias. A spring balance weighing design (SBWD) is specified by a $\mathbf{W} \in \{0, 1\}^{N\times K}$, where the (n, k)th element of $\mathbf{W}$ indicates that the kth object is placed on the balance (1), or absent (0) in the nth weighing. Such a design is called *circulant* if $\mathbf{W}$ is a circulant matrix. The information matrix $\mathbf{M}_b(\mathbf{W})$ for the K weights is equal to $\mathbf{W}^T(\mathbf{I}_N - N^{-1}\mathbf{J}_N)\mathbf{W}$. For each fMRI design $\mathbf{d}$, the matrix $\mathbf{X}_d$ clearly defines a circulant SBWD. Thus, the fMRI design issue formulated earlier is a sub-problem of the optimal SBWD problem: selecting an optimal design among circulant SBWDs.

A chemical balance weighing design (CBWD) is specified by a $\bar{\mathbf{W}} \in \{-1, 0, 1\}^{N\times K}$, where the (n, k)th element of $\bar{\mathbf{W}}$ indicates that the kth object is placed on the left pan (-1), right pan $(+1)$, or absent (0) in the nth weighing. Each SBWD matrix $\mathbf{W}$ can be transformed into a CBWD matrix $\tilde{\mathbf{W}}$ via $\tilde{\mathbf{W}} = \pm(\mathbf{J}_{N,K} - 2\mathbf{W})$, where $\mathbf{J}_{N,K} = \mathbf{j}_N\mathbf{j}_K^T$; that is, 0 and 1 are replaced by 1 and -1, or -1 and 1, respectively. Given an fMRI design $\mathbf{d} \in \mathcal{D}$, let $\tilde{\mathbf{d}} = \pm(\mathbf{j}_N - 2\mathbf{d})$ and $\mathbf{X}_{\tilde{d}} = [\tilde{\mathbf{d}}, \mathbf{U}\tilde{\mathbf{d}}, \dots, \mathbf{U}^{K-1}\tilde{\mathbf{d}}]$, where $\mathbf{U}$ is defined as in (2.3); then, $\tilde{\mathbf{d}} \in \tilde{\mathcal{D}} = \{-1, 1\}^N$, and $\mathbf{X}_{\tilde{d}}$ is a circulant CBWD matrix. Specifically, if we write $\mathbf{X}_d$ as $\mathbf{W}$, then $\mathbf{X}_{\tilde{d}} = \tilde{\mathbf{W}}$.

Cheng (2014) showed that

$$\text{(2.5)} \qquad \mathbf{M}_b(\mathbf{W}) = \tfrac{1}{4}\mathbf{M}_b(\tilde{\mathbf{W}}) \qquad \text{for all } \mathbf{W} \in \{0, 1\}^{N\times K}.$$

We thus have the following result.

LEMMA 2.1. *For any Φ satisfying (2.4) and any $\mathbf{W}_1, \mathbf{W}_2 \in \{0, 1\}^{N\times K}$, $\Phi(\mathbf{M}_b(\mathbf{W}_1)) \le \Phi(\mathbf{M}_b(\mathbf{W}_2))$ if and only if $\Phi(\mathbf{M}_b(\tilde{\mathbf{W}}_1)) \le \Phi(\mathbf{M}_b(\tilde{\mathbf{W}}_2))$. Therefore, an fMRI design $\mathbf{d}^* \in \mathcal{D}$ is Φ-optimal for estimating the HRF if and only if*

$$\Phi\{\mathbf{M}_b(\mathbf{X}_{\tilde{d}^*})\} = \min_{\tilde{d}\in\tilde{\mathcal{D}}} \Phi\{\mathbf{M}_b(\mathbf{X}_{\tilde{d}})\} \qquad \textit{where } \tilde{\mathbf{d}}^* = \pm(\mathbf{j}_N - 2\mathbf{d}^*).$$

Lemma 2.1 reduces the problem of finding optimal fMRI designs for estimating an HRF to that of identifying optimal biased circulant CBWDs without zero entries. In Section 3, we establish the connection of optimal designs for comparing two HRFs to that of optimal biased circulant CBWDs allowing zero entries.

2.3. *Main results*. Following Lemma 2.1, we tackle our first fMRI design issue by working on circulant CBWDs ($\mathbf{X}_{\tilde{d}}$ for $\tilde{\mathbf{d}} \in \tilde{\mathcal{D}}$) that contain no zero. We have the following result for such CBWDs.

LEMMA 2.2. *For $\tilde{\mathbf{d}} \in \tilde{\mathcal{D}}$, $\mathbf{M}(\mathbf{X}_{\tilde{d}}) = \mathbf{X}_{\tilde{d}}^T\mathbf{X}_{\tilde{d}}$ has diagonal elements equal to N and off-diagonal elements congruent to N modulo* 4. *When N is odd, $\mathbf{M}_b(\mathbf{X}_{\tilde{d}}) \preceq \mathbf{M}(\mathbf{X}_{\tilde{d}}) - N^{-1}\mathbf{J}_K$ and the equality holds if $\mathbf{j}_N^T\tilde{\mathbf{d}} = \pm 1$; here, $\preceq$ is the Löwner ordering, that is, $\mathbf{M}_1 \preceq \mathbf{M}_2$ if $\mathbf{M}_2 - \mathbf{M}_1$ is nonnegative definite.*

PROOF. All the diagonal elements of $\mathbf{M}(\mathbf{X}_{\tilde{d}})$ are clearly $\tilde{\mathbf{d}}^T\tilde{\mathbf{d}} = N$. In addition, for $q, r \in \{-1, +1\}$, let $n_k^{(rq)}$ be the number of times $(\tilde{d}_{n-k}, \tilde{d}_n) = (q, r)$, where $\tilde{d}_n$ is the nth element of $\tilde{\mathbf{d}}$ for $n = 1, \dots, N$, and $\tilde{d}_n$ is set to $\tilde{d}_{N+n}$ when $n \le 0$. We have, for any $i \neq j$ and $k = |i - j|$, the (i, j)th element of $\mathbf{M}(\mathbf{X}_{\tilde{d}})$ is $n_k^{(++)} + n_k^{(--)} - (n_k^{(+-)} + n_k^{(-+)}) = N - 4n_k^{(-+)}$, and is thus congruent to N modulo 4. Note that the above equality is a consequence of the fact that $\mathbf{X}_{\tilde{d}}$ is a circulant matrix. Moreover, $\mathbf{M}_b(\mathbf{X}_{\tilde{d}}) = \mathbf{M}(\mathbf{X}_{\tilde{d}}) - a^2N^{-1}\mathbf{J}_K$, where $a = \mathbf{j}_N^T\tilde{\mathbf{d}}$ with $a^2 \ge 1$ if N is odd. Thus, $\mathbf{M}(\mathbf{X}_{\tilde{d}}) - N^{-1}\mathbf{J}_K - \mathbf{M}_b(\mathbf{X}_{\tilde{d}}) = (a^2 - 1)N^{-1}\mathbf{J}_K$, and our claim follows. $\square$

We now provide some results for obtaining optimal circulant biased CBWDs with no zero, and hence, optimal fMRI designs for estimating the HRF. For cases with $N = 4t - 1$ (≥ 4), the following lemma due to Cheng (1992) is useful.

LEMMA 2.3. *Let $\mathcal{M}^N$ be a set of K-by-K symmetric and nonnegative definite matrices, $\mathcal{M}_m^N \subset \mathcal{M}^N$ be the set of matrices that have the maximum trace over $\mathcal{M}^N$, and $\mathcal{M}_{ms}^N$ be the set of $\mathbf{M}$ that minimize* $\mathrm{tr}(\mathbf{M}^2)$ *over $\mathcal{M}_m^N$. Suppose $A_N = \max_{M \in \mathcal{M}^N} \mathrm{tr}\{\mathbf{M}\}$ and $B_N = \min_{M \in \mathcal{M}_m^N} \mathrm{tr}\{\mathbf{M}^2\}$ are such that* (a) $\lim_{N\to\infty} A_N = \infty$, *and* (b) *for some $L > 0$, $|B_N - K^{-1}A_N^2| \le L$ for all N. In addition, let $\lambda_i(\mathbf{M})$ be as in Definition 2.2, and $\Phi_{(g)}\{\mathbf{M}\} = \sum_{i=1}^K g(\lambda_i(\mathbf{M}))$ for a real-function g satisfying the following two conditions*: (i) *g is thrice continuously differentiable in a neighborhood of* 1 *with $g'(1) < 0$ and $g''(1) > 0$ and* (ii) *for any $c > 0$, there are constants $\alpha(c) > 0$ and $\beta(c)$ such that $g(cx) = \alpha(c)g(x) + \beta(c)$ for all x. Then there exists an $N_0(K, g)$ such that whenever $N \ge N_0(K, g)$, for any $\mathbf{M}^* \in \mathcal{M}_{ms}^N$, we have $\Phi_{(g)}\{\mathbf{M}^*\} < \Phi_{(g)}\{\mathbf{M}\}$ for all $\mathbf{M} \notin \mathcal{M}_{ms}^N$.*

We note that when $g(x) = -\log x$, $\Phi_{(g)}$ is equivalent to the D-criterion, or equivalently, the Φ_p-criterion in Definition 2.3 with $p = 0$. This and the other Φ_p-criteria satisfy conditions (i) and (ii) in Lemma 2.3; see also Cheng (1992). We thus have the following result on Φ_p-optimal circulant biased CBWDs and Φ_p-optimal fMRI designs for $N = 4t - 1$.

THEOREM 2.1. *Let $N = 4t - 1$, $p_0 > 0$, and $\tilde{\mathbf{d}}^* \in \tilde{\mathcal{D}}$ be such that $\mathbf{M}_b(\mathbf{X}_{\tilde{d}^*}) = (N+1)[\mathbf{I}_K - N^{-1}\mathbf{J}_K]$. Then there exists an $N_0(K, p_0)$ such that, whenever $N \geq N_0(K, p_0)$, $\tilde{\mathbf{d}}^*$ is Φ_p-optimal over $\tilde{\mathcal{D}}$, and $\mathbf{d}^* = (\mathbf{j}_N \pm \tilde{\mathbf{d}}^*)/2$ is Φ_p-optimal over $\mathcal{D}$ for any $p \in [0, p_0]$.*

PROOF. We first work on $\mathbf{M}(\mathbf{X}_{\tilde{d}}) - N^{-1}\mathbf{J}_K$ for $\tilde{\mathbf{d}} \in \tilde{\mathcal{D}}$. Following Lemma 2.2, the diagonal elements of $\mathbf{M}(\mathbf{X}_{\tilde{d}}) - N^{-1}\mathbf{J}_K$ are $N_b = N - N^{-1}$, and the (i, j)th element of this matrix is $(c_{i,j} - N^{-1})$ with $c_{i,j} = 3 (\mathrm{mod}\ 4)$ for $i \neq j$. Thus, $\mathrm{tr}\{[\mathbf{M}(\mathbf{X}_{\tilde{d}}) - N^{-1}\mathbf{J}_K]^2\}$ is minimized when $c_{i,j} = -1$ for all $i \neq j$. This implies the (M, S)-optimality of $\mathbf{M}(\mathbf{X}_{\tilde{d}^*}) - N^{-1}\mathbf{J}_K$ over $\tilde{\mathcal{M}}_b = \{\mathbf{M}(\mathbf{X}_{\tilde{d}}) - N^{-1}\mathbf{J}_K | \tilde{\mathbf{d}} \in \tilde{\mathcal{D}}\}$. In addition, it can be seen that $A_N = KN_b$, and $B_N = KN_b^2 + (1 + N^{-1})^2 K(K - 1)$, where A_N and B_N are defined as in Lemma 2.3. Therefore, $\lim_{N\to\infty} A_N = \infty$, and $|B_N - K^{-1}A_N^2| = |KN_b^2 + (1 + N^{-1})^2 K(K - 1) - KN_b^2| = (1 + N^{-1})^2 K(K - 1)$ is bounded above by a positive number for all N. Following Lemmas 2.2, and 2.3, we then have, when $N \geq N_0(K, p_0)$ for some $N_0(K, p_0)$,

$$\begin{aligned}\Phi_{p_0}\{\mathbf{M}_b(\mathbf{X}_{\tilde{d}^*})\} &= \Phi_{p_0}\{\mathbf{M}(\mathbf{X}_{\tilde{d}^*}) - N^{-1}\mathbf{J}_K\} \\ &\leq \Phi_{p_0}\{\mathbf{M}(\mathbf{X}_{\tilde{d}}) - N^{-1}\mathbf{J}_K\} \leq \Phi_{p_0}\{\mathbf{M}_b(\mathbf{X}_{\tilde{d}})\}\end{aligned}$$

for any $\tilde{\mathbf{d}} \in \tilde{\mathcal{D}}$; here Φ_{p_0} is defined as in Definition 2.3. The Φ_p-optimality of $\tilde{\mathbf{d}}^*$ over $\tilde{\mathcal{D}}$ for $p \in [0, p_0]$ then follows from Corollary 3.3 of Cheng (1987) and the fact that $\mathbf{M}_b(\mathbf{X}_{\tilde{d}^*})$ has two eigenvalues, with the smaller one having multiplicity 1. Moreover, with Lemma 2.1, we obtain the Φ_p-optimality of $\mathbf{d}^*$ over $\mathcal{D}$. $\square$

In Theorem 2.2, we provide an $N_0(K, 1)$ for a design to be Φ_1-optimal (i.e., A-optimal). Our approach for deriving this bound for N is analogous to that of Galil and Kiefer (1980), and Sathe and Shenoy (1989). A proof is provided in the Appendix.

THEOREM 2.2. *Consider the same conditions as in Theorem 2.1. Let $N_0(K, 1)$ be the greatest real root of the cubic function $c(x) = 2x^3 + (10 - 7K)x^2 + 2(2K - 5)(K - 1)x + 4K^2 - 7K$. If $K \geq 4$ and $N \geq N_0(K, 1)$, then $\tilde{\mathbf{d}}^*$ is Φ_p-optimal over $\tilde{\mathcal{D}}$, and $\mathbf{d}^*$ is Φ_p-optimal over $\mathcal{D}$ for $0 \leq p \leq 1$.*

Recently, Kao (2014) studied the efficiency of Hadamard sequences, $\mathbf{d}_H$, in estimating $\mathbf{h}$ of model (2.2). A Hadamard sequence is a binary sequence constructed from a normalized Hadamard matrix $\mathbf{H} \in \{-1, 1\}^{(N+1)\times(N+1)}$ that contains a circulant core $\tilde{\mathbf{H}}$. Such an $\mathbf{H}$ is such that $\mathbf{H}\mathbf{H}^T = (N+1)\mathbf{I}_{N+1}$, the elements of its first row and column are all 1, and the bottom-right N-by-N sub-matrix $\tilde{\mathbf{H}}$ is a circulant matrix. These Hadamard matrices are known to exist when N is a prime, a product of twin primes, or $2^r - 1$ for an integer $r > 1$. They can be easily generated by, for example, the Paley, Singer, or twin prime power difference sets [Golomb and Gong (2005), Horadam (2007)]. Any column of the circulant core $\tilde{\mathbf{H}}$ is a vertex, $\tilde{\mathbf{d}}_H$, of the hypercube $\tilde{\mathcal{D}} = \{-1, 1\}^N$, and $\mathbf{d}_H = (\mathbf{j}_N - \tilde{\mathbf{d}}_H)/2$ forms a Hadamard sequence. The popularly used binary m-sequences [Buračas and Boynton (2002)] can be constructed by the same method when $N = 2^r - 1$, and are thus special cases of $\mathbf{d}_H$. The $\mathbf{d}_H$ has design length $N = 4t - 1$ for some integer t, and our results can be applied to establish the A- and D-optimality of these designs as stated in the following corollary.

COROLLARY 2.1. *A Hadamard sequences $\mathbf{d}_H$ of length N is A- and D-optimal for estimating the HRF if $K \geq 4$ and $N \geq N_0(K, 1)$. Here, $N_0(K, 1)$ is defined as in Theorem 2.2.*

PROOF. For $\mathbf{d}_H$, let $\tilde{\mathbf{d}}_H = \mathbf{j}_N - 2\mathbf{d}_H$. Then it can be seen from the construction of $\mathbf{d}_H$ that $\mathbf{X}_{\tilde{d}_H}$ is a ciruclant matrix consisting of K distinct columns of the circulant core of a normalized Hadamard matrix. Consequently, $\mathbf{M}_b(\mathbf{X}_{\tilde{d}_H}) = (N+1)[\mathbf{I}_K - N^{-1}\mathbf{J}_K]$. Our claim then follows from Lemma 2.1 and Theorem 2.2. □

Our results so far are for cases with $N = 4t - 1$. For $N = 4t$, if there exists a $\tilde{\mathbf{d}}$ with $\mathbf{M}_b(\mathbf{X}_{\tilde{d}}) = N\mathbf{I}_K$, then $\tilde{\mathbf{d}}$ is universally optimal over $\tilde{\mathcal{D}}$, and the corresponding $\mathbf{d} = (\mathbf{j}_N \pm \tilde{\mathbf{d}})/2$ is universally optimal in estimating the HRF over all fMRI designs. This fact follows directly from Proposition 1′ of Kiefer (1975). We note that $\tilde{\mathbf{d}}$ is universally optimal whenever the columns of $\mathbf{X}_{\tilde{d}}$ are pairwise orthogonal, and are all orthogonal to $\mathbf{j}_N$. The transpose of such a matrix $\mathbf{X}_{\tilde{d}}$ is called a *circulant partial Hadamard matrix* by Craigen et al. (2013). Clearly, the corresponding $\mathbf{X}_d$ with $\mathbf{d} = (\mathbf{j}_N \pm \tilde{\mathbf{d}})/2$ forms a two-symbol, N-run, K-factor *circulant orthogonal array (OA)* whose strength is ≥ 2; see Hedayat, Sloane and Stufken (1999) for an overview of OAs. A circulant partial Hadamard matrix, and thus a circulant OA, can be obtained by a computer search [Lin, Phoa and Kao (2014), Low et al. (2005)]. Here, we provide a systematic method for constructing a universally optimal $\mathbf{d}$.

THEOREM 2.3. *Let $\mathbf{d}_{1,g,H} \in \mathcal{D}$ be obtained by inserting a 0 to a run of g 0's in a Hadamard sequence $\mathbf{d}_H$. If $K \leq g + 1$, then $\mathbf{d}_{1,g,H}$ is universally optimal for estimating $\mathbf{h}$ of model (2.2).*

PROOF. Without loss of generality, we assume that a run of g 0's appears in the tail of $\mathbf{d}_H$, and $\mathbf{d}_{1,g,H}$ is obtained by adding a 0 to this run of 0's. Suppose

$K \le q+1$, and $\tilde{\mathbf{d}}_{1,g,H} = \mathbf{j}_N - 2\mathbf{d}_{1,g,H}$. It can be seen that $\mathbf{X}_{\tilde{d}_{1,g,H}}$ is an N-by-K circulant orthogonal array whose columns are some K distinct columns of a Hadamard matrix. □

For $N = 4t+1$, Theorem 4.1 of Cheng (2014) provides a guidance on the selection of $\Phi_{(f)}$-optimal biased CBWDs for any type 1 criterion $\Phi_{(f)}$. We describe this result in Lemma 2.5 with our notation. It is interesting to note that, under our setting, a simple alternative proof of Lemma 2.5 can be achieved by utilizing Theorem 2.1 of Cheng (1980) that is slightly rephrased in Lemma 2.4 below.

LEMMA 2.4. *Let $\mathbf{M}^*$ be a symmetric matrix with eigenvalues $\lambda_1(\mathbf{M}^*) > \lambda_2(\mathbf{M}^*) = \lambda_3(\mathbf{M}^*) = \cdots = \lambda_K(\mathbf{M}^*) > 0$ and $\mathcal{M}$ be a set of nonnegative definite symmetric matrices. If the following conditions are satisfied, then $\Phi_{(f)}\{\mathbf{M}^*\} \le \Phi_{(f)}\{\mathbf{M}\}$ for any $\mathbf{M} \in \mathcal{M}$ and any type 1 criterion $\Phi_{(f)}$:*

(a) $\operatorname{tr}\{\mathbf{M}^*\} \ge \operatorname{tr}\{\mathbf{M}\}$ *for any* $\mathbf{M} \in \mathcal{M}$;

(b) *for any* $\mathbf{M} \in \mathcal{M}$, $\operatorname{tr}\{\mathbf{M}^*\} - \sqrt{[K/(K-1)][\operatorname{tr}\{(\mathbf{M}^*)^2\} - (\operatorname{tr}\{\mathbf{M}^*\})^2/K]} \ge \operatorname{tr}\{\mathbf{M}\} - \sqrt{[K/(K-1)][\operatorname{tr}\{\mathbf{M}^2\} - (\operatorname{tr}\{\mathbf{M}\})^2/K]}$.

Note that since the condition $\lim_{x\to 0^+} f(x) = f(0) = \infty$ is required in Definition 2.2, there is no need to verify (2.2) in Theorem 2.1 of Cheng (1980); see Theorem 2.3 of Cheng (1978).

LEMMA 2.5. *Let $N = 4t+1$, and $\tilde{\mathbf{d}}^* \in \tilde{\mathcal{D}}$ be such that $\mathbf{M}_b(\mathbf{X}_{\tilde{d}^*}) = (N-1)[\mathbf{I}_K + N^{-1}\mathbf{J}_K]$. Then $\tilde{\mathbf{d}}^*$ is optimal over $\tilde{\mathcal{D}}$, and $\mathbf{d}^* = (\mathbf{j}_N \pm \tilde{\mathbf{d}}^*)/2$ is optimal for estimating the HRF in terms of any type 1 criterion.*

PROOF. From Lemma 2.2, we have that the diagonal elements of $\mathbf{M}(\mathbf{X}_{\tilde{d}}) = \mathbf{X}_{\tilde{d}}^T\mathbf{X}_{\tilde{d}}$ are N, and the off-diagonal elements are congruent to 1 modulo 4. In addition, $\mathbf{M}_b(\mathbf{X}_{\tilde{d}}) \preceq \mathbf{M}(\mathbf{X}_{\tilde{d}}) - N^{-1}\mathbf{J}_K$, and the equality holds when $\mathbf{j}_N^T\tilde{\mathbf{d}} = \pm 1$. It can then be easily seen that $\mathbf{M}(\mathbf{X}_{\tilde{d}^*}) - N^{-1}\mathbf{J}_K$ is (M,S)-optimal over $\tilde{\mathcal{M}}_b = \{\mathbf{M}(\mathbf{X}_{\tilde{d}}) - N^{-1}\mathbf{J}_K \mid \tilde{\mathbf{d}} \in \tilde{\mathcal{D}}\}$, and conditions (a) and (b) in Lemma 2.4 are satisfied if we replace $\mathbf{M}^*$, $\mathbf{M}$ and $\mathcal{M}$ there by $\mathbf{M}(\mathbf{X}_{\tilde{d}^*}) - N^{-1}\mathbf{J}_K$, $\mathbf{M}(\mathbf{X}_{\tilde{d}}) - N^{-1}\mathbf{J}_K$, and $\tilde{\mathcal{M}}_b$, respectively. Consequently, for any type 1 criterion $\Phi_{(f)}$, we have $\Phi_{(f)}\{\mathbf{M}_b(\mathbf{X}_{\tilde{d}^*})\} = \Phi_{(f)}\{\mathbf{M}(\mathbf{X}_{\tilde{d}^*}) - N^{-1}\mathbf{J}_K\} \le \Phi_{(f)}\{\mathbf{M}(\mathbf{X}_{\tilde{d}}) - N^{-1}\mathbf{J}_K\} \le \Phi_{(f)}\{\mathbf{M}_b(\mathbf{X}_{\tilde{d}})\}$. The optimality of $\tilde{\mathbf{d}}^*$ thus follows. By (2.5), this argument also applies to $\mathbf{d}^*$. □

We now provide a systematic method for constructing optimal fMRI designs for cases with $N = 4t+1$, followed by an example on an application of our results in this section.

TABLE 1
A Hadamard sequence $\mathbf{d}_H$, and a $\mathbf{d}_{1,g,H}$ for estimating $\mathbf{h}$ with $K \le 9$

	N	**Design**
$\mathbf{d}_H$	151	1 0 0 1 0 0 1 1 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 1 1 0 1 1 1 0 1 0 0 1 0 1 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 1 1 1 1 0 1 1 0 0 1 1 0 1 0 1 1 1 0 0 1 1 0 1 0 1 0 1 0 1 0 0 1 1 0 0 0 1 0 1 0 0 1 1 0 0 1 0 0 0 0 1 1 0 1 0 1 1 1 1 0 1 1 1 1 1 0 1 0 1 1 0 1 0 0 0 1 0 0 1 1 1 1 1 1 1 0 0 0 0 1 1 1 1 0 0 1 1 0 1 1
$\mathbf{d}_{1,g,H}$	132	1 0 1 0 0 0 1 0 1 0 1 0 0 0 1 0 0 1 1 1 0 0 1 1 1 0 1 0 0 1 1 1 1 0 0 0 0 1 0 0 1 0 1 0 0 0 0 1 0 0 1 1 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 0 0 1 0 1 1 0 0 1 1 0 1 1 1 1 0 1 0 1 1 0 1 1 1 1 0 0 0 0 1 1 0 1 0 0 0 1 1 0 0 0 1 1 0 1 1 1 0 1 0 1 0 1 1 1 0 1

THEOREM 2.4. *Let $\mathbf{d}_{2,g,H} \in \mathcal{D}$ be obtained by inserting two* 0*'s to a run of g* 0*'s in a Hadamard sequence $\mathbf{d}_H$. If $K \le g+1$, then $\mathbf{d}_{2,g,H}$ is optimal for estimating $\mathbf{h}$ of model* (2.2) *for all type* 1 *criteria.*

PROOF. With a similar argument as in the proof of Theorem 2.3, we have that, when $K \le g+1$ and $\tilde{\mathbf{d}}_{2,g,H} = \mathbf{j}_N - 2\mathbf{d}_{2,g,H}$, $\mathbf{M}_b(\mathbf{X}_{\tilde{d}_{2,g,H}}) = (N-1)[\mathbf{I}_K + N^{-1}\mathbf{J}_K]$. Our claim then follows from Lemma 2.5. $\square$

EXAMPLE 2.1. Consider an experiment where the stimulus can possibly occur every 4 seconds. Then $N = 4(38) - 1 = 151$ corresponds to a 10-minute experiment and $K = 9$ corresponds to a 32-second HRF. A $\mathbf{d}_H = (d_{1,H}, \dots, d_{N,H})^T$ can be obtained by a Paley difference set [Paley (1933)]. This is to set $d_{n,H} = 0$ if $(n-1) \in \{x^2 \pmod{N} | x = 1, \dots, (N-1)/2\}$ and $d_{n,H} = 1$, otherwise. The obtained design $\mathbf{d}_H$ is presented in Table 1. It is both A- and D-optimal for estimating $\mathbf{h}$ of model (2.2) since $N > N_0(K, 1) = 21.34$.

Following Theorem 2.3, we may insert a 0 to the longest run of 0's in the $\mathbf{d}_H$ presented in Table 1 to yield a universally optimal design. The resulting design can accommodate a $K \le 8$. For $K = 9$, we obtain a universally optimal $\mathbf{d}_{1,g,H}$ by extending a $\mathbf{d}_H$ of length $N = 131$. This $\mathbf{d}_{1,g,H}$ is presented in Table 1. We also obtain a $\mathbf{d}_{2,g,H}$ by inserting another 0 into the longest run of 0's in $\mathbf{d}_{1,g,H}$. Following Theorem 2.4, this $\mathbf{d}_{2,g,H}$ is optimal for any type 1 criterion in estimating $\mathbf{h}$ with $K \le 9$.

It is noteworthy that, by replacing 0 and 1 with 1 and 2, respectively, the $\mathbf{d}_{1,g,H}$ in Table 1 is equivalent to the design of the same design length in Table 3.1 of Kao (2015). The use of such a design whose elements are 1 or 2 is discussed in the next section.

3. Designs for contrasts between HRFs. We now consider optimal fMRI experimental designs for studies where the objective is on comparing HRFs of two stimulus types. For this situation, Kao (2015) presented some optimal designs for $N = 4t$ by considering the following extension of model (2.2):

$$\mathbf{y} = \gamma \mathbf{j}_N + \mathbf{X}_{u,1}\mathbf{h}_1 + \mathbf{X}_{u,2}\mathbf{h}_2 + \boldsymbol{\varepsilon}, \tag{3.1}$$

where $\mathbf{h}_q = (h_{q1}, \dots, h_{qK})^T$ is the vector of the K unknown HRF heights of the qth-type stimulus, $\mathbf{X}_{u,q}$ is the 0-1 design matrix obtained from the selected fMRI design $\mathbf{u} = (u_1, \dots, u_N)^T$ with $u_n \in \{0, 1, 2\}$, $q = 1, 2$, and the remaining terms are as in (2.2). Specifically, $u_n = q > 0$ indicates that a stimulus of the qth type appears at the nth time point, and $u_n = 0$ if no stimulus is present. In addition, for $q = 1, 2$, $\mathbf{X}_{u,q} = [\boldsymbol{\delta}_q, \mathbf{U}\boldsymbol{\delta}_q, \dots, \mathbf{U}^{K-1}\boldsymbol{\delta}_q]$, where $\mathbf{U}$ is defined in (2.3), and the nth element of $\boldsymbol{\delta}_q$ is 1 if $u_n = q$, and is 0 otherwise. The main interest lies in $\boldsymbol{\zeta} = \mathbf{h}_1 - \mathbf{h}_2$, and we may rewrite model (3.1) as

$$\mathbf{y} = \gamma \mathbf{j}_N + \mathbf{E}_u\boldsymbol{\eta} + \mathbf{F}_u\boldsymbol{\zeta} + \boldsymbol{\varepsilon}, \tag{3.2}$$

where $\mathbf{E}_u = (\mathbf{X}_{u,1} + \mathbf{X}_{u,2})/2$, $\boldsymbol{\eta} = \mathbf{h}_1 + \mathbf{h}_2$, $\mathbf{F}_u = (\mathbf{X}_{u,1} - \mathbf{X}_{u,2})/2$, and all the remaining terms are as in (3.1). The aim is thus at obtaining a design $\mathbf{u} \in \{0, 1, 2\}^N$ so that $\Phi\{\mathbf{M}_u\}$ is minimized, where $\mathbf{M}_u = \mathbf{F}_u^T(\mathbf{I}_N - \omega\{[\mathbf{j}_N, \mathbf{E}_u]\})\mathbf{F}_u$ and $\omega\{\mathbf{A}\}$ is the orthogonal projection matrix onto the space spanned by the columns of the matrix $\mathbf{A}$. The following lemma can be easily proved.

LEMMA 3.1. *For a given design $\mathbf{u} \in \{0, 1, 2\}^N$, let $\bar{\mathbf{d}}_u = (\bar{d}_{u,1}, \dots, \bar{d}_{u,N})^T \in \bar{\mathcal{D}} = \{-1, 0, 1\}^N$ be defined as $\bar{d}_{u,n} = 0$, 1 and -1 when $u_n = 0$, 1 and 2, respectively. Then*

$$\mathbf{M}_u \preceq \mathbf{F}_u^T(\mathbf{I}_N - N^{-1}\mathbf{J}_N)\mathbf{F}_u = \mathbf{X}_{\bar{d}_u}^T(\mathbf{I}_N - N^{-1}\mathbf{J}_N)\mathbf{X}_{\bar{d}_u}/4,$$

where $\mathbf{X}_{\bar{d}_u} = [\bar{\mathbf{d}}_u, \mathbf{U}\bar{\mathbf{d}}_u, \dots, \mathbf{U}^{K-1}\bar{\mathbf{d}}_u]$, and $\mathbf{U}$ is as in (2.3). In addition, $\mathbf{M}_u = \mathbf{X}_{\bar{d}_u}^T(\mathbf{I}_N - N^{-1}\mathbf{J}_N)\mathbf{X}_{\bar{d}_u}/4$ if $\mathbf{u}$ contains no zero.

Our approach for obtaining optimal fMRI designs for comparing HRFs is by working on the upper bound of $\mathbf{M}_u$ provided in Lemma 3.1. Specifically, we would like to obtain a $\bar{\mathbf{d}}_u \in \bar{\mathcal{D}}$, or equivalently a circulant CBWD $\mathbf{X}_{\bar{d}_u} \in \{-1, 0, 1\}^{N\times K}$, that minimizes $\Phi\{\mathbf{M}_b(\mathbf{X}_{\bar{d}_u})\}$. As pointed out at the end of Section 2.2, unlike the case of estimating an HRF, here we also need to consider circulant CBWDs with zero entries. If the obtained $\bar{\mathbf{d}}_u$ contains no zero, then the corresponding $\mathbf{u}$ is Φ-optimal. To identify such a $\bar{\mathbf{d}}_u$, we consider the following lemma. For convenience, we omit the subscript of $\bar{\mathbf{d}}_u$ hereinafter, but its dependence on $\mathbf{u}$ should be clear.

LEMMA 3.2. *Suppose $\bar{\mathbf{d}} \in \bar{\mathcal{D}}$ contains r zeros, and $\mathbf{j}_N^T\bar{\mathbf{d}} = a$. We have the following results*:

(i) *If* $N = 4t - 1$, *and* $\mathbf{M}_b(\mathbf{X}_{\bar{d}}) = (N + 1)[\mathbf{I}_K - N^{-1}\mathbf{J}_K]$, *then* $a^2 = 1$, $r = 0$, *and* $\mathbf{M}(\mathbf{X}_{\bar{d}}) = \mathbf{X}_{\bar{d}}^T\mathbf{X}_{\bar{d}} = (N + 1)\mathbf{I}_K - \mathbf{J}_K$.

(ii) *If* $N = 4t + 1$, *and* $\mathbf{M}_b(\mathbf{X}_{\bar{d}}) = (N - 1)[\mathbf{I}_K + N^{-1}\mathbf{J}_K]$, *then* $a^2 = 1$, $r = 0$, *and* $\mathbf{M}(\mathbf{X}_{\bar{d}}) = (N - 1)\mathbf{I}_K + \mathbf{J}_K$.

PROOF. We work only on (i) here. A similar argument can be applied to prove (ii). For (i), we have $\mathbf{M}_b(\mathbf{X}_{\bar{d}}) = \mathbf{M}(\mathbf{X}_{\bar{d}}) - (a^2/N)\mathbf{J}_K = (N + 1)[\mathbf{I}_K - N^{-1}\mathbf{J}_K]$. Since each diagonal element of $\mathbf{M}(\mathbf{X}_{\bar{d}})$ is an integer that is not greater than N, it can be seen that $a^2 \leq 1$. If $a = 0$, then $\mathbf{X}_{\bar{d}}^T\mathbf{X}_{\bar{d}} = (N + 1)[\mathbf{I}_K - N^{-1}\mathbf{J}_K]$. This leads to a contradiction since the diagonal elements of the latter matrix are $(N + 1)(1 - N^{-1})$. Therefore, $a^2 = 1$, $\mathbf{M}(\mathbf{X}_{\bar{d}}) = (N + 1)\mathbf{I}_K - \mathbf{J}_K$ and $r = 0$. $\square$

The first main result in this section is an extension of Theorem 2.1. We note that $\bar{N}_0(K, p_0)$ in Theorem 3.1 may not be the same as $N_0(K, p_0)$ in Theorem 2.1.

THEOREM 3.1. *Suppose* $\bar{\mathbf{d}}^* \in \bar{\mathcal{D}}$ *is a vector with* $N = 4t - 1$ *elements, and it satisfies* $\mathbf{M}_b(\mathbf{X}_{\bar{d}^*}) = (N + 1)[\mathbf{I}_K - N^{-1}\mathbf{J}_K]$. *For any positive number* p_0, *there exists an* $\bar{N}_0(K, p_0)$ *such that, if* $N \geq \bar{N}_0(K, p_0)$, *then* $\bar{\mathbf{d}}^*$ *is* Φ_p*-optimal over* $\bar{\mathcal{D}}$ *for any* $p \in [0, p_0]$.

PROOF. Let r be the number of zeros in $\bar{\mathbf{d}} \in \bar{\mathcal{D}}$. It is clear that $\mathrm{tr}\{\mathbf{M}(\mathbf{X}_{\bar{d}}) - N^{-1}\mathbf{J}_K\} = (N - r) - K/N$ is maximized when $r = 0$. With Lemma 3.2, we can easily see that $\mathbf{M}(\mathbf{X}_{\bar{d}^*}) - N^{-1}\mathbf{J}_K$ is (M, S)-optimal over $\bar{\mathcal{M}}_b = \{\mathbf{M}(\mathbf{X}_{\bar{d}}) - N^{-1}\mathbf{J}_K\}$. Following Lemma 2.3 and Corollary 3.3 of Cheng (1987), $\bar{\mathbf{d}}^*$ is Φ_p-optimal over $\bar{\mathcal{D}}$ for $p \in [0, p_0]$ when $N \geq \bar{N}_0(K, p_0)$ for some $\bar{N}_0(K, p_0)$. $\square$

An explicit lower bound $N_0(K, 1)$ for the A-criterion was given in Theorem 2.2. We show in the following theorem that one can take $\bar{N}_0(K, 1) = N_0(K, 1)$.

THEOREM 3.2. *Let* $\bar{\mathbf{d}}^* \in \bar{\mathcal{D}}$, *where* $N = 4t - 1$, *be such that* $\mathbf{M}_b(\mathbf{X}_{\bar{d}^*}) = (N + 1)[\mathbf{I}_K - N^{-1}\mathbf{J}_K]$. *If* $K \geq 4$ *and* $N \geq N_0(K, 1)$, *where* $N_0(K, 1)$ *is as given in Theorem* 2.2, *then* $\bar{\mathbf{d}}^*$ *is* A*-optimal* (*and* Φ_p*-optimal for all* $p \in [0, 1]$) *over* $\bar{\mathcal{D}}$.

PROOF. By Theorem 2.2, it is enough to show that if $K \geq 4$ and $N \geq N_0(K, 1)$, then for any $\bar{\mathbf{d}} \in \bar{\mathcal{D}}$ that has at least one zero entry, $\Phi_1\{\mathbf{M}_b(\mathbf{X}_{\bar{d}^*})\} \leq \Phi_1\{\mathbf{M}_b(\mathbf{X}_{\bar{d}})\}$. If $\bar{\mathbf{d}} \in \bar{\mathcal{D}}$ has at least one zero entry, then each diagonal entry of $\mathbf{M}_b(\mathbf{X}_{\bar{d}})$ is at most $N - 1$; thus $\Phi_1\{\mathbf{M}_b(\mathbf{X}_{\bar{d}})\} \geq K/(N - 1)$. On the other hand, since $\mathbf{M}_b(\mathbf{X}_{\bar{d}^*}) = (N + 1)[\mathbf{I}_K - N^{-1}\mathbf{J}_K]$ has two distinct eigenvalues $N + 1$ and $(N + 1)(N - K)/N$, with multiplicity $K - 1$ and 1, respectively, we have $\Phi_1\{\mathbf{M}_b(\mathbf{X}_{\bar{d}^*})\} = (K - 1)/(N + 1) + N/[(N + 1)(N - K)]$. It follows that $\Phi_1\{\mathbf{M}_b(\mathbf{X}_{\bar{d}^*})\} \leq \Phi_1\{\mathbf{M}_b(\mathbf{X}_{\bar{d}})\}$ provided $(K - 1)/(N + 1) + N/[(N + 1)(N - K)] \leq K/(N - 1)$. The latter is the same as $N \geq 2K - 1$. Therefore, it remains to show that $2K-1 \le N_0(K,1)$. Since $N_0(K,1)$ is the greatest real root of the cubic function $c(x)=2x^3+(10-7K)x^2+2(2K-5)(K-1)x+4K^2-7K$, $c(x)>0$ for all $x>N_0(K,1)$. One can verify that $c(2K-1)<0$ when $K\ge 4$. It follows that in this case $2K-1<N_0(K,1)$. $\square$

For $N=4t$, circulant OAs or equivalently circulant partial Hadamard matrices described in Section 2.3 can be used to construct $\bar{\mathbf{d}}^*$ that has $\mathbf{M}_b(\mathbf{X}_{\bar{d}^*})=N\mathbf{I}_K$. Such a $\bar{\mathbf{d}}^*$ can be easily seen to be universally optimal in $\bar{\mathcal{D}}$.

Theorem 3.3 below helps to identify some optimal $\bar{\mathbf{d}}$ for $N=4t+1$. For deriving this theorem, we again consider Lemmas 2.4 and 2.5.

THEOREM 3.3. *For $N=4t+1$, let $\bar{\mathbf{d}}^*\in\bar{\mathcal{D}}$ have $\mathbf{M}_b(\mathbf{X}_{\bar{d}^*})=(N-1)[\mathbf{I}_K+N^{-1}\mathbf{J}_K]$. Then, $\bar{\mathbf{d}}^*$ is optimal over $\bar{\mathcal{D}}$ for any type 1 criterion.*

PROOF. $\mathbf{M}_b(\mathbf{X}_{\bar{d}^*})$ has two nonzero eigenvalues, and the smaller eigenvalue has multiplicity $K-1$. It can also be seen that condition (a) in Lemma 2.4 is satisfied by $\bar{\mathbf{d}}^*$. In addition, $\operatorname{tr}\{\mathbf{M}_b(\mathbf{X}_{\bar{d}^*})\}=K(N-N^{-1})$, and $\operatorname{tr}\{\mathbf{M}_b^2(\mathbf{X}_{\bar{d}^*})\}-(\operatorname{tr}\{\mathbf{M}_b(\mathbf{X}_{\bar{d}^*})\})^2/K=K(K-1)(1-N^{-1})^2$. Thus, condition (b) of Lemma 2.4 is satisfied if and only if

$$K\big(N-N^{-1}\big)-A_{\bar{d}}\ge\sqrt{\frac{K}{K-1}}\left[(1-N^{-1})\sqrt{K(K-1)}-\sqrt{B_{\bar{d}}-\frac{A_{\bar{d}}^2}{K}}\right]\tag{3.3}$$

for $\bar{\mathbf{d}}\in\bar{\mathcal{D}}$, where $A_{\bar{d}}=\operatorname{tr}\{\mathbf{M}_b(\mathbf{X}_{\bar{d}})\}$, and $B_{\bar{d}}=\operatorname{tr}\{\mathbf{M}_b^2(\mathbf{X}_{\bar{d}})\}$. Clearly, (3.3) holds for the class of $\bar{\mathbf{d}}\in\bar{\mathcal{D}}$ that satisfy $A_{\bar{d}}\le K(N-N^{-1})-\sqrt{K/(K-1)}(1-N^{-1})\sqrt{K(K-1)}=K(N-1)$. Thus, all $\bar{\mathbf{d}}$'s in this class are outperformed by $\bar{\mathbf{d}}^*$ with respect to any type 1 criterion. For any other $\bar{\mathbf{d}}$, let r be the number of zeros in $\bar{\mathbf{d}}$ and $a=\mathbf{j}_N^T\bar{\mathbf{d}}$. Then we have $A_{\bar{d}}>K(N-1)$, and $(1-r-a^2/N)>0$ since $A_{\bar{d}}=K[(N-r)-a^2/N]$. Consequently, $r=0$, and $\bar{\mathbf{d}}$ contains no zero. Following Lemma 2.5, $\bar{\mathbf{d}}^*$ is also optimal over the class of designs with no zero for any type 1 criterion. Our claim then follows. $\square$

With these results, we can derive the following theorem for identifying some optimal fMRI designs for studying contrasts between two HRFs.

THEOREM 3.4. *Suppose $\mathbf{d}_H$ is a Hadamard sequence, $\mathbf{d}_{1,g,H}$ is defined as in Theorem 2.3, $\mathbf{d}_{2,g,H}$ is as in Theorem 2.4, and $\mathbf{M}_u$ is the information matrix for $\boldsymbol{\zeta}$ in model (3.2) for a design $\mathbf{u}\in\{0,1,2\}^N$. We have the following results:*

(a) *Suppose $N=4t-1$, and $\mathbf{u}^*=\mathbf{j}_N+\mathbf{d}_H$ or $\mathbf{u}^*=2\mathbf{j}_N-\mathbf{d}_H$. If $p_0>0$, $p\in[0,p_0]$, and $N\ge\bar{N}_0(K,p_0)$ for some $\bar{N}_0(K,p_0)>0$, then $\Phi_p\{\mathbf{M}_{u^*}\}\le\Phi_p\{\mathbf{M}_u\}$ for any $\mathbf{u}\in\{0,1,2\}^N$; here, Φ_p is defined as in Definition 2.3. For $K\ge 4$, we can take $\bar{N}_0(K,1)$ to be the $N_0(K,1)$ given in Theorem 2.2.*

(b) *Suppose* $N = 4t$, *and* $\mathbf{u}^* = \mathbf{j}_N + \mathbf{d}_{1,g,H}$ *or* $\mathbf{u}^* = 2\mathbf{j}_N - \mathbf{d}_{1,g,H}$. *If* $K \le g+1$, *and* Φ *is any criterion satisfying the conditions in Definition* 2.1, *then* $\Phi\{\mathbf{M}_{u^*}\} \le \Phi\{\mathbf{M}_u\}$ *for any* $\mathbf{u} \in \{0, 1, 2\}^N$.

(c) *Suppose* $N = 4t + 1$, *and* $\mathbf{u}^* = \mathbf{j}_N + \mathbf{d}_{2,g,H}$ *or* $\mathbf{u}^* = 2\mathbf{j}_N - \mathbf{d}_{2,g,H}$. *If* $K \le g+1$, *and* $\Phi_{(f)}$ *is any type* 1 *criterion defined in Definition* 2.2, *then* $\Phi_{(f)}\{\mathbf{M}_{u^*}\} \le \Phi_{(f)}\{\mathbf{M}_u\}$ *for any* $\mathbf{u} \in \{0, 1, 2\}^N$.

PROOF. For all the designs $\mathbf{u}^*$ in (a), (b) and (c), we have $\mathbf{M}_{u^*} = \mathbf{X}^T_{\bar{d}_{u^*}}(\mathbf{I}_N - N^{-1}\mathbf{J}_N)\mathbf{X}_{\bar{d}_{u^*}}/4$, where $\bar{\mathbf{d}}_{u^*}$ is defined as in Lemma 3.1. When $\mathbf{u}^* = \mathbf{j}_N + \mathbf{d}_H$ (or, resp., $\mathbf{u}^* = 2\mathbf{j}_N - \mathbf{d}_H$), we have $\mathbf{X}_{\bar{d}_{u^*}} = \mathbf{X}_{\bar{d}_H}$ with $\bar{\mathbf{d}}_H = \mathbf{j}_N - 2\mathbf{d}_H$ (or, resp., $\bar{\mathbf{d}}_H = 2\mathbf{d}_H - \mathbf{j}_N$). We thus have that, if $N \ge \bar{N}_0(K, p_0)$ and $p \in [0, p_0]$, then

$$\Phi_p\{\mathbf{M}_{u^*}\} = \Phi_p\{\mathbf{M}_b(\mathbf{X}_{\bar{d}_H})/4\} \le \Phi_p\{\mathbf{M}_b(\mathbf{X}_{\bar{d}_u})/4\} \le \Phi_p\{\mathbf{M}_u\}$$

for any $\mathbf{u} \in \{0, 1, 2\}^N$. This completes the proof for (a). Similar arguments can be used to prove (b) and (c) and are omitted. □

4. Conclusion. Neuroimaging experiments utilizing the pioneering fMRI technology are widely conducted in a variety of research fields for gaining better knowledge about human brain functions. One of the key steps to ensure the success of such an experiment is to judiciously select an optimal fMRI design. Existing studies on obtaining optimal fMRI designs primarily focus on computational approaches. However, insightful analytical results, while important, are rather scarce and scattered. To address this important issue, we conduct a systematic and analytical analysis to characterize some optimal fMRI designs for estimating the HRF of a stimulus type and for comparing HRFs of two stimulus types. Under certain conditions, we show that the popularly used binary m-sequences as well as the more general Hadamard sequences are optimal in some statistically meaningful senses. We also identify several new classes of high-quality fMRI designs and present systematic methods for constructing them. These designs exist in many design lengths where good fMRI designs have not been reported previously. There, however, are many research challenges that need to be overcome. For example, our results provide good designs for design lengths of $N = 4t - 1$, $4t$ and $4t + 1$. A future research of interest is on identifying optimal fMRI designs for cases with $N = 4t + 2$. In addition, our experience indicates that the designs that we present here remain quite efficient under some violations of model assumptions [cf. Kao (2014)]. Nevertheless, it still is of interest to analytically study optimal designs for other situations (e.g., with an autoregressive error process). Extending current results to cases with a greater number of stimulus types is also a future research of interest. Many research opportunities exist in this new and wide-open research area.

APPENDIX: A PROOF OF THEOREM 2.2

For $N = 3 \pmod 4) \geq 4$ and $K \geq 4$, we consider the following set of K-by-K nonnegative definite matrices:

$$\Xi_{N,K} = \{\mathbf{E}_K = ((e_{ij}))_{i,j=1,\dots,K} | e_{ij} = 3 \ (\mathrm{mod}\ 4)\ \forall i, j, e_{ii} = N, \mathbf{E}_K \succ N^{-1}\mathbf{J}_K\},$$

where $\mathbf{E}_K \succ N^{-1}\mathbf{J}_K$ indicates that $\mathbf{E}_K - N^{-1}\mathbf{J}_K$ is positive definite. With Lemma 2.2, it can be seen that $\mathbf{M}(\mathbf{X}_{\tilde{d}}) \in \Xi_{N,K}$ for any $\tilde{\mathbf{d}} \in \tilde{\mathcal{D}}$ having a nonsingular $\mathbf{M}_b(\mathbf{X}_{\tilde{d}})$. The idea for proving Theorem 2.2 is then to show that an $\mathbf{E}_K \in \Xi_{N,K}$ minimizing $\mathrm{tr}\{[\mathbf{E}_K - N^{-1}\mathbf{J}_K]^{-1}\}$ is similar (with some permutations of rows and columns) to a block matrix $\mathbf{B} \in \Xi_{N,K}$ to be defined in Definition A.1 below. We also will show in Lemma A.3 that, when the condition in Theorem 2.2 is satisfied, we have $\mathrm{tr}\{[\mathbf{M}(\mathbf{X}_{\tilde{d}*}) - N^{-1}\mathbf{J}_K]^{-1}\} = \min_{B\in\mathcal{B}} \mathrm{tr}\{[\mathbf{B} - N^{-1}\mathbf{J}_K]^{-1}\}$, where $\mathcal{B} \subset \Xi_{N,K}$ is the set of all block matrices. With these facts and Lemma 2.2, we have

$$\begin{aligned}
&\Phi_1\{\mathbf{M}_b(\mathbf{X}_{\tilde{d}*})\}\\
&\quad = \Phi_1\{\mathbf{M}(\mathbf{X}_{\tilde{d}*}) - N^{-1}\mathbf{J}_K\} = \min_{E_k\in\Xi_{N,K}} \Phi_1\{\mathbf{E}_K - N^{-1}\mathbf{J}_K\} \qquad \text{(A.1)}\\
&\quad \leq \min_{\tilde{d}\in\tilde{\mathcal{D}}} \Phi_1\{\mathbf{M}(\mathbf{X}_{\tilde{d}}) - N^{-1}\mathbf{J}_K\} \leq \min_{\tilde{d}\in\tilde{\mathcal{D}}} \Phi_1\{\mathbf{M}_b(\mathbf{X}_{\tilde{d}})\}.
\end{aligned}$$

Our claim in Theorem 2.2 then follows from (A.1), and Corollary 3.3 of Cheng (1987). This approach is similar to that of Sathe and Shenoy (1989), and Galil and Kiefer (1980), where weighing designs under the unbiasedness assumption are considered. We now present the details of our proof.

DEFINITION A.1. A block matrix $\mathbf{B} \in \Xi_{N,K}$ is of the form

$$\mathbf{B} = \bigoplus_{i=1}^{m} [(N-3)\mathbf{I}_{r_i} + 4\mathbf{J}_{r_i}] - \mathbf{J}_K$$

for an integer $m \in \{1, \dots, K\}$ representing the number of "blocks." Here, $\bigoplus$ is the matrix direct sum, and $r_1, r_2, \dots, r_m$ are the block sizes of $\mathbf{B}$ that satisfy $r_i \geq 1$, and $\sum_{i=1}^{m} r_i = K$.

For these block matrices, the following result is an extension of Theorem 2.1(a) of Masaro (1988) and equation (1.1) of Sathe and Shenoy (1989). Clearly, this result also implies that $\mathrm{tr}\{\mathbf{B}_b^{-1}\}$ is invariant to a rearrangement of the block sizes $r_1, \dots, r_m$, which facilitates the derivation of the subsequent results; here $\mathbf{B}_b = \mathbf{B} - N^{-1}\mathbf{J}_K$.

LEMMA A.1. *Let $\mathcal{B} \subset \Xi_{N,K}$ be the set of all block matrices. Then for $\mathbf{B} \in \mathcal{B}$, we have*

$$\operatorname{tr}\{\mathbf{B}_b^{-1}\} = \sum_{i=1}^{m} L_i^{-1} + \frac{K-m}{N-3} + \frac{\sum_{i=1}^m r_i L_i^{-2}}{(1+N^{-1})^{-1} - \sum_{i=1}^m r_i L_i^{-1}}$$
$$= \frac{1}{4t}\left\{K - \sum_{i=1}^{m} \frac{r_i}{t+r_i} + \frac{t\sum_{i=1}^m r_i/(t+r_i)^2}{4/(1+N^{-1}) - \sum_{i=1}^m r_i/(t+r_i)}\right\}, \tag{A.2}$$

where $L_i = N - 3 + 4r_i$, and $t = (N-3)/4 \geq 1$.

PROOF. Let $\mathbf{C}_{r_i} = (N-3)\mathbf{I}_{r_i} + 4\mathbf{J}_{r_i}$, we have $\mathbf{B}_b = \bigoplus_{i=1}^m \mathbf{C}_{r_i} - (1+N^{-1})\mathbf{J}_K$, and

$$\mathbf{B}_b^{-1} = \left[\bigoplus_{i=1}^{m} \mathbf{C}_{r_i} - (1+N^{-1})\mathbf{j}_K\mathbf{j}_K'\right]^{-1}$$
$$= \bigoplus_{i=1}^{m} \mathbf{C}_{r_i}^{-1} + \frac{(1+N^{-1})[\bigoplus_{i=1}^m \mathbf{C}_{r_i}^{-1}]\mathbf{j}_K\mathbf{j}_K'[\bigoplus_{i=1}^m \mathbf{C}_{r_i}^{-1}]}{1-(1+N^{-1})\sum_{i=1}^m \mathbf{j}_{r_i}\mathbf{C}_{r_i}^{-1}\mathbf{j}_{r_i}}.$$

The two equalities in (A.2) can then be derived by some simple algebra. □

The following lemma indicates that a block matrix minimizing the trace of $\mathbf{B}_b^{-1} = [\mathbf{B} - N^{-1}\mathbf{J}_K]^{-1}$ can be found over a small subset of $\mathcal{B}$ [cf. Theorem 2.1(b) of Masaro (1988)].

LEMMA A.2. *Let $\mathcal{B}_s \subset \mathcal{B}$ be the set of block matrices having blocks of only one size or two contiguous sizes. Then*

$$\min_{\mathbf{B}\in\mathcal{B}_s} \operatorname{tr}\{\mathbf{B}_b^{-1}\} = \min_{\mathbf{B}\in\mathcal{B}} \operatorname{tr}\{\mathbf{B}_b^{-1}\}.$$

PROOF. Among the block matrices that yield the minimum $\operatorname{tr}\{\mathbf{B}_b^{-1}\}$ over $\mathcal{B}$, let $\mathbf{B}_m$ be the one with the smallest number of blocks, m. Clearly, we only need to consider cases where $m \geq 2$. Without loss of generality (see also the statement above Lemma A.1), we may assume that the first two block sizes r_1 and r_2 are, respectively, the largest and the smallest block sizes among the m block sizes. With Lemma A.1, we can then write $\operatorname{tr}\{\mathbf{B}_{m,b}^{-1}\} = [K + f(x)]/4t$, where $\mathbf{B}_{m,b} = \mathbf{B}_m - N^{-1}\mathbf{J}_K$,

$$f(x) = \frac{t[x/(t+x)^2 + (r-x)/(t+r-x)^2 + \beta]}{4/(1+N^{-1}) - [x/(t+x) + (r-x)/(t+r-x) + \alpha]}$$
$$- \left[\frac{x}{t+x} + \frac{r-x}{t+r-x} + \alpha\right],$$

$r = r_1 + r_2$, $x = r_1$ (or r_2) with $0 < x < r$, $\alpha = \sum_{i=3}^{m} r_i/(t + r_i)$, and $\beta = \sum_{i=3}^{m} r_i/(t + r_i)^2$. We note that this expression of $\mathrm{tr}\{\mathbf{B}_b^{-1}\}$ applies to any block matrix, and that $\alpha = \beta = 0$ for block matrices with two or fewer blocks. Since the number of blocks of $\mathbf{B}_m$ is $m \geq 2$, we have $\mathrm{tr}\{\mathbf{B}_{m,b}^{-1}\} < \mathrm{tr}\{\mathbf{B}_b^{-1}\}$ for those block matrices $\mathbf{B}$ with only one block; thus, $f(x) < f(0) = f(r)$. With some simple algebra similar to that of Masaro (1988), we also have

$$f(x) = h(y) = \frac{ay^2 + b}{cy^2 + d} \quad \text{and} \quad f'(x) = h'(y) = \frac{2(ad - bc)y}{(cy^2 + d)^2},$$

where $y = x - 0.5r \in (-0.5r, 0.5r)$, and a, b, c, d are some constants. Along with the fact that $f(x) < f(0) = f(r)$, and f is symmetric about $0.5r$, the minimum of $f(x)$ occurs when x is the integer closest to $0.5r$. Consequently, $r_1 = r_2$ when r is even, and $r_1 = r_2 + 1$ when r is odd. □

With these results, we can now work on $\mathcal{B}_{0,s} \subset \mathcal{B}_s$ that consists of block matrices having $m(< K)$ block sizes with $r_1 - 1 = \cdots = r_v - 1 = r_{v+1} = \cdots = r_m = r \geq 1$, and $v \geq 1$. We note that, for any $m_0 \geq 1$ and $r_0 \geq 2$, a block matrix having $(m, v, r) = (m_0, 0, r_0)$ can be treated as a block matrix with $(m, v, r) = (m_0, m_0, r_0 - 1)$, and is thus in $\mathcal{B}_{0,s}$; see also Sathe and Shenoy (1989). Consequently, the only block matrix in $\mathcal{B}_s$ that is left out from $\mathcal{B}_{0,s}$ is $\mathbf{B}^* = (N+1)\mathbf{I}_K - \mathbf{J}_K$, which has $(m, v, r) = (K, 0, 1)$, or equivalently, $(m, v, r) = (K, K, 0)$. Under the condition described in the following lemma, we have $\mathrm{tr}\{(\mathbf{B}^* - N^{-1}\mathbf{J}_K)^{-1}\} \leq \mathrm{tr}\{(\mathbf{B}_s - N^{-1}\mathbf{J}_K)^{-1}\}$ for any other $\mathbf{B}_s \in \mathcal{B}_s$.

LEMMA A.3. *Let* $\mathbf{B}^* = (N+1)\mathbf{I}_K - \mathbf{J}_K$, $\mathbf{B}_s \in \mathcal{B}_{0,s}$ *be a previously described block matrix,* $\mathbf{B}_b^* = \mathbf{B}^* - N^{-1}\mathbf{J}_K$, $\mathbf{B}_{s,b} = \mathbf{B}_s - N^{-1}\mathbf{J}_K$, *and* $N_0(K, 1)$ *be the greatest real root of the cubic function* $c(x) = 2x^3 + (10 - 7K)x^2 + 2(2K - 5)(K - 1)x + 4K^2 - 7K$. *If* $N \geq N_0(K, 1)$, *then*

$$\mathrm{tr}\{(\mathbf{B}_b^*)^{-1}\} \leq \mathrm{tr}\{\mathbf{B}_{s,b}^{-1}\}.$$

PROOF. Let $u = m - v$, and $\mathbf{B}$ be obtained by replacing a block of size $r + 1$ in $\mathbf{B}_s$ with a block of size r and a block of size 1. From (A.2), we have

$$\mathrm{tr}\{\mathbf{B}_{s,b}^{-1}\} = \frac{K - m}{N - 3} + \frac{u - 1}{L} + \frac{v - 1}{L + 4} + \frac{(2L + 4)(1 + N^{-1})^{-1} - K}{g(v)},$$

where $L = N - 3 + 4r$, and $g(v) = 4v(r+1) + (L+4)\xi$, and $\xi = [L(1+N^{-1})^{-1} - K]$. In addition,

$$\mathrm{tr}\{\mathbf{B}_b^{-1}\} = \frac{K - m - 1}{N - 3} + \frac{u}{L} + \frac{v - 2}{L + 4} + \frac{1}{N + 1} + \frac{(2L + 4)(1 + N^{-1})^{-1} - K + 16r(r - 1)(N + 1)^{-2}}{g(v) - 4r(N + 1 + L)(N + 1)^{-1}}.$$

Since $g(v) \geq g(1)$, we have

$$
\begin{aligned}
&\operatorname{tr}\{\mathbf{B}_{s,b}^{-1} - \mathbf{B}_b^{-1}\} \\
&\quad = \frac{4r}{L(N-3)} + \frac{(2L+4)(1+N^{-1})^{-1} - K}{g(v)} + \frac{1}{L+4} - \frac{1}{N+1} \\
&\qquad - \frac{(2L+4)(1+N^{-1})^{-1} - K + 16r(r-1)(N+1)^{-2}}{g(v) - 4r(N+1+L)(N+1)^{-1}} \\
&\quad \geq \frac{4r}{L(N-3)} + \frac{(2L+4)(1+N^{-1})^{-1} - K}{g(1)} \\
&\qquad - \frac{(2L+4-4r)(1+N^{-1})^{-1} - K}{(N+1)\xi - 4(r-1)} = \Delta(r).
\end{aligned}
$$

The equality holds when $v = 1$. With some algebra, we have

$$
\begin{aligned}
\Delta(r) = {} & \frac{8r(N+12r-11)(N+1)^{-1}c(N)}{L(N-3)[(N+1)\xi - 4(r-1)][(L+4)\xi + 4(r+1)]} \\
& + 16r(r-1) \\
& \quad \times \frac{N^2 - 7N + 12 - 8(1+N^{-1})^{-1} + 16(r-1)^2 N(3N-1)(N+1)^{-2}}{L(N-3)[(N+1)\xi - 4(r-1)][(L+4)\xi + 4(r+1)]} \\
& + 16r(r-1) \\
& \quad \times \frac{4(r-1)(N+1)^{-1}[(7N-K)(N-K) + 5(N^2-1) + 8N]}{L(N-3)[(N+1)\xi - 4(r-1)][(L+4)\xi + 4(r+1)]},
\end{aligned}
$$

where $c(N) = 2N^3 + (10-7K)N^2 + 2(2K-5)(K-1)N + 4K^2 - 7K$. With $N = 3 \pmod 4) \geq 4$ and $K \geq 4$, it can be seen that, when $N \geq N_0(K, 1)$, we have $c(N) \geq 0$, $\Delta(r) > 0$ for $r > 1$, and $\Delta(1) \geq 0$. Consequently, for a $\mathbf{B}_s$, we either (i) find a $\mathbf{B} \notin \mathcal{B}_s$ with $\operatorname{tr}\{\mathbf{B}_{s,b}^{-1}\} > \operatorname{tr}\{\mathbf{B}_b^{-1}\}$, or (ii) can keep splitting each block of size $r+1 = 2$ in $\mathbf{B}_s$ into two blocks of size 1 without increasing the objective function ($\operatorname{tr}\{\mathbf{B}_b^{-1}\}$). For the first case, $\mathbf{B}_s (\neq \mathbf{B}^*)$ is obviously not an optimal block matrix of our interest. For the second case, we can continue the process until $\mathbf{B} = \mathbf{B}^*$. Our claim thus follows. $\square$

The results we have so far suggest that, under the condition of Lemma A.3, $\mathbf{B}^* = (N+1)\mathbf{I}_K - \mathbf{J}_K$ minimizes $\operatorname{tr}\{\mathbf{B}_b^{-1}\}$ over all block matrices $\mathbf{B}$. With the following Lemma A.4, our proof of Theorem 2.2 is then complete.

LEMMA A.4. *Let* $\mathbf{B}^* = (N+1)\mathbf{I}_K - \mathbf{J}_K$. *If* $N \geq N_0(K, 1)$ *for the* $N_0(K, 1)$ *defined in Lemma* A.3, *then*

$$
\operatorname{tr}\{(\mathbf{B}_b^*)^{-1}\} = \min_{\mathbf{E}_K \in \Xi_{N,K}} \operatorname{tr}\{[\mathbf{E}_K - N^{-1}\mathbf{J}_K]^{-1}\}.
$$

The proof of Lemma A.4 is lengthy, but otherwise is a simple extension of that of Theorem 2.2 of Sathe and Shenoy (1989). The main idea is to show that an $\mathbf{E}_K^* \in \Xi_{N,K}$ minimizing $\mathrm{tr}\{(\mathbf{E}_{b,K}^{-1}\}$ is similar to a block matrix after some permutations of rows and columns. Lemma A.4 then follows from Lemma A.3. To that end, we need the following lemmas, which are extensions of results in Sathe and Shenoy (1989). Lemma A.5 is a well-known result, and the proof is omitted. We also use the following notation: $\mathbf{E}_K^* \in \Xi_{N,K}$ is a matrix such that $\mathrm{tr}\{(\mathbf{E}_{b,K}^*)^{-1}\} = \min_{\mathbf{E}_K \in \Xi_{N,K}} \mathrm{tr}\{\mathbf{E}_{b,K}^{-1}\}$, $N_b = N - N^{-1}$, $3_b = 3 - N^{-1}$, $c_b = c - N^{-1}$ for some $c = 3 \pmod 4$, and $\boldsymbol{\mu}_{b,i} = \boldsymbol{\mu}_i - N^{-1}\mathbf{j}_{K-2}$ and $\boldsymbol{\mu}_{b,j} = \boldsymbol{\mu}_j - N^{-1}\mathbf{j}_{K-2}$ for some $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ whose elements are congruent to 3 modulo 4. In addition, $a_{b,i,j} = \boldsymbol{\mu}_{b,i}^T \mathbf{E}_{b,K-2}^{-1} \boldsymbol{\mu}_{b,j}^T$, $b_{b,i,j} = \boldsymbol{\mu}_{b,i}^T \mathbf{E}_{b,K-2}^{-2} \boldsymbol{\mu}_{b,j}^T$, $A_{b,i,j} = N_b - a_{b,i,j}$, $z_{b,i,j}(c_b) = c_b - a_{b,i,j}$, and

$$f_{b,i,j}(c_b) = \frac{(A_{b,i,i} + A_{b,j,j}) + A_{b,i,i} b_{b,j,j} + A_{b,j,j} b_{b,i,i} - 2b_{b,i,j} z_{b,i,j}(c_b)}{A_{b,i,i} A_{b,j,j} - z_{b,i,j}^2(c_b)}. \tag{A.3}$$

LEMMA A.5. *Let* $\mathbf{E} = ((\mathbf{E}_{ij}))$ *for* $i, j = 1, 2$ *be a partitioned positive definite matrix, where* $\mathbf{E}_{11}$ *and* $\mathbf{E}_{22}$ *are square matrices. We have*

$$\mathrm{tr}\{\mathbf{E}^{-1}\} = \mathrm{tr}\{\mathbf{E}_{22}^{-1}\} + \mathrm{tr}\{\mathbf{V}[\mathbf{I} + \mathbf{E}_{12}\mathbf{E}_{22}^{-2}\mathbf{E}_{21}]\},$$

where $\mathbf{V} = (\mathbf{E}_{11} - \mathbf{E}_{12}\mathbf{E}_{22}^{-1}\mathbf{E}_{21})^{-1}$.

LEMMA A.6. $\mathrm{tr}\{(\mathbf{E}_{b,K}^*)^{-1}\} < \mathrm{tr}\{(\mathbf{E}_{b,K-1}^*)^{-1}\} + (N-3)^{-1}$.

PROOF. For $K = 2$, $\mathrm{tr}\{\mathbf{E}_{b,2}^{-1}\} = 2N_b/(N_b^2 - c_b^2)$ is minimized when $c = -1$, or equivalently, $c_b = -1 - N^{-1}$. Thus, $\mathrm{tr}\{(\mathbf{E}_{b,2}^*)^{-1}\} - \mathrm{tr}\{(\mathbf{E}_{b,1}^*)^{-1}\} =$

$$\frac{2N_b}{N_b^2 - (1+N^{-1})^2} - N_b = \frac{N^2 - 2N + 2}{(N+1)(N-1)(N-2)} < \frac{1}{N-3}.$$

Suppose $\mathrm{tr}\{(\mathbf{E}_{b,K-1}^*)^{-1}\} < \mathrm{tr}\{(\mathbf{E}_{b,K-2}^*)^{-1}\} + (N-3)^{-1}$. We would like to show that $\mathrm{tr}\{(\mathbf{E}_{b,K}^*)^{-1}\} < \mathrm{tr}\{(\mathbf{E}_{b,K-1}^*)^{-1}\} + (N-3)^{-1}$. To that end, we write

$$\mathbf{E}_{b,K-1}^* = \begin{bmatrix} N_b & \boldsymbol{\mu}_{b,j}^T \\ \boldsymbol{\mu}_{b,j} & \mathbf{E}_{b,K-2} \end{bmatrix}.$$

With Lemma A.5 and the fact that $\mathbf{E}_{b,K-1}^*$ is positive definite, we have $b_{b,j,j} > 0$ and

$$\begin{aligned} A_{b,j,j}^{-1} < A_{b,j,j}^{-1}(1 + b_{b,j,j}) &= \mathrm{tr}\{(\mathbf{E}_{b,K-1}^*)^{-1}\} - \mathrm{tr}\{\mathbf{E}_{b,K-2}^{-1}\} \\ &\le \mathrm{tr}\{(\mathbf{E}_{b,K-1}^*)^{-1}\} - \mathrm{tr}\{(\mathbf{E}_{b,K-2}^*)^{-1}\} < (N-3)^{-1}. \end{aligned}$$

Thus, $A_{b,j,j} = N_b - a_{b,j,j} > N - 3$, and this in turn implies $z_{b,j,j}(3_b) = 3_b - a_{b,j,j} > 0$. Following this fact and some simple algebra, we can show that the following matrix $\mathbf{E}_{b,K}(j)$, which is obtained by adding a row and a column to $\mathbf{E}^*_{b,K-1}$, is positive definite, and is thus in $\Xi_{N,K}$:

$$\mathbf{E}_{b,K}(j) = \begin{bmatrix} N_b & 3_b & \boldsymbol{\mu}^T_{b,j} \\ 3_b & N_b & \boldsymbol{\mu}^T_{b,j} \\ \boldsymbol{\mu}_{b,j} & \boldsymbol{\mu}_{b,j} & \mathbf{E}_{b,K-2} \end{bmatrix}.$$

With Lemma A.5, we also have

$$\begin{aligned} \operatorname{tr}\{\mathbf{E}^{-1}_{b,K}(j)\} &= \operatorname{tr}\{\mathbf{E}^{-1}_{b,K-2}\} + f_{b,j,j}(3_b) \\ &= \operatorname{tr}\{(\mathbf{E}^*_{b,K-1})^{-1}\} - A^{-1}_{b,j,j}(1 + b_{b,j,j}) + f_{b,j,j}(3_b). \end{aligned} \tag{A.4}$$

By noting that $f_{b,j,j}(3_b)$ in (A.3) can be written as

$$f_{b,j,j}(3_b) = \frac{2[z_{b,j,j}(3_b) + (N-3)(1 + b_{b,j,j})]}{(N-3)(N-3+2z_{b,j,j}(3_b))},$$

we can show that $f_{b,j,j}(3_b) - A^{-1}_{b,j,j}(1 + b_{b,j,j}) < (N-3)^{-1}$. The proof is then completed by the fact that

$$\begin{aligned} \operatorname{tr}\{(\mathbf{E}^*_{b,K})^{-1}\} \le \operatorname{tr}\{\mathbf{E}^{-1}_{b,K}(j)\} &= \operatorname{tr}\{(\mathbf{E}^*_{b,K-1})^{-1}\} + f_{b,j,j}(3_b) - A^{-1}_{b,j,j}(1 + b_{b,j,j}) \\ &< \operatorname{tr}\{(\mathbf{E}^*_{b,K-1})^{-1}\} + (N-3)^{-1}. \end{aligned}$$ □

LEMMA A.7. *Write* $\mathbf{E}^*_{b,K} = \mathbf{E}^*_K - N^{-1}\mathbf{J}_K$ *in the form of*

$$\mathbf{E}^*_{b,K} = \begin{bmatrix} N_b & c_b & \boldsymbol{\mu}^T_{b,i} \\ c_b & N_b & \boldsymbol{\mu}^T_{b,j} \\ \boldsymbol{\mu}_{b,i} & \boldsymbol{\mu}_{b,j} & \mathbf{E}_{b,K-2} \end{bmatrix} = \begin{bmatrix} N & c & \boldsymbol{\mu}^T_i \\ c & N & \boldsymbol{\mu}^T_j \\ \boldsymbol{\mu}_i & \boldsymbol{\mu}_j & \mathbf{E}_{K-2} \end{bmatrix} - N^{-1}\mathbf{J}_K. \tag{A.5}$$

We have (i) $\operatorname{tr}\{(\mathbf{E}^*_{b,K})^{-1}\} = \operatorname{tr}\{\mathbf{E}^{-1}_{b,K-2}\} + f_{b,i,j}(c_b)$ *and* (ii) $f_{b,i,j}(c_b) < 2(N-3)^{-1}$.

PROOF. We first replace $\mathbf{E}$, $\mathbf{E}_{11}$ and $\mathbf{E}_{22}$ in Lemma A.5 by $\mathbf{E}^*_{b,K}$,

$$\mathbf{E}_{11} = \begin{bmatrix} N_b & c_b \\ c_b & N_b \end{bmatrix},$$

and $\mathbf{E}_{b,K-2}$, respectively. This allows to verify (i). In addition, we have from Lemma A.6 that

$$\begin{aligned} f_{b,i,j}(c_b) &= \operatorname{tr}\{(\mathbf{E}^*_{b,K})^{-1}\} - \operatorname{tr}\{\mathbf{E}^{-1}_{b,K-2}\} \le \operatorname{tr}\{(\mathbf{E}^*_{b,K})^{-1}\} - \operatorname{tr}\{(\mathbf{E}^*_{b,K-2})^{-1}\} \\ &\le \operatorname{tr}\{(\mathbf{E}^*_{b,K})^{-1}\} - \operatorname{tr}\{(\mathbf{E}^*_{b,K-1})^{-1}\} + \operatorname{tr}\{(\mathbf{E}^*_{b,K-1})^{-1}\} - \operatorname{tr}\{(\mathbf{E}^*_{b,K-2})^{-1}\} \\ &< 2(N-3). \end{aligned}$$

This proves (ii). $\square$

We now are ready to prove that $\mathbf{E}_K^*$ is similar to a block matrix. This is done by considering the expression of $\mathbf{E}_{b,K}^*$ in (A.5). We then will show that if $|c| > 3$, then $f_{b,i,j}(c_b) \geq 2(N-1)^{-1}$, which contradicted with Lemma A.7(ii), a necessary condition of Lemma A.6. Since the same argument can be applied after permuting the rows and columns of $\mathbf{E}_{b,K}^*$, $\mathbf{E}_K^*$ must have off-diagonal elements equal to -1 or 3, and is thus similar to a block matrix. We begin this procedure by deriving some useful results. With Lemma A.7(i) and equation (A.4), we have

$$\begin{aligned}\operatorname{tr}\{(\mathbf{E}_{b,K}^*)^{-1}\} &= \operatorname{tr}\{\mathbf{E}_{b,K-2}^{-1}\} + f_{b,i,j}(c_b)\\ &\leq \min\{\operatorname{tr}\{\mathbf{E}_{b,K-2}^{-1}\} + f_{b,i,i}(3_b), \operatorname{tr}\{\mathbf{E}_{b,K-2}^{-1}\} + f_{b,j,j}(3_b)\}.\end{aligned}$$

Thus, $f_{b,i,j}(c_b) \leq \min\{f_{b,i,i}(3_b), f_{b,j,j}(3_b)\}$. Let $z_{b,g}(c_b) = \sqrt{z_{b,i,i}(c_b) z_{b,j,j}(c_b)}$; then

$$\begin{aligned}&\frac{z_{b,i,i}(3_b) b_{b,j,j} + z_{b,j,j}(3_b) b_{b,i,i}}{2}\\ &\quad\geq \sqrt{z_{b,i,i}(3_b) b_{b,j,j} z_{b,j,j}(3_b) b_{b,i,i}}\\ &\quad= \sqrt{z_{b,i,i}(3_b) z_{b,j,j}(3_b)}\sqrt{b_{b,i,i} b_{b,j,j}} \geq z_{b,g}(3_b)|b_{b,i,j}|.\end{aligned}$$

The last inequality is due to the Cauchy–Schwarz inequality [see also, Theorem 14.10.1 of Harville (1997)]. With the same reason, we also have $a_{b,i,j}^2 \leq a_{b,i,i} a_{b,j,j}$. Thus, for $|c| > 3$,

$$\begin{aligned}|z_{b,i,j}(c_b)| &= |c_b - a_{b,i,j}| \geq |c| - |N^{-1}| - |a_{b,i,j}| > 3_b - \sqrt{a_{b,i,i} a_{b,j,j}}\\ &\geq 3_b - \frac{a_{b,i,i} + a_{b,j,j}}{2} = \frac{3_b - a_{b,i,i} + 3_b - a_{b,j,j}}{2}\\ &= \frac{z_{b,i,i}(3_b) + z_{b,j,j}(3_b)}{2}.\end{aligned}$$

We note that $a_{b,i,i}$, and $a_{b,j,j}$ are positive since $\mathbf{E}_{K-2}$ is positive definite. Let $z_{b,a}(c_b) = (z_{b,i,i}(c_b) + z_{b,j,j}(c_b))/2$. We have, for $|c| > 3$, $|z_{b,i,j}(c_b)| > z_{b,a}(3_b) \geq z_{b,g}(3_b)$. It also can be easily seen that

$$f_{b,j,j}(3_b) = \frac{2[A_{b,j,j} + (N-3) b_{b,j,j}]}{(N-3)[A_{b,j,j} + z_{b,j,j}(3_b)]}.$$

With these facts and some algebra similar to that in Sathe and Shenoy (1989), we can see that, for $|c| > 3$,

$$\begin{aligned}&[A_{b,i,i} A_{b,j,j} - z_{b,i,j}^2(c_b)] f_{b,i,j}(c_b)\\ &\quad= [(N-3)^2 + 2(N-3) z_{b,a}(3_b) + z_{b,g}^2(3_b) - z_{b,i,j}^2(c_n)] f_{b,i,j}(c_b)\end{aligned}$$

$$\geq 2\big[N-3+z_{b,a}(3_b)\big] + \big[(N-3)z_{b,a}(3_b)+z_{b,g}^2(3_b)-z_{b,i,j}^2(c_b)\big]f_{b,i,j}(c_b).$$

The first equality is due to $A_{b,i,j}=N-c+z_{b,i,j}(c_b)$. This in turn leads to $f_{b,i,j}(c_b)\geq 2(N-3)^{-1}$. With Lemma A.7(ii), we thus can conclude that $|c|\leq 3$ and $\mathbf{E}_K^*$ is similar to a block matrix. The proof of Lemma A.4 is then completed by using Lemma A.3.

REFERENCES

BOYNTON, G. M., ENGEL, S. A., GLOVER, G. H. and HEEGER, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *J. Neurosci*. **16** 4207–4221.

BURAČAS, G. T. and BOYNTON, G. M. (2002). Efficient design of event-related fMRI experiments using M-sequences. *NeuroImage* **16** 801–813.

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

CHENG, C. S. (1980). Optimality of some weighing and 2^n fractional factorial designs. *Ann. Statist*. **8** 436–446. MR0560739

CHENG, C.-S. (1987). An optimization problem with applications to optimal design theory. *Ann. Statist*. **15** 712–723. MR0888435

CHENG, C.-S. (1992). On the optimality of (M, S)-optimal designs in large systems. *Sankhyā Ser. A* **54** 117–125. MR1234687

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

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

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

ECCLESTON, J. A. and HEDAYAT, A. (1974). On the theory of connected designs: Characterization and optimality. *Ann. Statist*. **2** 1238–1255. MR0362672

FRISTON, K. J., ZARAHN, E., JOSEPHS, O., HENSON, R. N. and DALE, A. M. (1999). Stochastic designs in event-related fMRI. *NeuroImage* **10** 607–619.

GALIL, Z. and KIEFER, J. (1980). *D*-optimum weighing designs. *Ann. Statist*. **8** 1293–1306. MR0594646

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

HARVILLE, D. A. (1997). *Matrix Algebra from a Statistician's Perspective*. Springer, New York. MR1467237

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

HORADAM, K. J. (2007). *Hadamard Matrices and Their Applications*. Princeton Univ. Press, Princeton, NJ. MR2265694

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

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

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

KAO, M.-H., MANDAL, A. and STUFKEN, J. (2008). Optimal design for event-related functional magnetic resonance imaging considering both individual stimulus effects and pairwise contrasts. *Stat. Appl.* (*N. S.*) **6** 225–241.

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.

KIEFER, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* **2** 849–879. MR0356386

KIEFER, J. (1975). Construction and optimality of generalized Youden designs. In *A Survey of Statistical Design and Linear Models* (*Proc. Internat. Sympos.*, *Colorado State Univ.*, *Ft. Collins*, *Colo.*, 1973) (J. N. Srivastava, ed.) 333–353. North-Holland, Amsterdam. MR0395079

KUSHNER, H. B. (1997). Optimal repeated measurements designs: The linear optimality equations. *Ann. Statist.* **25** 2328–2344. MR1604457

LAZAR, N. A. (2008). *The Statistical Analysis of Functional MRI Data*, *Statistics for Biology and Health*. Springer, New York.

LIN, Y.-L., PHOA, F. K. H. and KAO, M.-H. (2014). Partial Hadamard matrices: Construction via general difference sets and application to fMRI designs. Unpublished manuscript.

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.

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.

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

MASARO, J. C. (1988). On A-optimal block matrices and weighing designs when $N \equiv 3 \pmod 4$. *J. Statist. Plann. Inference* **18** 363–370. MR0926638

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

MIEZIN, F. M., MACCOTTA, L., OLLINGER, J. M., PETERSEN, S. E. and BUCKNER, R. L. (2000). Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *NeuroImage* **11** 735–759.

PALEY, R. (1933). On orthogonal matrices. *Journal of Mathematics and Physics* **12** 311–320.

SATHE, Y. S. and SHENOY, R. G. (1989). A-optimal weighing designs when $N = 3 \pmod 4$. *Ann. Statist.* **17** 1906–1915. MR1026319

WAGER, T. D. and NICHOLS, T. E. (2003). Optimization of experimental design in fMRI: A general framework using a genetic algorithm. *NeuroImage* **18** 293–309.

WORSLEY, K. J., LIAO, C. H., ASTON, J., PETRE, V., DUNCAN, G. H., MORALES, F. and EVANS, A. C. (2002). A general statistical analysis for fMRI data. *NeuroImage* **15** 1–15.

INSTITUTE OF STATISTICAL SCIENCE
ACADEMIA SINICA
TAIPEI 11529
TAIWAN
E-MAIL: cheng@stat.sinica.edu.tw

SCHOOL OF MATHEMATICAL AND STATISTICAL SCIENCES
ARIZONA STATE UNIVERSITY
TEMPE, ARIZONA 85287
USA
E-MAIL: mkao3@asu.edu