Optimal and Efficient Designs for Functional Brain Imaging Experiments

Ching-Shui Cheng^a, Ming-Hung Kao^{b,*}, Federick Kin Hing Phoa^a

^a*Institute of Statistical Science, Academia Sinica, 11529 Taipei, Taiwan*

^b*School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287, USA*

Abstract

This article concerns optimal experimental designs for neuroimaging studies in which the pioneering functional magnetic resonance imaging (fMRI) technology is used to investigate functions of the human brain. We develop analytical results for identifying designs that allow precise estimates of the contrasts between hemodynamic response functions (HRFs), each describing the effect of a mental stimulus over time. We also provide systematic methods for constructing optimal and very efficient designs, and show that fMRI designs obtained by relabeling the symbols of m -sequences can be very efficient in comparing the HRFs.

Keywords: Circulant orthogonal array, De Bruijn sequence, Hadamard sequence, Schur optimality, Type 1 criterion

1. Introduction

Functional neuroimaging experiments in which the experimenter utilizes the functional magnetic resonance imaging (fMRI) technology to study human brain activity in response to mental stimuli are widely conducted in many fields. In such an experiment, the subject is exposed to a sequence of mental stimuli of one or more types while an fMRI scanner repeatedly scans the subject's brain to collect the blood-oxygen-level-dependent (BOLD) signals for making statistical inference about the inner workings of the brain. An important design goal is to obtain an optimal sequence of stimuli (or an fMRI design) to allow an efficient collection of informative data for rendering a precise inference. See Lazar (2008) for an overview of statistical methods

*Corresponding author

Preprint submitted to Journal of Statistical Planning and Inference
Email address: mkao3@asu.edu (Ming-Hung Kao)

July 21, 2016

in fMRI, and Kao & Stufken (2015) for a review of some previous works on the design of fMRI experiments.

For mathematical convenience, we represent an fMRI design as a column vector $\mathbf{d} = (d_1, \dots, d_N)^T \in \{0, 1, \dots, Q\}^N$. Here, N (typically tens or hundreds) is the number of equally spaced time points where a stimulus can possibly appear, and Q is the number of stimulus types. The value of Q can be any integer ≥ 1 , and is pre-determined by the experimenter based on the research question(s) of interest. For example, one might repeatedly present to the experimental subject a flashing checkerboard interweaved with a control condition (i.e., $Q=1$; Blamire et al., 1992), or study the effects of two stimulus types such as pictures of faces and houses (Summerfield et al., 2006). It also is not uncommon to consider three or more types of stimuli (e.g. Zhang & Yu, 2008; Brendel et al., 2010). When d_n of $\mathbf{d}$ is a positive integer $q \leq Q$, a stimulus of the q th type will be presented to the experimental subject at the n th time point. There will be no stimulus presentation at that time point if $d_n = 0$. Each stimulus evokes a change in the ratio of oxy- to deoxy-blood concentrations in some cerebral blood vessels of the subject. At each brain voxel (3D imaging unit), this change is detected by the fMRI scanner with much noise, and is described by a hemodynamic response function (HRF). When a brain region is activated by the stimulus, the HRF of the corresponding voxel will start to increase after the stimulus onset and then descend to baseline, possibly with an undershooting before their complete return to baseline. This process typically takes several time points. The HRFs, which help to study the underlying brain activity, are of the main interest in most fMRI studies. At the design stage, we would like to select an optimal $\mathbf{d}$ for making a precise inference about the HRFs. This however is a challenging task.

Optimal designs have been playing a crucial role in fMRI studies. Experimenters often seek designs best suited to their study objectives before conducting the experiment; see, e.g., Jiang et al. (2007), Kubilius et al. (2011), Eck et al. (2013), and Jansma et al. (2013) for some applications of optimal design techniques in real fMRI experiments. A popular way for finding an optimal fMRI design is by computational approaches (Wager & Nichols, 2003; Liu, 2004; Kao et al., 2009; Delzell et al., 2012; Kao et al., 2012, 2013; Kao & Mittelman, 2014). Recent years have also seen some advances in the development of analytical results for guiding the selection of good fMRI designs. In particular, Liu (2004) and Maus et al. (2010) approximated the optimal frequency of each stimulus type that an optimal design

59 can possess for (i) estimating the individual HRFs, which will be referred to
 60 as Objective I hereinafter, and (ii) comparing HRFs, which will be referred
 61 to as Objective II. The focus of Liu (2004) is on the A -optimality criterion,
 62 and that of Maus et al. (2010) is on the D -optimality criterion. These two
 63 design selection criteria are the most dominant in the current fMRI design
 64 literature. Working on Objective I, Kao (2013) established the optimality of
 65 the extended m -sequences, or more generally the de Bruijn sequences; see,
 66 e.g., Buracas & Boynton (2002) and Aguirre et al. (2011) for the applica-
 67 tions of these designs in fMRI. For the same study objective but with $Q = 1$,
 68 Kao (2014) put forward the use of Hadamard sequences that exist in many
 69 design lengths for which the binary (extended) m -sequences and de Bruijn
 70 sequences are unavailable. By focusing on Objective II with $Q = 2$, Kao
 71 (2015) identified some optimal $\mathbf{d}$ with design length $N = 4t$ for some pos-
 72 itive integer t . Cheng & Kao (2015) conducted a systematic study to give
 73 a collection of optimal designs for Objective I with $Q = 1$ and Objective II
 74 with $Q = 2$ when $N = 4t$ and $N = 4t \pm 1$.

75 There remain many challenging cases where the development of new op-
 76 timal fMRI design theory is called for. The present article contributes to this
 77 research line with a focus on Objective II. Specifically, for $Q = 2$, we develop
 78 optimal designs for cases with $N = 4t + 2$. By relabeling the symbols, the
 79 designs that we obtain are optimal for Objective I with $Q = 1$. This first
 80 result thus completes the cases discussed in Cheng & Kao (2015). We then
 81 extend the results of Kao (2015) to find optimal and very efficient designs for
 82 cases with $Q \geq 3$. Our result for these cases suggests that the m -sequences,
 83 which have been known useful for Objective I, can easily be modified to gen-
 84 erate good designs for Objective II. This gives a simple systematic method
 85 for constructing high-quality designs for Objective II when $Q \geq 3$. Such a
 86 design construction method is handy and useful, but to our knowledge, is
 87 previously unavailable.

88 When evaluating the goodness of designs, we mainly follow the previous
 89 studies to focus on the A - and D -optimality criteria. Nevertheless, for most
 90 designs that we study, we are able to prove strong optimality results by
 91 considering some large classes of optimality criteria that include both A -
 92 and D -criteria (and many other criteria) as special cases. The A - and D -
 93 criteria may respectively be written as $\Phi_A(\mathbf{M}) = \{\sum_{i=1}^m 1/\lambda_i(\mathbf{M})\}/m$, and
 94 $\Phi_D(\mathbf{M}) = \{\prod_{i=1}^m 1/\lambda_i(\mathbf{M})\}^{1/m}$; here, $\mathbf{M}$ may be the information matrix
 95 of interest or its upper bound to be specified latter in this paper, $\lambda_i(\mathbf{M})$
 96 is the i th greatest eigenvalue of $\mathbf{M}$, and m is the rank of the information

matrix that makes the parametric functions of interest estimable. When $\mathbf{M}$ is the information matrix of interest, minimizing Φ_A and Φ_D are respectively equivalent to minimizing the average variance and generalized variance of the estimates of the parametric functions of interest, and thus allows a precise statistical inference.

In the next section, we provide optimal and very efficient designs for Objective II with $Q = 2$. In Section 3, optimal designs for the same objective with $Q \geq 3$ are studied. We then investigate the performance of some m -sequence-based designs in Section 4 followed by a conclusion in Section 5.

2. Designs for $Q = 2$

Let $y_1, \dots, y_N$ be the BOLD signals collected by using an fMRI scanner to repeatedly scan a voxel of the subject's brain while a stimulus sequence $\mathbf{d} = (d_1, \dots, d_N)^T$ is presented to the subject. The following model is considered:

$$y_n = \gamma + \sum_{k=1}^K (x_{1,n-k+1} h_{1k} + x_{2,n-k+1} h_{2k}) + \varepsilon_n, \quad n = 1, \dots, N. \quad (1)$$

Here, γ is a unknown nuisance parameter, h_{qk} is the height of the HRF of the q th-type stimulus that is evaluated at the $(k - 1)$ th time point after the stimulus onset; $q = 1, 2$ and $k = 1, \dots, K$. The value of K is selected so that the HRF completely returns to baseline after K time points. The value of $x_{q,n}$ is determined by the design $\mathbf{d}$; it is 1 if $d_n = q$ and is 0 if $d_n \neq q$, $n = 1, \dots, N$. With the same assumption as in Kao (2015) and Cheng & Kao (2015), $x_{q,\ell}$ is set to $x_{q,N+\ell}$ when $\ell \leq 0$, and $\varepsilon_1, \dots, \varepsilon_N$ are considered as uncorrelated Gaussian noise with mean zero and variance σ^2 ; see also Kao (2014) for a discussion on the model assumption.

The HRF of the q th-type is represented by the unknown parameter vector $\mathbf{h}_q = (h_{q1}, \dots, h_{qK})^T$. With a focus on Objective II, our main design goal is to find an optimal design yielding the most precise estimate of the K -by-1 parameter vector $\boldsymbol{\theta} = \mathbf{h}_1 - \mathbf{h}_2$. Some such optimal designs have recently been obtained by Kao (2015) and Cheng & Kao (2015). However, these previous results are only for $N = 4t$ or $N = 4t \pm 1$. Here, we work on cases with $N = 4t + 2$ by considering the following equivalent model of model (1)

$$y_n = \gamma + \sum_{k=1}^K (a_{n,k} \eta_k + b_{n,k} \theta_k) + \varepsilon_n, \quad n = 1, \dots, N, \quad (2)$$

126 where $a_{n,k} = (x_{1,n-k+1} + x_{2,n-k+1})/2$, $b_{n,k} = (x_{1,n-k+1} - x_{2,n-k+1})/2$, $\eta_k =$
 127 $h_{1k} + h_{2k}$, and $\theta_k = h_{1k} - h_{2k}$ is the k th element of the parameter vector
 128 $\boldsymbol{\theta}$. The remaining terms of (2) are as in (1). Instead of focusing on the
 129 popularly used Φ_A and Φ_D criteria, we will consider here the class of type 1
 130 criteria of Cheng (1978) when evaluating the goodness of competing designs.
 131 This class of criteria includes both Φ_A and Φ_D (or their equivalent forms)
 132 as special cases. For a non-negative definite matrix $\mathbf{M}$ of rank m , a type 1
 133 criterion has the form of $\Phi_f(\mathbf{M}) = \sum_{i=1}^m f\{\lambda_i(\mathbf{M})\}$, where f is a function
 134 on $[0, \infty)$ such that (i) f is continuously differentiable in $(0, \infty)$ with $f' < 0$,
 135 $f'' > 0$, and $f''' < 0$, and (ii) $\lim_{x \rightarrow 0^+} f(x) = f(0) = \infty$. Our aim is then
 136 at a design minimizing $\Phi_f\{\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})\}$ over $\mathcal{D} = \{0, 1, 2\}^N$. With a given
 137 $\mathbf{d} \in \mathcal{D}$, $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta}) = \mathbf{B}(\mathbf{d})^T w^\perp\{\mathbf{j}_N, \mathbf{A}(\mathbf{d})\} \mathbf{B}(\mathbf{d})$ is the information matrix
 138 of $\boldsymbol{\theta}$, where $\mathbf{A}(\mathbf{d})$ and $\mathbf{B}(\mathbf{d})$ are N -by- K matrices whose (n, k) th elements
 139 are $a_{n,k}$ and $b_{n,k}$, respectively, $\mathbf{j}_N$ is the vector of N ones, $[\mathbf{j}_N, \mathbf{A}(\mathbf{d})]$ is the
 140 N -by- $(K+1)$ matrix formed by $\mathbf{j}_N$ and $\mathbf{A}(\mathbf{d})$, $w^\perp(\mathbf{G}) = \mathbf{I}_\ell - w(\mathbf{G})$ for a
 141 matrix $\mathbf{G}$ of order ℓ , $w(\mathbf{G}) = \mathbf{G}(\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T$, and $\mathbf{I}_\ell$ is the identity matrix of
 142 order ℓ . It is noteworthy that $\mathbf{A}(\mathbf{d})$ and $\mathbf{B}(\mathbf{d})$ are circulant matrices whose
 143 $(k+1)$ th column is obtained by cyclically shifting the k th column one position
 144 downward. The elements of $\mathbf{A}(\mathbf{d})$ are either 0 or 0.5, and those of $\mathbf{B}(\mathbf{d})$ can
 145 be -0.5 , 0, or 0.5. The following tool due to Cheng (1978) facilitates the
 146 identification of optimal designs.

147 **Lemma 1.** *Let $\mathbf{M}^*$ be a symmetric matrix with eigenvalues $\lambda_1(\mathbf{M}^*) >$
 148 $\lambda_2(\mathbf{M}^*) = \dots = \lambda_m(\mathbf{M}^*) > 0$ and $\mathcal{M}$ be a set of non-negative definite
 149 symmetric matrices of order m . Then, $\mathbf{M}^*$ is optimal over $\mathcal{M}$ under any
 150 type 1 criterion if, for any $\mathbf{M} \in \mathcal{M}$, $\text{tr}(\mathbf{M}) \leq \text{tr}(\mathbf{M}^*)$ and*

$$\begin{aligned}
 & \text{tr}(\mathbf{M}) - \sqrt{[m/(m-1)][\text{tr}(\mathbf{M}^2) - \text{tr}(\mathbf{M})^2/m]} \\
 & \leq \text{tr}(\mathbf{M}^*) - \sqrt{[m/(m-1)][\text{tr}\{(\mathbf{M}^*)^2\} - \text{tr}(\mathbf{M}^*)^2/m]}.
 \end{aligned}$$

151 Directly applying Lemma 1 to the rather complex $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})$ is quite chal-
 152 lenging. To reduce the complexity of our problem, we obtain manageable
 153 upper bounds of $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})$ in Lemma 2. There, $n_q(\mathbf{d})$ is the number of ap-
 154 pearances of q in design $\mathbf{d}$, and $\leq_L$ represents the Löwner ordering with
 155 $\mathbf{G}_1 \leq_L \mathbf{G}_2$ when $\mathbf{G}_2 - \mathbf{G}_1$ is non-negative definite.

156 **Lemma 2.** *For any $\mathbf{d} \in \mathcal{D}$, we have (i) $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta}) \leq_L \mathbf{B}(\mathbf{d})^T \mathbf{B}(\mathbf{d})$; and (ii)
 157 if $n_0(\mathbf{d}) = 1$, $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta}) \leq_L \mathbf{B}_0(\mathbf{d})^T \mathbf{B}_0(\mathbf{d})$, where $\mathbf{B}_0(\mathbf{d})$ is an $(N-K)$ -by- K
 158 matrix formed by the $N-K$ rows of $\mathbf{B}(\mathbf{d})$ that have no zero element.*

159 PROOF. (i) is straightforward. For (ii), without loss of generality, we consider
 160 those $\mathbf{d}$'s with $d_1 = 0$ and $d_n = 1$ or 2 for $n \geq 2$. Following the fact that $\mathbf{A}(\mathbf{d})$
 161 is now an N -by- K circulant matrix obtained by cyclic shifts of the column
 162 vector $(0, \mathbf{j}_{N-1}^T/2)^T$, we have

$$w^\perp\{[\mathbf{j}_N, \mathbf{A}(\mathbf{d})]\} = \begin{bmatrix} \mathbf{O}_{K,K} & \mathbf{O}_{K,N-K} \\ \mathbf{O}_{N-K,K} & w^\perp(\mathbf{j}_{N-K}) \end{bmatrix},$$

163 where $\mathbf{O}_{a,b}$ is the a -by- b zero matrix. Thus,

$$\mathbf{M}(\mathbf{d}; \boldsymbol{\theta}) = \mathbf{B}(\mathbf{d})^T w^\perp([\mathbf{j}_N, \mathbf{A}(\mathbf{d})]) \mathbf{B}(\mathbf{d}) = \mathbf{B}_0(\mathbf{d})^T w^\perp(\mathbf{j}_{N-K}) \mathbf{B}_0(\mathbf{d}) \leq_L \mathbf{B}_0(\mathbf{d})^T \mathbf{B}_0(\mathbf{d}),$$

164 where $\mathbf{B}_0(\mathbf{d})$ is the lower $(N-K)$ -by- K sub-matrix of $\mathbf{B}(\mathbf{d})$ that has no zero
 165 element.

166 We now present our first main result of this section. We note that, when
 167 identifying optimal designs for $N = 4t + 2$, we utilize the same tool (Lemma
 168 1) as that for $N = 4t + 1$ of Cheng & Kao (2015). Nevertheless, our derivation
 169 here is slightly more involved than that of the latter case due to the need for
 170 using different upper bounds of the information matrix for different values
 171 of $n_0(\mathbf{d})$; see also Lemma 2. Our proof of Theorem 1 is then complete by
 172 considering all the possible values of $n_0(\mathbf{d})$.

173 **Theorem 1.** For $N = 4t + 2$ and $K \geq 2$, let $\mathbf{d}^* \in \mathcal{D}$ be such that $\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta}) =$
 174 $\{(N-2)\mathbf{I}_K + 2\mathbf{J}_K\}/4$, where $\mathbf{J}_K = \mathbf{j}_K \mathbf{j}_K^T$. Then, for any type 1 criterion,
 175 $\mathbf{d}^*$ is optimal for estimating $\boldsymbol{\theta} = (\theta_1, \dots, \theta_K)^T$ of model (2).

176 PROOF. To show that $\Phi_f\{\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})\} = \min_{\mathbf{d} \in \mathcal{D}} \Phi_f\{\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})\}$ for any Φ_f , we
 177 consider three cases. In Case 1, we work on the subclass $\mathcal{D}_1 = \{1, 2\}^N \subset \mathcal{D}$
 178 of designs having no zero. We then focus on $\mathcal{D}_2 = \{\mathbf{d} \in \mathcal{D} : n_0(\mathbf{d}) = 1\}$ in
 179 Case 2, and $\mathcal{D}_3 = \{\mathbf{d} \in \mathcal{D} : n_0(\mathbf{d}) \geq 2\}$ in Case 3. In all the three cases, we
 180 set $\mathbf{M}^*$ in Lemma 1 to $\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})$, and $m = K$. The matrix $\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})$ has
 181 two positive eigenvalues and the greater one has multiplicity 1.

182 Case 1. For any $\mathbf{d} \in \mathcal{D}_1$, the corresponding $\mathbf{B}(\mathbf{d})$ has $\text{tr}\{\mathbf{B}(\mathbf{d})^T \mathbf{B}(\mathbf{d})\} =$
 183 $\text{tr}\{\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})\}$, and each off-diagonal element of $\mathbf{B}(\mathbf{d})^T \mathbf{B}(\mathbf{d})$ is $(4s + 2)/4$
 184 for some integer s ; see also Lemma 2.2 of Cheng & Kao (2015). We thus
 185 have $\text{tr}\{\mathbf{M}^2(\mathbf{d}^*; \boldsymbol{\theta})\} = \min_{\mathbf{d} \in \mathcal{D}_1} \text{tr}[\mathbf{B}(\mathbf{d})^T \mathbf{B}(\mathbf{d})]^2$. The conditions in Lemma
 186 1 are thus satisfied by replacing $\mathbf{M}$ with $\mathbf{B}(\mathbf{d})^T \mathbf{B}(\mathbf{d})$, and $\mathcal{M}$ with $\mathcal{D}_1$. In
 187 addition, with Lemma 2(i), we then have for any type 1 criterion Φ_f

$$\Phi_f\{\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})\} = \min_{\mathbf{d} \in \mathcal{D}_1} \Phi_f\{\mathbf{B}(\mathbf{d})^T \mathbf{B}(\mathbf{d})\} \leq \min_{\mathbf{d} \in \mathcal{D}_1} \Phi_f\{\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})\}.$$

Case 2. While Lemma 2(i) remains true for this case, the sharper upper bound of $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})$ provided in Lemma 2(ii) is needed. It then can be seen that, for any $\mathbf{d} \in \mathcal{D}_2$,

$$\Phi_f\{(N-K)\mathbf{I}_K/4\} \leq \Phi_f\{\mathbf{B}_0(\mathbf{d})^T \mathbf{B}_0(\mathbf{d})\} \leq \Phi_f\{\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})\}.$$

In addition, with $K \geq 2$, the condition in Lemma 1 is satisfied when we set $\mathbf{M} = (N-K)\mathbf{I}_K/4$, and $\mathcal{M}$ contains only this $\mathbf{M}$. Thus, for any $\mathbf{d} \in \mathcal{D}_2$,

$$\Phi_f\{\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})\} \leq \Phi_f\{(N-K)\mathbf{I}_K/4\} \leq \Phi_f\{\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})\}.$$

Case 3. Since $n_0(\mathbf{d}) \geq 2$ for any $\mathbf{d} \in \mathcal{D}_3$, we have

$$\begin{aligned} \text{tr}\{\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})\} - \sqrt{[K/(K-1)][\text{tr}\{\mathbf{M}^2(\mathbf{d}^*; \boldsymbol{\theta})\} - \text{tr}\{\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})\}^2/K]} \\ = K(N-2)/4 \geq K\{N - n_0(\mathbf{d})\}/4 = \text{tr}\{\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})\}. \end{aligned}$$

With Lemmas 1 and 2, this implies $\Phi_f\{\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})\} \leq \Phi_f\{\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})\}$ for any $\mathbf{d} \in \mathcal{D}_3$ and Φ_f .

Remark 1. With a proof similar to Case 1, we can show that designs $(\mathbf{d}^* - \mathbf{j}_N)$ and $(2\mathbf{j}_N - \mathbf{d}^*)$ in $\{0, 1\}^N$ are optimal under any Φ_f -criterion for estimating the HRF $\mathbf{h}_1 = (h_{11}, \dots, h_{1K})^T$, i.e. Objective I, when $Q = 1$ and $N = 4t + 2$.

The following corollary of Theorem 1 provides an additional insight into the structure of the optimal design $\mathbf{d}^*$. There, for any design $\mathbf{d}$, we define $n_{k,p,q}(\mathbf{d}) = \text{card}\{j : (d_{j-k}, d_j) = (q, p), j = 1, \dots, N\}$, where d_j is interpreted as d_{N+j} if $j \leq 0$, and $\text{card}\{\cdot\}$ is the cardinal number. If a design with the specified structure is found, it is optimal under any Φ_f criterion. For cases where such a design is unavailable, our result can still be used to obtain a lower bound for the efficiencies of other designs.

Corollary 1. A design $\mathbf{d}^* \in \mathcal{D}$ of length $N = 4t + 2$ is optimal under any type 1 criterion for comparing the two HRFs if, for some integer t ,

$$n_{k,1,1}(\mathbf{d}^*) - 1 = n_{k,2,2}(\mathbf{d}^*) - 1 = n_{k,1,2}(\mathbf{d}^*) = n_{k,2,1}(\mathbf{d}^*) = t, \quad k = 1, \dots, K-1.$$

PROOF. For any $\mathbf{d} \in \mathcal{D}$, the (i, j) th element of $\mathbf{B}(\mathbf{d})^T \mathbf{B}(\mathbf{d})$ is $\{N - n_0(\mathbf{d})\}/4$ for $i = j$, and is $\{n_{k,1,1}(\mathbf{d}) + n_{k,2,2}(\mathbf{d}) - n_{k,1,2}(\mathbf{d}) - n_{k,2,1}(\mathbf{d})\}/4$ for $|i - j| = k \geq 1$. Now, $n_1(\mathbf{d}^*) = n_{k,1,1}(\mathbf{d}^*) + n_{k,1,2}(\mathbf{d}^*) = 2t + 1$, and $n_2(\mathbf{d}^*) = n_{k,2,1}(\mathbf{d}^*) + n_{k,2,2}(\mathbf{d}^*) = 2t + 1$. We have $n_0(\mathbf{d}^*) = N - n_1(\mathbf{d}^*) - n_2(\mathbf{d}^*) = 0$. Consequently $\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta}) = \mathbf{B}(\mathbf{d}^*)^T \mathbf{B}(\mathbf{d}^*) = \{(N-2)\mathbf{I}_K + 2\mathbf{J}_K\}/4$. The optimality of $\mathbf{d}^*$ then follows from Theorem 1.

215 Kao (2014) and Cheng & Kao (2015) considered Hadamard sequences
 216 for constructing optimal fMRI designs. A Hadamard sequence, $\mathbf{d}_H$, is ob-
 217 tained from a normalized Hadamard matrix that has a circulant core. Such
 218 a Hadamard matrix has the form of

$$\mathbf{H} = \begin{bmatrix} 1 & \mathbf{j}_\ell^T \\ \mathbf{j}_\ell & \tilde{\mathbf{H}} \end{bmatrix},$$

219 where $\tilde{\mathbf{H}} \in \{\pm 1\}^{\ell \times \ell}$ is a circulant matrix, called the circulant core of H , with
 220 $\tilde{\mathbf{H}}^T \tilde{\mathbf{H}} = (\ell + 1)\mathbf{I}_\ell - \mathbf{J}_\ell$ and $\mathbf{j}_\ell^T \tilde{\mathbf{H}} = -\mathbf{j}_\ell^T$. Any column of $(\mathbf{j}_\ell - \tilde{\mathbf{H}})/2$ can be
 221 used as a $\mathbf{d}_H$. Such an $\mathbf{H}$, and thus $\tilde{\mathbf{H}}$ and $\mathbf{d}_H$, are known to exist when ℓ
 222 is a prime, a product of twin primes, or $2^s - 1$ for an integer s . They can be
 223 generated by the Paley, Singer, or twin-prime-power difference sets (Golomb
 224 & Gong, 2005; Horadam, 2007).

225 For convenience, we hereinafter use 1 and 2 to respectively replace the
 226 two symbols 0 and 1 in $\mathbf{d}_H$, and set $\ell = N - 3 = 4t - 1$. For a given g , we also
 227 obtain $\mathbf{d}_{iH}$ by adding i 1's to any subsequence of g 1's in $\mathbf{d}_H$. The statistical
 228 optimality of $\mathbf{d}_H$, $\mathbf{d}_{1H}$ and $\mathbf{d}_{2H}$ is established in Cheng & Kao (2015), but
 229 our Theorem 1 suggests that $\mathbf{d}_{3H}$ can be sub-optimal. Nevertheless, the $\mathbf{d}_{3H}$
 230 still gives very efficient fMRI designs for $K \leq g + 1$. This can be seen by the
 231 fact that, for $K \leq g + 1$,

$$\mathbf{M}(\mathbf{d}_{3H}; \theta) = \frac{(N - 2)\mathbf{I}_K + (2 - 4/N)\mathbf{J}_K}{4}.$$

232 We can then use the possibly hypothetical $\mathbf{d}^*$ to obtain a lower bound for
 233 the A - and D -efficiencies of $\mathbf{d}_{3H}$. That is

$$A_{eff}(\mathbf{d}_{3H}) \geq \frac{\text{tr}\{\mathbf{M}^{-1}(\mathbf{d}^*; \theta)\}}{\text{tr}\{\mathbf{M}^{-1}(\mathbf{d}_{3H}; \theta)\}} = 1 - \frac{4}{(N - 2 + 2K)^2};$$

234

$$D_{eff}(\mathbf{d}_{3H}) \geq \left\{ \frac{|\mathbf{M}(\mathbf{d}_{3H}; \theta)|}{|\mathbf{M}(\mathbf{d}^*; \theta)|} \right\}^{1/K} = \left\{ 1 - \frac{4K}{N(N - 2 + 2K)} \right\}^{1/K}$$

235 In typical fMRI experiments, N is tens or hundreds and is much greater than
 236 K . Both $A_{eff}(\mathbf{d}_{3H})$ and $D_{eff}(\mathbf{d}_{3H})$ are thus at least 99% in most cases as
 237 presented in the following example.

238 **Example 1.** *We consider an experimental setting similar to that in the Ex-*
 239 *ample 2.1 of Cheng & Kao (2015) where $(N, K) = (134, 9)$, a stimulus can*

possibly occur every 4 seconds, the duration of the experiment is just short of 9 minutes, and the HRF evoked by each stimulus is assumed to be 32 seconds, counting from the onset of the stimulus to the HRF's complete return to baseline. Suppose the experiment involves two types of stimuli, and the experimenter is interested in comparing the HRFs of the two stimulus types. For this study objective, we construct the following sequence $\mathbf{d}_{3H}$ by adding three 1's to a subsequence of $g(=8)$ 1's in a Hadamard sequence of length $N-3=131$.

$$\mathbf{d}_{3H} = \begin{pmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 & 1 & 2 & 1 & 2 & 2 & 1 & 1 & 2 & 2 & 1 & 2 & 2 & 2 & 2 & 1 & 2 & 1 & 2 & 2 \\ 1 & 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 2 & 2 & 1 & 2 & 1 & 1 & 1 & 2 & 2 & 1 & 1 & 1 & 2 & 2 & 1 & 2 & 2 & 2 & 1 \\ 2 & 1 & 2 & 1 & 2 & 2 & 2 & 1 & 2 & 2 & 1 & 2 & 1 & 1 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 1 & 1 & 2 & 1 & 1 & 2 & 1 \\ 2 & 1 & 1 & 2 & 2 & 2 & 1 & 2 & 1 & 1 & 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 2 & 1 & 1 & 2 & 1 & 2 & 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 2 & 2 & 1 & 1 & 2 & 1 & 2 & 2 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}^T.$$

Both the A- and D-efficiencies of this $\mathbf{d}_{3H}$ are at least 99.98%.

A systematic construction of an optimal design $\mathbf{d}^*$ for $N = 4t + 2$ is also possible by slightly modifying the above-mentioned method. We have the following simple two-step procedure to construct such an optimal design.

- Step 1 Obtain a Hadamard sequence $\mathbf{d}_H \in \{1, 2\}^{N-3}$. For this $\mathbf{d}_H$ or its cyclic shifts, identify a subsequence of g_1 1's followed by g_2 2's that yields the maximum $g = \min(g_1, g_2)$.
- Step 2 At the end of the g_1 1's, insert $(1, 1, 2)^T$ to obtain a design $\mathbf{d}_{3H}^*$ of length N .

We now prove the statistical optimality of the $\mathbf{d}_{3H}^*$ obtained by the two-step procedure.

Theorem 2. *The $\mathbf{d}_{3H}^* \in \{1, 2\}^N$ obtained by the previously described two-step procedure is optimal for comparing two HRFs with $K \leq g + 1$ under any type 1 criterion.*

PROOF. Without loss of generality, we consider a $\mathbf{d}_H$ of length $(N-3) = 4t-1$ that has g_1 1's at the tail and g_2 2's on top and $g_2 \leq g_1$ so that $g = g_2 = \min\{g_1, g_2\}$. Adding two 1's to the end of this $\mathbf{d}_H$ yields a $\mathbf{d}_{2H}$ of length $N = 4t + 1$ that has (see also Cheng & Kao, 2015)

$$n_{k,1,1}(\mathbf{d}_{2H}) - 1 = n_{k,2,2}(\mathbf{d}_{2H}) = n_{k,1,2}(\mathbf{d}_{2H}) = n_{k,2,1}(\mathbf{d}_{2H}) = t, \quad k = 1, \dots, g_1.$$

267 With an additional 2 at the end of $\mathbf{d}_{2H}$, it can be seen that $\mathbf{d}_{3H}^*$ has $n_{k,1,1}(\mathbf{d}_{3H}^*) =$
 268 $n_{k,1,1}(\mathbf{d}_{2H}) = t + 1$ for $k \leq g(= g_2)$, but $n_{g+1,1,1}(\mathbf{d}_{3H}^*) = n_{g+1,1,1}(\mathbf{d}_{2H}) - 1 = t$.
 269 Moreover, for $q = 1, 2$,

$$N/2 = 2t + 1 = n_q(\mathbf{d}_{3H}^*) = n_{k,q,1}(\mathbf{d}_{3H}^*) + n_{k,q,2}(\mathbf{d}_{3H}^*) = n_{k,1,q}(\mathbf{d}_{3H}^*) + n_{k,2,q}(\mathbf{d}_{3H}^*).$$

270 Thus, for $k \leq g$,

$$\begin{aligned} n_{k,1,2}(\mathbf{d}_{3H}^*) &= n_{k,2,1}(\mathbf{d}_{3H}^*) = n_1(\mathbf{d}_{3H}^*) - n_{k,1,1}(\mathbf{d}_{3H}^*) = 2t + 1 - (t + 1) = t; \\ n_{k,2,2}(\mathbf{d}_{3H}^*) &= n_2(\mathbf{d}_{3H}^*) - n_{k,2,1}(\mathbf{d}_{3H}^*) = 2t + 1 - t = t + 1. \end{aligned}$$

271 The condition in Corollary 1 is thus satisfied when $K \leq g + 1$. Our claim
 272 then follows.

273 **Remark 2.** Following the previous argument, a slightly more flexible method
 274 for constructing optimal designs is to insert two 1's into any subsequence of
 275 g_1 1's and a 2 into any subsequence of g_2 2's in a $\mathbf{d}_H$. Here, the two subse-
 276 quences of 1's and 2's may be separated.

277 Theorem 2 suggests that, by using a 2 to replace the last element of $\mathbf{d}_{3H}$ in
 278 Example 1, we obtain an optimal design $\mathbf{d}_{3H}^*$ for estimating θ when $N = 134$
 279 and $K \leq 9$. For cases with $K > g + 1$, we may again use Theorem 1 to
 280 calculate a lower bound for the A - and D -efficiencies of $\mathbf{d}_{3H}$ and $\mathbf{d}_{3H}^*$. Our
 281 experience suggests that both types of designs are often highly efficient for
 282 practical (N, K) -values, although $\mathbf{d}_{3H}$ outperforms $\mathbf{d}_{3H}^*$ in some cases while
 283 the latter design is better than the former one in other cases. For $K > g + 1$,
 284 an evaluation of the criterion values of $\mathbf{d}_{3H}$ and $\mathbf{d}_{3H}^*$ is thus recommended
 285 before selecting in between these two designs.

286 The designs $\mathbf{d}_{3H}^*$ belong to the design class $\mathcal{D}_0 = \{\mathbf{d} \in \mathcal{D} : n_1(\mathbf{d}) =$
 287 $n_2(\mathbf{d}) = N/2 \text{ \& } |N - 4n_{k,1,2}(\mathbf{d})| = 2, k = 1, \dots, K - 1\}$, which includes
 288 all the designs whose information matrix has diagonal elements $N/4$ and off-
 289 diagonal elements ± 0.5 . For cases where $\mathbf{d}_{3H}^*$ is unavailable, $\mathcal{D}_0$ may contain
 290 other designs for experimenters to choose from; see Lin et al. (2016) for a
 291 construction of such designs. This is especially true when $N/2 = 2t + 1$ is
 292 an odd prime. For this case, we always can find $\mathbf{d}^\dagger \in \mathcal{D}_0$ whose information
 293 matrix is $\mathbf{M}(\mathbf{d}^\dagger; \theta) = \{(N + 2)\mathbf{I}_K - 2\mathbf{J}_K\}/4$ whenever $K \leq N/2$ (Lin et al.,
 294 2016). As shown in the next proposition which is a direct consequence of
 295 the results of Cheng (1978), $\mathbf{d}^\dagger$ is actually the worst design in $\mathcal{D}_0$ under any
 296 type 1 criterion. Nevertheless, our numerical results to be presented below

also suggest that $\mathbf{d}^\dagger$ is already highly efficient for typical fMRI experimental settings. Our results here thus establish the good performance of all the designs in $\mathcal{D}_0$, which are recommended for practical use.

Proposition 1. For $N = 4t + 2$ and $1 < K < 2t + 2$, let $\mathbf{d}^\dagger \in \mathcal{D}_0$ be such that $\mathbf{M}(\mathbf{d}^\dagger; \boldsymbol{\theta}) = \{(N + 2)\mathbf{I}_K - 2\mathbf{J}_K\}/4$, or equivalently,

$$n_{k,1,1}(\mathbf{d}^\dagger) = n_{k,2,2}(\mathbf{d}^\dagger) = n_{k,1,2}(\mathbf{d}^\dagger) - 1 = n_{k,2,1}(\mathbf{d}^\dagger) - 1 = t, \quad k = 1, \dots, K - 1.$$

Then, for any type 1 criterion Φ_f and any $\mathbf{d} \in \mathcal{D}_0$, we have $\Phi_f\{\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})\} \leq \Phi_f\{\mathbf{M}(\mathbf{d}^\dagger; \boldsymbol{\theta})\}$.

PROOF. As suggested by Cheng (1978), a design whose information matrix $\mathbf{M}$ has two different positive eigenvalues with the smaller one being simple is the ‘ Φ_f -worst’ over the class of designs attaining the same $\text{tr}\{\mathbf{M}\}$ and $\text{tr}\{\mathbf{M}^2\}$ (see also Kunert, 1985). Now, $\mathbf{M}(\mathbf{d}^\dagger; \boldsymbol{\theta})$ has eigenvalues $(N + 2)$ and $(N + 2 - 2K)$ with the latter one having multiplicity 1, and for any $\mathbf{d} \in \mathcal{D}_0$, $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})$ has diagonal elements of $N/4$ and off-diagonal elements of ± 0.5 . Thus, $\text{tr}\{\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})\} = \text{tr}\{\mathbf{M}(\mathbf{d}^\dagger; \boldsymbol{\theta})\}$, and $\text{tr}\{\mathbf{M}^2(\mathbf{d}; \boldsymbol{\theta})\} = \text{tr}\{\mathbf{M}^2(\mathbf{d}^\dagger; \boldsymbol{\theta})\}$. Our claim then follows.

With Theorem 1, we also have $\Phi_f\{\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})\} \geq \Phi_f\{\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})\}$ for any design $\mathbf{d} \in \mathcal{D}$. A lower bound of A - and D -efficiencies for any $\mathbf{d} \in \mathcal{D}_0$ can then be obtained by this fact and Proposition 1. We thus have, for any $\mathbf{d} \in \mathcal{D}_0$,

$$A_{eff}(\mathbf{d}) \geq \frac{\text{tr}\{\mathbf{M}^{-1}(\mathbf{d}^*; \boldsymbol{\theta})\}}{\text{tr}\{\mathbf{M}^{-1}(\mathbf{d}^\dagger; \boldsymbol{\theta})\}} = \frac{(N + 2)(N + 2 - 2K)\{N - 2 + (K - 1)(N - 2 + 2K)\}}{(N - 2)(N - 2 + 2K)\{N + 2 + (K - 1)(N + 2 - 2K)\}},$$

316

$$D_{eff}(\mathbf{d}) \geq \left\{ \frac{|\mathbf{M}(\mathbf{d}^\dagger; \boldsymbol{\theta})|}{|\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})|} \right\}^{1/K} = \left(\frac{N + 2}{N - 2} \right) \left\{ \frac{(N - 2)(N + 2 - 2K)}{(N + 2)(N - 2 + 2K)} \right\}^{1/K}.$$

It can then be seen that any design $\mathbf{d} \in \mathcal{D}_0$ attains at least 99% A - and D -efficiencies for $K = 9$ when $N \geq 50$, and for $K = 17$ when $N \geq 78$. The former value of K is used in Example 1. It corresponds to the typically assumed 32-second HRFs with a discretization interval of 4 seconds; that is the time between two consecutive fMRI scans of the same voxel is 4 seconds. The latter value of K is for the same HRF duration with a 2-second discretization interval. These two settings are common in practice.

324 3. Designs for $Q \geq 3$

325 For experiments with $Q (\geq 3)$ types of stimuli, model (1) is extended to

$$y_n = \gamma + \sum_{k=1}^K (x_{1,n-k+1}h_{1k} + x_{2,n-k+1}h_{2k} + \cdots + x_{Q,n-k+1}h_{Qk}) + \varepsilon_n, \quad n = 1, \dots, N. \quad (3)$$

326 For Objective II, we consider the standardized contrasts $\boldsymbol{\theta}$ of the HRFs,
 327 where $\boldsymbol{\theta} = \mathbf{E}^T \mathbf{h}$, $\mathbf{E} = w^\perp(\mathbf{j}_Q) \otimes \mathbf{I}_K$, $\otimes$ is the Kronecker product, and
 328 $\mathbf{h} = (\mathbf{h}_1^T, \dots, \mathbf{h}_Q^T)^T$. That is, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \dots, \boldsymbol{\theta}_Q^T)^T$ with $\boldsymbol{\theta}_q = \mathbf{h}_q - \bar{\mathbf{h}}$, and $\bar{\mathbf{h}} =$
 329 $(\mathbf{h}_1 + \mathbf{h}_2 + \cdots + \mathbf{h}_Q)/Q$. We would like to find a design $\mathbf{d} = (d_1, \dots, d_N)^T \in \tilde{\mathcal{D}}$
 330 that yields the most precise estimate of $\boldsymbol{\theta}$; here, $\tilde{\mathcal{D}} \subset \{0, 1, \dots, Q\}^N$ is the
 331 class of designs making $\boldsymbol{\theta}$ estimable. To this end, we will use the Schur
 332 ordering to rank the information matrices of competing designs (see also
 333 Cheng, 1995). Our design goal is then to identify a $\mathbf{d}^*$ such that $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta}) \leq_S$
 334 $\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})$ for any other $\mathbf{d} \in \tilde{\mathcal{D}}$. Here, $\leq_S$ represents the Schur ordering, and
 335 for any two symmetric, non-negative definite matrices $\mathbf{G}_1$ and $\mathbf{G}_2$ of order
 336 m , a necessary and sufficient condition for $\mathbf{G}_1 \leq_S \mathbf{G}_2$ is that there exists a
 337 $\mathbf{G} \in \text{conv}\{\mathbf{W}\mathbf{G}_1\mathbf{W}^T : \mathbf{W} \in \mathcal{O}(m)\}$ such that $\mathbf{G} \leq_L \mathbf{G}_2$, where $\text{conv}\{\cdot\}$
 338 denotes the convex hull, and $\mathcal{O}(m)$ is the group of orthogonal matrices of
 339 order m ; see Harman (2008) and references therein for further details. When
 340 such a design $\mathbf{d}^*$ exists, it can be shown optimal for estimating $\boldsymbol{\theta}$ under [any](#)
 341 [Schur-convex, componentwise-decreasing function of the eigenvalues of the](#)
 342 [information matrix \(see also Cheng, 1995\); these functions include \$\Phi_A\$ and](#)
 343 [\$\Phi_D\$.](#)

344 To achieve our design goal, we find it useful to consider the following two
 345 lemmas, which provide manageable upper bounds of $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta})$. Lemma 3 is
 346 due to Bailey & Druilhet (2004), and is derived under the Löwner ordering;
 347 it is rephrased based on our setting. The bound provided in Lemma 4 is
 348 obtained with the Schur ordering.

349 **Lemma 3.** $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta}) \leq_L \mathbf{F}(\mathbf{d}) = [\mathbf{E}^T \mathbf{E}]^+ \mathbf{E}^T \mathbf{M}(\mathbf{d}; \mathbf{h}) \mathbf{E} [\mathbf{E}^T \mathbf{E}]^+$, where $\boldsymbol{\theta} =$
 350 $\mathbf{E}^T \mathbf{h}$, $[\mathbf{E}^T \mathbf{E}]^+$ is the Moore-Penrose inverse of $\mathbf{E}^T \mathbf{E}$, and $\mathbf{M}(\mathbf{d}; \mathbf{h})$ is the
 351 information matrix for the parameter vector $\mathbf{h}$. The equality holds if and
 352 only if $\mathbf{M}(\mathbf{d}; \mathbf{h})$ commutes with $w(\mathbf{E})$.

353 **Lemma 4.** Let $\mathcal{G}$ be the symmetric group of all the $Q!$ permutations on
 354 $\{1, 2, \dots, Q\}$. Suppose $\mathbf{d} \in \{0, 1, \dots, Q\}^N$ and $\mathbf{d}_g$ is obtained from $\mathbf{d}$ by per-

355 muting the labels $1, \dots, Q$ based on a $g \in \mathcal{G}$, and keeping 0's intact. Then,
 356 $\mathbf{F}(\mathbf{d}) \leq_S \bar{\mathbf{F}}(\mathbf{d}) = \sum_{g \in \mathcal{G}} \mathbf{F}(\mathbf{d}_g)/Q!$, where $\mathbf{F}(\mathbf{d})$ is defined as in Lemma 3.

357 PROOF. With $\mathbf{E} = w^\perp(\mathbf{j}_Q) \otimes \mathbf{I}_K$, we have $\mathbf{F}(\mathbf{d}) = \mathbf{E}\mathbf{M}(\mathbf{d}; \mathbf{h})\mathbf{E}$. Let $\mathbf{W}_g$
 358 be the Q -by- Q permutation matrix corresponding to $g \in \mathcal{G}$. Then,

$$\mathbf{F}(\mathbf{d}_g) = \mathbf{E}\mathbf{M}(\mathbf{d}_g; \mathbf{h})\mathbf{E} = \mathbf{E}(\mathbf{W}_g \otimes \mathbf{I}_K)\mathbf{M}(\mathbf{d}; \mathbf{h})(\mathbf{W}_g^T \otimes \mathbf{I}_K)\mathbf{E} = (\mathbf{W}_g \otimes \mathbf{I}_K)\mathbf{F}(\mathbf{d})(\mathbf{W}_g^T \otimes \mathbf{I}_K).$$

359 Our claim then follows from the fact that $(\mathbf{W}_g \otimes \mathbf{I}_K)$ is an orthogonal matrix.

360 To derive $\bar{\mathbf{F}}(\mathbf{d})$ for $\mathbf{d} \in \tilde{D}$, we utilize the following result, which is a
 361 simple extension of Lemma 3.1 of Shah et al. (2005). The proof is omitted.

362 **Lemma 5.** Let $\mathcal{G}$ and $\mathbf{W}_g$ be defined as in Lemma 4, and $\mathbf{M} = ((\mathbf{M}_{p,q}))_{p,q=1,\dots,Q}$
 363 be a QK -by- QK partitioned matrix, where $\mathbf{M}_{p,q}$ is of order K . Then,

$$\bar{\mathbf{M}} = \sum_{g \in \mathcal{G}} \frac{(\mathbf{W}_g \otimes \mathbf{I}_K)\mathbf{M}(\mathbf{W}_g^T \otimes \mathbf{I}_K)}{Q!} = \mathbf{I}_Q \otimes (\mathbf{M}_0 - \mathbf{M}_1) + \mathbf{j}_Q \otimes \mathbf{M}_1,$$

364 where $\mathbf{M}_0 = \sum_{q=1}^Q \mathbf{M}_{q,q}/Q$, and $\mathbf{M}_1 = (\sum_{p=1}^Q \sum_{q=1}^Q \mathbf{M}_{p,q} - Q\mathbf{M}_0)/\{Q(Q -$
 365 $1)\}$.

366 With model (3), we also have $\mathbf{M}(\mathbf{d}; \mathbf{h}) = ((\mathbf{M}_{d,pq}))_{p,q=1,\dots,Q}$, where

$$\mathbf{M}_{d,pq} = \begin{bmatrix} 0 & n_{1,p,q}(\mathbf{d}) & \cdots & n_{K-1,p,q}(\mathbf{d}) \\ n_{1,q,p}(\mathbf{d}) & 0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & n_{1,p,q}(\mathbf{d}) \\ n_{K-1,q,p}(\mathbf{d}) & \cdots & n_{1,q,p}(\mathbf{d}) & 0 \end{bmatrix} + \delta_{pq}n_p(\mathbf{d})\mathbf{I}_K - \frac{n_p(\mathbf{d})n_q(\mathbf{d})}{N}\mathbf{J}_K,$$

367 where δ_{pq} is 1 if $p = q$ and is 0 if $p \neq q$. With some algebra, we then have
 368 $\bar{\mathbf{F}}(\mathbf{d}) = w^\perp(\mathbf{j}_Q) \otimes \bar{\mathbf{A}}(\mathbf{d})$, where

$$\bar{\mathbf{A}}(\mathbf{d}) = \begin{bmatrix} \alpha_0 & \alpha_1 & \cdots & \alpha_{K-1} \\ \alpha_1 & \alpha_0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \alpha_1 \\ \alpha_{K-1} & \cdots & \alpha_1 & \alpha_0 \end{bmatrix} - \beta\mathbf{J}_K,$$

369 $\alpha_0 = \{N - n_0(\mathbf{d})\}/Q$, $\alpha_k = \{Q \sum_{q=1}^Q n_{k,q,q}(\mathbf{d}) - N + 2n_0(\mathbf{d}) - n_{k,0,0}(\mathbf{d})\}/\{Q(Q -$
 370 $1)\}$, $k = 1, \dots, K - 1$, and $\beta = [Q \sum_{q=1}^Q n_q(\mathbf{d})^2 - \{N - n_0(\mathbf{d})\}^2]/\{NQ(Q - 1)\}$.
 371 We now provide an optimal design for estimating $\boldsymbol{\theta}$ in the following theorem.

372 **Theorem 3.** Let $\mathbf{d}^* \in \tilde{\mathcal{D}}$ be such that $\mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta}) = w^\perp(\mathbf{j}_Q) \otimes (N/Q)\mathbf{I}_K$.
 373 Then, $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta}) \leq_S \mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})$ for any other $\mathbf{d} \in \tilde{\mathcal{D}}$.

374 **PROOF.** For any $\mathbf{d} \in \tilde{\mathcal{D}}$, the average of the orthogonal transformations of
 375 $\bar{\mathbf{A}}(\mathbf{d})$ is $(\alpha_0 - \beta)\mathbf{I}_K$ (see also Cheng, 1985). Since $(\alpha_0 - \beta)\mathbf{I}_K \leq_L (N/Q)\mathbf{I}_K$,
 376 it can be seen that

$$\bar{\mathbf{F}}(\mathbf{d}) = w^\perp(\mathbf{j}_Q) \otimes \bar{\mathbf{A}}(\mathbf{d}) \leq_S w^\perp(\mathbf{j}_Q) \otimes (N/Q)\mathbf{I}_K = \mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta}).$$

377 With Lemmas 3 and 4 and the fact that $\mathbf{G}_1 \leq_L \mathbf{G}_2$ implies $\mathbf{G}_1 \leq_S \mathbf{G}_2$, we
 378 then have

$$\mathbf{M}(\mathbf{d}; \boldsymbol{\theta}) \leq_S \mathbf{F}(\mathbf{d}) \leq_S \bar{\mathbf{F}}(\mathbf{d}) \leq_S \mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})$$

379 for any $\mathbf{d} \in \tilde{\mathcal{D}}$. This completes the proof.

380 Following the argument at the beginning of this session, Theorem 3 im-
 381 plies the optimality of $\mathbf{d}^*$ under many optimality criteria, including A - and
 382 D -criteria. In other words, $\mathbf{d}^*$ simultaneously minimizes the average variance
 383 and generalized variance of the estimates of the standard contrasts $\boldsymbol{\theta} = \mathbf{E}^T \mathbf{h}$
 384 of the HRFs. This same $\mathbf{d}^*$ can also be shown to be universally optimal
 385 when the focus is on each $\tilde{\boldsymbol{\theta}}_k = (\theta_{1,k}, \dots, \theta_{Q,k})^T$, where $\theta_{q,k} = h_{q,k} - \bar{h}_k$, and
 386 $\bar{h}_k = \sum_{q=1}^Q h_{q,k}/Q$. To see this, we may set $\mathbf{E}$ to $w^\perp(\mathbf{j}_Q) \otimes \mathbf{e}_k^T$ in Lemma
 387 3, where $\mathbf{e}_k^T$ is the k th row of $\mathbf{I}_K$. It then can be seen that $\mathbf{M}(\mathbf{d}^*; \tilde{\boldsymbol{\theta}}_k)$
 388 has the maximum trace and is completely symmetric. The universal opti-
 389 mality of $\mathbf{d}^*$ then follows from Proposition 1 of Kiefer (1975). This guar-
 390 antees the A - and D -optimality of $\mathbf{d}^*$ for studying the standard contrasts
 391 of the k th heights of the HRFs. In addition, when a design $\mathbf{d} \in \tilde{\mathcal{D}}$ has
 392 $n_{k,p,q}(\mathbf{d}) = N/Q^2$ for any $k = 1, \dots, K-1$ and $p, q = 1, \dots, Q$, one can show
 393 that $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta}) = \mathbf{M}(\mathbf{d}; \mathbf{h}) = \mathbf{F}(\mathbf{d}) = \bar{\mathbf{F}}(\mathbf{d}) = w^\perp(\mathbf{j}_Q) \otimes (N/Q)\mathbf{I}_K$, and thus,
 394 $\mathbf{d}$ satisfies the condition in Theorem 3. We then have the following corollary.

395 **Corollary 2.** Suppose $N = Q^2 t$ for some integer t , and $\mathbf{d}^* \in \tilde{\mathcal{D}}$ has

$$n_{k,p,q}(\mathbf{d}^*) = t, \quad p, q = 1, \dots, Q, \quad k = 1, \dots, K-1.$$

396 Then, we have $\mathbf{M}(\mathbf{d}; \boldsymbol{\theta}) \leq_S \mathbf{M}(\mathbf{d}^*; \boldsymbol{\theta})$ for any other $\mathbf{d} \in \tilde{\mathcal{D}}$.

397 While being optimal for Objective II, the design $\mathbf{d}^*$ is not recommended
 398 for Objective I since it does not allow the estimation of the HRFs $\mathbf{h}$. Some
 399 optimal designs for this latter objective can be found in Kao (2013), but they

400 tend to be sub-optimal for Objective II. When these two competing objec-
 401 tives are simultaneously of interest, one may seek a multi-objective design.
 402 Finding multi-objective designs is beyond the scope of the present study,
 403 and readers are referred to Kao et al. (2008) for some discussions on such
 404 designs. Although the designs of Kao (2013) are sub-optimal for Objective
 405 II, a similar design construction method can be considered for generating
 406 optimal designs $\mathbf{d}^*$ that satisfy the condition in Corollary 2. In particular,
 407 we follow Kao (2013) to consider an m -sequence $\mathbf{v} = (v_1, \dots, v_{N-1})^T$, where
 408 $v_n \in \{0, 1, \dots, Q-1\}$, $N = Q^K$ and Q is a prime power. One way to gen-
 409 erate such an m -sequence $\mathbf{v}$ is by 1) specifying a non-zero initial K -tuple
 410 $(v_1, \dots, v_K)$, and 2) obtaining the subsequent elements of $\mathbf{v}$ with a linear
 411 recurrence relation: $v_{n+K} = a_1 v_{n+K-1} + a_2 v_{n+K-2} + \dots + a_K v_n \pmod{Q}$,
 412 $n = 1, \dots, N - K - 1$. Here, the coefficients $a_1, \dots, a_K \in \{0, 1, \dots, Q-1\}$ are
 413 determined by a primitive polynomial $f(x) = x^K - a_1 x^{K-1} - \dots - a_K x^0$
 414 over the finite field $GF(Q)$ of order Q . As a toy example, we may consider
 415 the primitive polynomial, $f(x) = x^2 - 2x - 1$, of $GF(3)$, which gives the
 416 linear recurrence relation of $v_{n+2} = 2v_{n+1} + v_n \pmod{3}$, $n = 1, \dots, 6$; here,
 417 $N = 3^2 = 9$ and $K = 2$. With the initial $(v_1, v_2) = (1, 1)$, we obtain the
 418 m -sequence $\mathbf{v}^T = (1, 1, 0, 1, 2, 2, 0, 2)$ of length $N - 1 = 8$. With such an $\mathbf{v}$,
 419 a $\mathbf{d}^*$ for Q types of stimuli can then be achieved by inserting an additional
 420 0 to any $(K - 1)$ -tuple of 0's in $\mathbf{v}$, and then replacing all the 0 in the ob-
 421 tained sequence by Q . The resulting $\mathbf{d}^* = (d_1^*, \dots, d_N^*)^T$ is a Q -ary de Bruijn
 422 sequence of order K , and every possible Q -tuple of $1, \dots, Q$ appears equally
 423 often in the set $\{(d_1^*, \dots, d_K^*), (d_2^*, \dots, d_{K+1}^*), \dots, (d_N^*, d_1^*, \dots, d_{K-1}^*)\}$. Any design
 424 possessing this property satisfies the condition in Theorem 3, and thus is
 425 optimal for studying $\boldsymbol{\theta}$. One possible drawback of such designs is that the
 426 design length can be long. In Section 4, we provide another type of very
 427 efficient m -sequence-based designs that, for the same K , can have a much
 428 shorter design length.

429 Another possible way to construct an optimal design is by using a circu-
 430 lant orthogonal array (OA) of strength 2 (see also Hedayat et al., 1999). In
 431 particular, for given N , K , and Q , a circulant OA of strength 2 is an $K \times N$
 432 array with entries $1, \dots, Q$ so that the Q^2 possible pairs appear equally often,
 433 say c times, as columns in every $2 \times N$ subarray, and the k th row of the array
 434 is obtained by cyclically shifting the $(k - 1)$ th row one position to the right.
 435 As a toy example, a circulant OA of strength 2 with $(N, K, Q) = (18, 4, 3)$
 436 and $c = 2$ can be formed by the following sequence and its cyclic shifts:
 437 $\mathbf{v} = (2, 1, 3, 3, 3, 2, 3, 1, 1, 3, 1, 2, 2, 2, 3, 2, 1, 1)^T$. By setting $\mathbf{d}^*$ to this $\mathbf{v}$, we

438 have $n_{k,p,q}(\mathbf{d}^*) = N/Q^2 = c$ for all $p, q = 1, \dots, Q$ and $k = 1, \dots, K - 1$. This
 439 design thus satisfies the condition in Corollary 2. Some circulant OAs can
 440 be found in Lin et al. (2016).

441 4. The efficiency of m -sequence-based designs

442 Since its introduction by Buracas & Boynton (2002) into the field, the
 443 m -sequences have been widely used as very efficient designs in real fMRI
 444 experiments for Objective I. For this study objective, the good performance
 445 of m -sequences has been reported in several studies (Liu & Frank, 2004;
 446 Liu, 2004; Maus et al., 2010; Kao, 2014). By focusing on Objective II with
 447 $Q = 2$, Cheng & Kao (2015) showed that an optimal design can be obtained
 448 by relabeling the two symbols of a binary m -sequence from 0 and 1 to 1 and
 449 2, respectively. With this fact and our results in the previous section, it is
 450 conjectured that, by relabeling the symbols of a Q -ary m -sequence, one can
 451 have an optimal design for Objective II with $Q \geq 3$. An analytical proof
 452 of this seems challenging, and is an ongoing research of interest. In what
 453 follows, we use an analysis similar to that in the previous section to obtain a
 454 lower bound for the A - and D -efficiencies of such m -sequence-based designs.

455 For $N = Q^r - 1$ with Q being a prime power, let $\mathbf{v} \in \{0, 1, \dots, Q - 1\}^N$ be
 456 an m -sequence of length N . With this $\mathbf{v}$, we construct an m -sequence-based
 457 design $\mathbf{d}_m$ by setting $\mathbf{d}_m = \mathbf{v} + \mathbf{j}_N$. Following the two-tuple balance property
 458 of m -sequences (see, e.g., Golomb & Gong, 2005, Section 5.6), we have, for
 459 $k = 1, \dots, N/(Q - 1) - 1$,

$$n_{k,1,1}(\mathbf{d}_m) = \frac{N+1}{Q^2} - 1; \quad n_{k,p,q}(\mathbf{d}_m) = \frac{N+1}{Q^2}, \quad (p, q) \neq (1, 1).$$

460 For this $\mathbf{d}_m$, the information matrix of $\mathbf{h}$ is $\mathbf{M}(\mathbf{d}_m; \mathbf{h}) = \mathbf{C}_1 \otimes \mathbf{I}_K - \mathbf{C}_2 \otimes \mathbf{J}_K$,
 461 where

$$\mathbf{C}_1 = \frac{N+1}{Q^2} [Q\mathbf{I}_Q - \mathbf{J}_Q]; \quad \mathbf{C}_2 = \frac{N+1}{NQ^2} \begin{bmatrix} (Q-1)^2 & (1-Q)\mathbf{j}_{Q-1}^T \\ (1-Q)\mathbf{j}_{Q-1} & \mathbf{J}_{Q-1} \end{bmatrix}.$$

462 In addition,

$$\mathbf{M}(\mathbf{d}_m; \boldsymbol{\theta}) = w^\perp(\mathbf{j}_Q) \otimes \frac{N+1}{Q} \mathbf{I}_K - \mathbf{C}_2 \otimes \mathbf{J}_K.$$

Table 1: Lower bounds of A - and D -efficiencies of $\mathbf{d}_m$ for some cases

	K	$Q = 3$		$Q = 4$		$Q = 5$	
		$N = 80$	$N = 242$	$N = 63$	$N = 255$	$N = 124$	$N = 624$
A_{eff} (%)	9	99.66	99.97	98.87	99.96	99.68	99.99
	17	99.11	99.93	93.80	99.90	99.04	99.98
D_{eff} (%)	9	99.84	99.98	99.53	99.98	99.86	99.99
	17	99.63	99.97	98.36	99.96	99.64	99.99

463 Following a similar argument as in the proof of Theorem 3, we can show that,
 464 for any $\mathbf{d} \in \tilde{\mathcal{D}}$,

$$\begin{aligned} \mathbf{M}(\mathbf{d}; \boldsymbol{\theta}) &\leq_S w^\perp(\mathbf{j}_Q) \otimes [\max_d \text{tr}\{\bar{\mathbf{\Lambda}}(\mathbf{d})\}/K] \mathbf{I}_K = w^\perp(\mathbf{j}_Q) \otimes [\text{tr}\{\bar{\mathbf{\Lambda}}(\mathbf{d}_m)\}/K] \mathbf{I}_K \\ &= w^\perp(\mathbf{j}_Q) \otimes [(N^2 - 1)/(NQ)] \mathbf{I}_K. \end{aligned}$$

465 While this derived upper bound of the information matrix is not attainable
 466 under the current setting, it can still be used to obtain a lower bound for
 467 the A - and D -efficiencies of any design $\mathbf{d}$, and thus of $\mathbf{d}_m$. Since the non-
 468 zero eigenvalues of $\mathbf{M}(\mathbf{d}_m; \boldsymbol{\theta})$ are $(N + 1)/Q$ with multiplicity $K(Q - 1) - 1$
 469 and $\{N - K(Q - 1)\}(N + 1)/(NQ)$ with multiplicity 1, we have, for $1 <$
 470 $K(Q - 1) < N$, the following lower bounds for A - and D -efficiencies of $\mathbf{d}_m$.

$$A_{eff}(\mathbf{d}_m) \geq \left(\frac{N}{N - 1} \right) \left(\frac{N - K(Q - 1)}{N - K(Q - 1) + 1} \right);$$

$$D_{eff}(\mathbf{d}_m) \geq \left(\frac{N}{N - 1} \right) \left(\frac{N - K(Q - 1)}{N} \right)^{1/\{K(Q - 1)\}}.$$

472 Table 1 presents these lower bounds for some combinations of Q , N and
 473 K . As presented there, the m -sequence-based designs are almost always
 474 very efficient under the A - and D -criteria. In addition to evaluating the
 475 performance of the m -sequence-based designs, the analysis presented above
 476 can also be used to study the efficiencies of other fMRI designs such as the
 477 ones obtained from the circulant almost-OAs introduced in Lin et al. (2016).

478 5. Conclusion

479 We derive some optimal and very efficient designs for comparing the ef-
 480 fects of different types of mental stimuli in functional neuroimaging experi-
 481 ments where the fMRI technology is used. The designs that we provide are

shown optimal or highly efficient for minimizing the average variance and generalized variance of the estimates of the contrasts of HRFs. Systematic methods for constructing such designs are also provided.

Acknowledgement

The research of Ming-Hung Kao was in part supported by National Science Foundation grants DMS-13-52213, and DMS-15-02640, and that of Frederick Kin Hing Phoa was in part supported by the Ministry of Science and Technology, Taiwan (project numbers: MOST-102-2628-M-001-002-MY3, MOST-104-2118-M-001-016-MY2, and MOST-105-2118-M-001-007-MY2), and by Academia Sinica (project number: AS-103-CDA-M04).

References

- AGUIRRE, G. K., MATTAR, M. G. & MAGIS-WEINBERG, L. (2011). De Bruijn cycles for neural decoding. *Neuroimage* **56**, 1293–1300.
- BAILEY, R. A. & DRUILHET, P. (2004). Optimality of neighbor-balanced designs for total effects. *Annals of Statistics* **32**, 1650–1661.
- BLAMIRE, A. M., OGAWA, S., UGURBIL, K., ROTHMAN, D., MCCARTHY, G., ELLERMANN, J. M., HYDER, F., RATTNER, Z., & SHULMAN, R. G. (1992). Dynamic mapping of the human visual-cortex by high-speed magnetic-resonance-imaging. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 11069–11073.
- BRENDEL, B., HERTRICH, I., ERB, M., LINDNER, A., RIECKER, A., GRODD, W. & ACKERMANN, H. (2010). The contribution of mesiofrontal cortex to the preparation and execution of repetitive syllable productions: An fMRI study. *Neuroimage* **50**, 1219–1230.
- BURACAS, G. T. & 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. *Annals of Statistics* **6**, 1239–1261.
- CHENG, C.-S. (1985). Comments on four papers by Jack Kiefer. In *Jack Carl Kiefer Collected Papers III*, L. D. Brown, I. Olkin, J. Sacks & H. P. Wynn, eds. New York: Springer, pp. 695–700.

- 513 CHENG, C.-S. (1995). Complete class results for the moment matrices of
514 designs over permutation-invariant sets. *Annals of Statistics* **23**, 41–54.
- 515 CHENG, C.-S. & KAO, M.-H. (2015). Optimal experimental designs for
516 fMRI via circulant biased weighing designs. *Annals of Statistics* **43**, 2565–
517 2587.
- 518 DELZELL, D. A. P., GUNST, R. F., SCHUCANY, W. R., CARMACK, P. S.,
519 LIN, Q., SPENCE, J. S. & HALEY, R. W. (2012). Key properties of
520 D-optimal designs for event-related functional MRI experiments with ap-
521 plication to nonlinear models. *Statistics in Medicine* **31**, 3907–3920.
- 522 ECK, J., KAAS, A. L. & GOEBEL, R. (2013). Crossmodal interactions of
523 haptic and visual texture information in early sensory cortex. *Neuroimage*
524 **75**, 123–135.
- 525 GOLOMB, S. W. & GONG, G. (2005). *Signal Design for Good Correlation*
526 *for Wireless Communication, Cryptography, and Radar*. Cambridge ; New
527 York: Cambridge University Press.
- 528 HARMAN, R. (2008). Equivalence theorem for Schur optimality of experi-
529 mental designs. *Journal of Statistical Planning and Inference* **138**, 1201–
530 1209.
- 531 HEDAYAT, A., SLOANE, N. J. A. & STUFKEN, J. (1999). *Orthogonal*
532 *Arrays: Theory and Applications*. Springer Series in Statistics. New York:
533 Springer.
- 534 HORADAM, K. J. (2007). *Hadamard Matrices and Their Applications*.
535 Princeton, N.J.: Princeton University Press.
- 536 JANSMA, J. M., DE ZWART, J. A., VAN GELDEREN, P., DUYN, J. H.,
537 DREVETS, W. C. & FUREY, M. L. (2013). In vivo evaluation of the
538 effect of stimulus distribution on FIR statistical efficiency in event-related
539 fMRI. *Journal of Neuroscience Methods* **215**, 190–195.
- 540 JIANG, X., BRADLEY, E., RINI, R. A., ZEFFIRO, T., VANMETER, J. &
541 RIESENHUBER, M. (2007). Categorization training results in shape- and
542 category-selective human neural plasticity. *Neuron* **53**, 891–903.

- 543 KAO, M.-H. (2013). On the optimality of extended maximal length linear
544 feedback shift register sequences. *Statistics & Probability Letters* **83**, 1479–
545 1483.
- 546 KAO, M.-H. (2014). A new type of experimental designs for event-related
547 fMRI via hadamard matrices. *Statistics & Probability Letters* **84**, 108–112.
- 548 KAO, M.-H. (2015). Universally optimal fMRI designs for comparing hemo-
549 dynamic response functions. *Statistica Sinica* **25**, 499–506.
- 550 KAO, M.-H., MAJUMDAR, D., MANDAL, A. & STUFKEN, J. (2013). Max-
551 imin and maximin-efficient event-related fMRI designs under a nonlinear
552 model. *Annals of Applied Statistics* **7**, 1940–1959.
- 553 KAO, M.-H., MANDAL, A., LAZAR, N. & STUFKEN, J. (2009). Multi-
554 objective optimal experimental designs for event-related fMRI studies.
555 *NeuroImage* **44**, 849–856.
- 556 KAO, M.-H., MANDAL, A. & STUFKEN, J. (2008). Optimal design for
557 event-related functional magnetic resonance imaging considering both in-
558 dividual stimulus effects and pairwise contrasts. *Statistics and Applications*
559 **6**, 225–241.
- 560 KAO, M.-H., MANDAL, A. & STUFKEN, J. (2012). Constrained multi-
561 objective designs for functional MRI experiments via a modified nondom-
562 inated sorting genetic algorithm. *Journal of the Royal Statistical Society:*
563 *Series C (Applied Statistics)* **61**, 515–534.
- 564 KAO, M.-H. & MITTELMANN, H. D. (2014). A fast algorithm for construct-
565 ing efficient event-related functional magnetic resonance imaging designs.
566 *Journal of Statistical Computation and Simulation* **84**, 2391–2407.
- 567 KAO, M.-H. & STUFKEN, J. (2015). Optimal design for event-related fMRI
568 studies. In *Handbook of Design and Analysis of Experiments*, A. Dean,
569 M. M. J. Stufken & D. Bingham, eds. Boca Raton: CRC Press, pp. 895–
570 924.
- 571 KIEFER, J. (1975). Construction and optimality of generalized youden de-
572 signs. In *A Survey of Statistical Designs and Linear Models*, J. N. Srivas-
573 tava, ed. Amsterdam: North-Holland, pp. 333–353.

- 574 KUBILIUS, J., WAGEMANS, J. & OP DE BEECK, H. P. (2011). Emer-
575 gence of perceptual gestalts in the human visual cortex: The case of the
576 configural-superiority effect. *Psychological Science* **22**, 1296–1303.
- 577 KUNERT, J. (1985). Optimal repeated measurements designs for correlated
578 observations and analysis by weighted least squares. *Biometrika* **72**, 375–
579 389.
- 580 LAZAR, N. A. (2008). *The Statistical Analysis of Functional MRI Data*.
581 Statistics for Biology and Health. New York: Springer.
- 582 LIN, Y.-L., PHOA, F. K. H. & KAO, M.-H. (2016). Optimal design of
583 fMRI experiments using circulant (almost-)orthogonal arrays. *submitted* .
- 584 LIU, T. T. (2004). Efficiency, power, and entropy in event-related fMRI
585 with multiple trial types: Part II: Design of experiments. *NeuroImage* **21**,
586 401–413.
- 587 LIU, T. T. & FRANK, L. R. (2004). Efficiency, power, and entropy in event-
588 related fMRI with multiple trial types: Part I: Theory. *NeuroImage* **21**,
589 387–400.
- 590 MAUS, B., VAN BREUKELN, G. J. P., GOEBEL, R. & BERGER, M. P. F.
591 (2010). Robustness of optimal design of fMRI experiments with application
592 of a genetic algorithm. *Neuroimage* **49**, 2433–2443.
- 593 SHAH, K. R., BOSE, M. & RAGHAVARAO, D. (2005). Universal optimality
594 of Patterson’s crossover designs. *Annals of Statistics* **33**, 2854–2872.
- 595 SUMMERFIELD, C., EGNER, T., MANGELS, J. & HIRSCH, J. (2006). Mis-
596 taking a house for a face: Neural correlates of misperception in healthy
597 humans. *Cerebral Cortex* **16**, 500–508.
- 598 WAGER, T. D. & NICHOLS, T. E. (2003). Optimization of experimental de-
599 sign in fMRI: A general framework using a genetic algorithm. *NeuroImage*
600 **18**, 293–309.
- 601 ZHANG, C. M. & YU, T. (2008). Semiparametric detection of significant
602 activation for brain fMRI. *Annals of Statistics* **36**, 1693–1725.