The Annals of Applied Statistics
2013, Vol. 7, No. 4, 1940–1959
DOI: 10.1214/13-AOAS658

MAXIMIN AND MAXIMIN-EFFICIENT EVENT-RELATED FMRI DESIGNS UNDER A NONLINEAR MODEL

BY MING-HUNG KAO, DIBYEN MAJUMDAR, ABHYUDAY MANDAL[1] AND JOHN STUFKEN[2]

Arizona State University, University of Illinois at Chicago and University of Georgia

Previous studies on event-related functional magnetic resonance imaging experimental designs are primarily based on linear models, in which a known shape of the hemodynamic response function (HRF) is assumed. However, the HRF shape is usually uncertain at the design stage. To address this issue, we consider a nonlinear model to accommodate a wide spectrum of feasible HRF shapes, and propose efficient approaches for obtaining maximin and maximin-efficient designs. Our approaches involve a reduction in the parameter space and a search algorithm that helps to efficiently search over a restricted class of designs for good designs. The obtained designs are compared with traditional designs widely used in practice. We also demonstrate the usefulness of our approaches via a motivating example.

1. Introduction. Functional magnetic resonance imaging (fMRI) is a pioneering, noninvasive brain mapping technology for studying brain functions [Culham (2006), D'Esposito, Zarahn and Aguirre (1999)]. It is arguably one of the most important advances in neuroscience and has many important clinical potentials such as early identification of Alzheimer's disease, pre-neurosurgical planning, and post-neurosurgical evaluations; see, Bookheimer (2007) and Wierenga and Bondi (2007). This cutting-edge technology has been applied in a wide variety of disciplines [Lazar (2008), Lindquist (2008)].

In a typical fMRI experiment, a predetermined sequence of mental stimuli (e.g., pictures or sounds) is presented to a subject. While the subject is exposed to the stimuli, an MR scanner repeatedly scans the subject's brain to collect a blood oxygenated level dependent (BOLD) time series from each brain voxel (three-dimensional imaging unit). A study usually involves multiple (e.g., $64 \times 64 \times 30$) voxels, resulting in multiple time series. These time series reflect the MR signal changes evoked by the underlying brain activity and are analyzed to make statistical inference about the inner workings of the brain. A crucial first step for rendering

Received February 2013; revised May 2013.
[1]Supported in part by NSF Grant DMS-09-05731 and NSA Grant H98230-13-1-0251.
[2]Supported in part by NSF Grants DMS-07-06917 and DMS-10-07507.
Key words and phrases. A-optimality, cyclic permutation, design efficiency, genetic algorithms, hemodynamic response function, information matrix.

a valid and precise inference is to select a high quality experimental design for the fMRI experiment.

Here, we focus on event-related (ER) fMRI designs with brief mental stimuli. Such designs are very popular due to their flexibility [Huettel (2012), Josephs and Henson (1999)], and are the primary focus of existing research on fMRI designs [e.g., Kao et al. (2009), Liu (2004), Maus et al. (2010b), Wager and Nichols (2003)]. Current knowledge about the performance of ER-fMRI designs is mainly based on general linear models. While popular, the use of general linear models is criticized by some researchers [Handwerker, Ollinger and D'Esposito (2004), Loh, Lindquist and Wager (2008), Worsley and Taylor (2006)]. A major criticism is the assumption of a fixed, known shape of the hemodynamic response function (HRF), a function of time describing the noise-free MR signal change evoked by one, single stimulus. This assumption is not always valid. Studies showed that the HRF shape may vary across brain voxels, and that a misspecified shape can lead to incorrect conclusions. To allow for uncertain HRF shapes, analysis methods such as the use of nonlinear models have been seen in the literature [e.g., Handwerker, Ollinger and D'Esposito (2004), Lindquist and Wager (2007), Miezin et al. (2000)]. However, not much work has been done to address this important issue at the design stage.

Kao (2009) investigated the performance of ER-fMRI designs under a nonlinear model (Section 2.1) that can accommodate a wide variety of feasible HRF shapes. With such a model, the optimality criterion for evaluating the performance of designs typically depends on unknown model parameters. Kao (2009) assumed the availability of a prior distribution of unknown parameters and put forward an approach for obtaining designs optimizing a (pseudo-)Bayesian design criterion, which is the expected value of the specific optimality criterion. Maus et al. (2012) considered a maximin-type approach that focuses on the worst case scenario over a prespecified parameter space containing possible values of the model parameters. Specifically, they targeted designs that maximize the worst relative efficiency over the parameter space. Here, a relative efficiency is the relative value of the specific optimality criterion with respect to a locally optimal design that is optimal for a given parameter vector value. Following Müller (1995), the obtained designs will be termed as maximin-efficient designs.

In contrast to maximin-efficient designs, maximin designs optimize the worst value of the optimality criterion. In other words, the maximin criterion focuses directly on the worst performance of designs over the parameter space, and the maximin-efficient criterion can be viewed as a "weighted" version of maximin criterion. The weights are determined by locally optimal designs or, more precisely, the best possible value of the optimality criterion evaluated at each parameter vector value. Both criteria are considered in a wide variety of design problems [e.g., Berger, King and Wong (2000), Berger and Wong (2009), Chen, Wong and Li (2008), Huang and Lin (2006), King and Wong (2000), Silvey (1980), Sitter (1992)]. Unfortunately, obtaining maximin-type designs optimizing these criteria

is very challenging. One typically needs to deal with an optimization problem that is mathematically intractable and computationally difficult, if not infeasible [Chen et al. (2011), Dette, Haines and Imhof (2007)]. An efficient approach is thus crucially important.

In this paper, we propose approaches for obtaining maximin and maximin-efficient designs for fMRI experiments to allow uncertain HRF shapes. We derive useful results and develop efficient strategies for obtaining high-quality designs. Our strategies involve a reduced parameter space, a restricted class of ER-fMRI designs, and an efficient search algorithm for searching over the restricted design class for good designs. The usefulness of our approaches is demonstrated via case studies and a real example.

We note that Maus et al. (2012) obtained D-optimal maximin-efficient designs for one stimulus type. Here, we develop approaches for obtaining both maximin and maximin-efficient designs, and apply our methods to find maximin-type designs under A-optimality for cases with one or more stimulus types. D-optimal designs help to control the volume of a confidence ellipsoid of the parameters. By contrast, A-optimality aims at maximizing the average estimation precision. While D-optimality is not uncommon in fMRI, the A-optimality criterion is widely accepted by researchers in the field [see also Dale (1999), Friston et al. (1999), Kao et al. (2009), Maus et al. (2010b)]. We also note that our proposed methods can be applied to all optimality criteria that are invariant under simultaneous permutation of rows and columns of the information matrix. Both A- and D-optimality criteria possess this invariance property.

The remainder of the article is organized as follows. In Section 2 we provide a brief introduction about ER-fMRI designs. We then introduce our methods, including the underlying statistical model, optimality criteria, and our proposed strategies for obtaining maximin and maximin-efficient designs. Case studies and a real example are provided in Section 3. The paper closes with a conclusion in Section 4.

2. Background and methodology.

2.1. *A nonlinear model.* An ER-fMRI design is a finite sequence of brief stimuli interlaced with control to be presented to an experimental subject. Each stimulus may last several milliseconds to a few seconds. Times between consecutive stimulus onsets are multiples of a prespecified time, called the inter stimulus interval (ISI; e.g., 4 s). The control (e.g., periods of fixation or rest) fills in the time when no stimulus is being presented. We may use a sequence of finite numbers, for example, $d = \{101210 \cdots 1\}$, to represent an ER-fMRI design. An integer q $(\neq 0)$ at the kth position indicates an onset of a qth-type stimulus at time $(k-1)$ISI. A "0" means no stimulus onset at that time point.

At an activated brain voxel, each stimulus evokes a change in the MR signal intensity. The signal intensity takes about 25 to 30 seconds to rise and decay. This

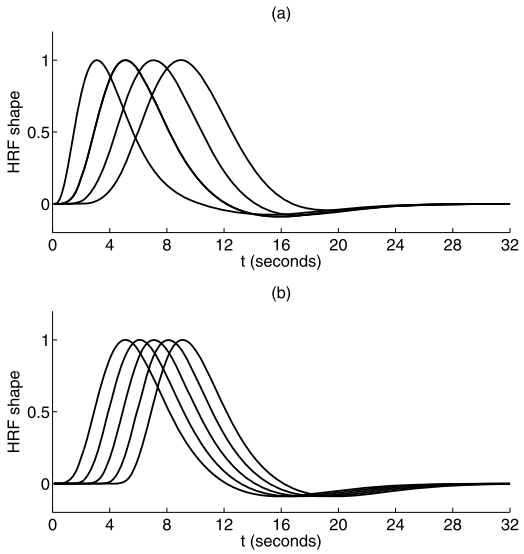

FIG. 1. *The HRF shapes* $g(t; \mathbf{p} = (p_1, p_6))$ *of* (2.2) *with, from left to right*, (a) $p_6 = 0$ *and* $p_1 = 4$ *to* 10 *in steps of* 2; *and* (b) $p_1 = 6$ *and* $p_6 = 0$ *to* 4 *in steps of* 1.

change is typically described by an HRF having an assumed shape with an unknown amplitude (maximal height); see Figure 1 for some possible HRF shapes. When the next stimulus occurs before the cessation of the current HRF, the evoked HRFs accumulate. Along with nuisance signals and noise, the accumulated HRF is acquired by an MR scanner every TR (time-to-repetition; e.g., 2 s) to form the BOLD time series. Denoting the time series of a voxel by a T-by-1 vector $\mathbf{y}$, we consider the following nonlinear model:

$$\mathbf{y} = \sum_{q=1}^{Q} \mathbf{X}_{d,q}\mathbf{h}(\mathbf{p})\theta_q + \mathbf{S}\boldsymbol{\gamma} + \mathbf{e}. \tag{2.1}$$

Here, Q is the number of stimulus types. $\mathbf{X}_{d,q}\mathbf{h}(\mathbf{p})\theta_q$ represents the accumulated HRF evoked by the qth-type stimuli of a design d. The scalar θ_q is the unknown HRF amplitude. The vector $\mathbf{h}(\mathbf{p})$, indexed by an unknown parameter vector $\mathbf{p}$, depicts the heights of the HRF shape after every ΔT seconds following a stimulus onset; ΔT is the greatest value making both $(\mathrm{ISI}/\Delta T)$ and $(\mathrm{TR}/\Delta T)$ integers. $\mathbf{X}_{d,q}$ is the 0–1 design matrix with 1 indicating the heights of the HRF that contribute to each BOLD measurement; a construction of $\mathbf{X}_{d,q}$ can be found in the Appendix of Kao, Mandal and Stufken (2012). The nuisance term $\mathbf{S}\boldsymbol{\gamma}$ allows a drift/trend over time with an unknown parameter vector $\boldsymbol{\gamma}$. The correlated noise is represented by $\mathbf{e}$. For detecting brain voxels activated by the stimuli, the focus is typically on the amplitudes, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_Q)$, which reflect the "strengths" of

brain activation. A large θ_q-value signals a voxel that is highly activated by the qth-type stimuli, $q = 1, \ldots, Q$.

With unknown $\mathbf{p}$, model (2.1) allows for an uncertain HRF shape $\mathbf{h}(\mathbf{p})$. The vector $\mathbf{h}(\mathbf{p})$ is determined by a continuous function $g(t; \mathbf{p})$ with t representing time elapsed after a stimulus onset. There are many choices for $g(t; \mathbf{p})$. Our selected $g(t; \mathbf{p})$ has the same form as the double-gamma function of SPM (http://www.fil.ion.ucl.ac.uk/spm/), a popular computer software package for analyzing fMRI data:

$$g(t; \mathbf{p}) = \frac{g_0(t; \mathbf{p})}{\max_s g_0(s; \mathbf{p})}, \tag{2.2}$$

where

$$g_0(t; \mathbf{p}) = f\left(t - p_6, \frac{p_1}{p_3}, p_3\right) - p_5 f\left(t - p_6, \frac{p_2}{p_4}, p_4\right);$$

$$f(x, \alpha, \beta) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}};$$

$\Gamma(\cdot)$ is the gamma function; and $f(x, \alpha, \beta)$ is the probability density function of the gamma distribution, gamma(α, β). The double-gamma function of SPM fixes $(p_1, p_2, \ldots, p_6) = (6, 16, 1, 1, 1/6, 0)$. This function is completely known and is commonly used in the general linear model approach for describing the HRF shape. By contrast, we allow an uncertain HRF shape and follow Wager et al. (2005) to treat the two most influential HRF parameters, namely, p_1, time-to-peak, and p_6, time-to-onset, as free parameters, while keeping the less sensitive parameters (p_2, p_3, p_4, p_5) fixed at $(16, 1, 1, 1/6)$. The HRF shapes with selected (p_1, p_6)-values can be found in Figure 1. For brevity, we will omit the fixed parameters p_2, p_3, p_4, and p_5 from $\mathbf{p}$ and write $\mathbf{p} = (p_1, p_6)$, although $\mathbf{p}$ should really include six parameters. The jth element of the vector $\mathbf{h}(\mathbf{p})$ is then $g((j-1) \times (\Delta T); \mathbf{p})$. The length of $\mathbf{h}(\mathbf{p})$ is set to $1 + \lfloor 32/\Delta T \rfloor$ since a typical HRF is nearly zero after 32 seconds; here, $\lfloor a \rfloor$ is the integer part of a.

2.2. *Optimality criteria.* We aim at a good design for detecting activation (or studying $\boldsymbol{\theta}$) with model (2.1). The performance of a design will be evaluated by $1/\operatorname{trace}(\operatorname{Cov}[\hat{\boldsymbol{\theta}}])$, the reciprocal of the average variance of the generalized least squares estimators $\hat{\boldsymbol{\theta}}$, that is, A-optimality. Following a popular technique [Box and Lucas (1959), Fedorov and Hackl (1997)], we first linearize model (2.1) and then use the linearized model to approximate $\operatorname{Cov}[\hat{\boldsymbol{\theta}}]$. The approximated covariance matrix is proportional to $\mathbf{M}^{-1}(d; \boldsymbol{\theta}, \mathbf{p})$, where

$$\begin{aligned}
\mathbf{M}(d; \boldsymbol{\theta}, \mathbf{p}) &= \mathbf{E}_d(\mathbf{p})'\big[\mathbf{I}_T - w\{\mathbf{L}_d(\boldsymbol{\theta}, \mathbf{p})\}\big]\mathbf{E}_d(\mathbf{p}),\\
\mathbf{E}_d(\mathbf{p}) &= \big[\mathbf{I}_T - w\{\mathbf{VS}\}\big]\mathbf{V}\mathbf{X}_d\big[\mathbf{I}_Q \otimes \mathbf{h}(\mathbf{p})\big],\\
\mathbf{L}_d(\boldsymbol{\theta}, \mathbf{p}) &= [\mathbf{L}_1, \mathbf{L}_6],
\end{aligned}$$

$$\mathbf{L}_i = [\mathbf{I}_T - w\{\mathbf{VS}\}]\mathbf{V}\mathbf{X}_d\left[\mathbf{I}_Q \otimes \frac{\partial \mathbf{h}(\mathbf{p})}{\partial p_i}\right]\boldsymbol{\theta}', \qquad i = 1, 6,$$

$\mathbf{I}_a$ is the a-by-a identity matrix, $w\{\mathbf{A}\} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'$ is the orthogonal projection matrix onto the column space of $\mathbf{A}$, $\mathbf{A}^{-}$ is a generalized inverse matrix of $\mathbf{A}$, $\mathbf{X}_d = [\mathbf{X}_{d,1}, \ldots, \mathbf{X}_{d,Q}]$, $\mathbf{V}$ is selected so that $\mathbf{Ve}$ is white noise, $\otimes$ is the Kronecker product, and the vector $(\partial \mathbf{h}(\mathbf{p})/\partial p_i)$ is determined by the partial derivative of $g(t; \mathbf{p})$ with respect to p_i, $i = 1, 6$.

We would like a design maximizing $\Phi_A(d; \boldsymbol{\theta}, \mathbf{p}) \equiv 1/\operatorname{trace}(\mathbf{M}^{-1}(d; \boldsymbol{\theta}, \mathbf{p}))$. The answer will depend on the unknown $\boldsymbol{\theta}$ and $\mathbf{p}$. This makes such a nonlinear design problem notoriously difficult. One way for tackling such a problem is by obtaining a locally optimal design [Chernoff (1953)] that is optimal for a given $(\boldsymbol{\theta}, \mathbf{p})$-value. However, this approach is unsatisfactory for fMRI. This is because a good guess for the parameter vector value is almost always unavailable. More importantly, the selected design should be efficient for the various parameter values (or HRF shapes) associated with all the brain voxels of interest. We thus resort to the maximin and maximin-efficient approaches.

The maximin approach seeks designs maximizing

$$\min_{(\boldsymbol{\theta}, \mathbf{p}) \in \Theta \times \mathcal{P}} \Phi_A(d; \boldsymbol{\theta}, \mathbf{p}), \tag{2.3}$$

where $\Theta \times \mathcal{P}$ is a specified parameter space of $(\boldsymbol{\theta}, \mathbf{p})$. A maximin design thus maximizes the worst average precision in estimating $\boldsymbol{\theta}$ by taking the uncertainty of both $\boldsymbol{\theta}$ and $\mathbf{p}$ into account. On the other hand, the maximin-efficient criterion is

$$\min_{(\boldsymbol{\theta}, \mathbf{p}) \in \Theta \times \mathcal{P}} \mathrm{RE}(d; \boldsymbol{\theta}, \mathbf{p}) = \min_{(\boldsymbol{\theta}, \mathbf{p}) \in \Theta \times \mathcal{P}} \frac{\Phi_A(d; \boldsymbol{\theta}, \mathbf{p})}{\Phi_A(d^*_{\boldsymbol{\theta}, \mathbf{p}}; \boldsymbol{\theta}, \mathbf{p})}, \tag{2.4}$$

where $d^*_{\boldsymbol{\theta}, \mathbf{p}}$ is a locally optimal design maximizing Φ_A for given $(\boldsymbol{\theta}, \mathbf{p})$. To reflect that the HRF typically increases in 0–2 s after the stimulus onset, reaches the peak in 5–8 s, and then falls back to baseline [Lindquist (2008), Rosen, Buckner and Dale (1998)], we set $\mathcal{P} = \{(p_1, p_6) \mid p_1 \in [6, 9], p_6 \in [0, 2]\}$. This choice also follows the fact that the mode of the gamma distribution gamma$(\alpha, 1)$ is $(\alpha - 1)$ for $\alpha > 1$. Other $\mathcal{P}$ can also be considered. With no further information, we consider $\mathbb{R}^Q$ as the parameter space of $\boldsymbol{\theta}$, which can be greatly reduced using the results presented in the next subsection.

2.3. *Strategies to find maximin and maximin-efficient designs.* Obtaining maximin or maximin-efficient ER-fMRI designs is computationally challenging. Results in this section help to reduce the computational burden. We first discuss results useful for the maximin approach. Some of these results can also be applied to the maximin-efficient approach. Additional results that facilitate the maximin-efficient approach are then described.

LEMMA 1. $\mathbf{M}^{-1}(d;\mathbf{0},\mathbf{p}) \le \mathbf{M}^{-1}(d;\boldsymbol{\theta},\mathbf{p})$ *in Löwner ordering for any* $\boldsymbol{\theta}$, $\mathbf{p}$, *and a design* d *that ensures the existence of* $\mathbf{M}^{-1}(d;\boldsymbol{\theta},\mathbf{p})$.

LEMMA 2. $\mathbf{M}(d;c\boldsymbol{\theta},\mathbf{p}) = \mathbf{M}(d;\boldsymbol{\theta},\mathbf{p})$ *for any scalar* $c \neq 0$.

The first lemma follows from Theorem 18.3.4 of Harville (1997), and allows us to leave out $\mathbf{0}$ from the parameter space of $\boldsymbol{\theta}$ when obtaining maximin designs. We note that the existence of $\mathbf{M}^{-1}(d;\mathbf{0},\mathbf{p})$ is guaranteed by the nonsingularity of $\mathbf{M}(d;\boldsymbol{\theta},\mathbf{p})$. Lemma 2 is linked to an observation made by Bose and Stufken (2007). It suggests that the Φ_A-value depends on the direction of $\boldsymbol{\theta}$, but not on its length. Thus, when $Q = 1$, $\Phi_A(d;\theta_1,\mathbf{p}) = \Phi_A(d;1,\mathbf{p})$ for any $\mathbf{p}$ and $\theta_1 \neq 0$. The parameter space can then be reduced to $\{1\} \times \mathcal{P}$ from $\mathbb{R} \times \mathcal{P}$. For $Q > 1$, we represent $\boldsymbol{\theta}$ using the hyper-spherical coordinate system, and focus only on the surface of the Q-dimensional unit hemisphere centered at the origin. Specifically, for $Q = 2$, the parameter space of $\boldsymbol{\theta}$ can be reduced to $\Theta = \{(\cos\varphi_1, \sin\varphi_1) \mid \varphi_1 \in (-\pi/2, \pi/2]\}$. For $Q = 3$, $\Theta = \{(\cos\varphi_1, \sin\varphi_1\cos\varphi_2, \sin\varphi_1\sin\varphi_2) \mid \varphi_i \in (-\pi/2, \pi/2]\}$ can be used. For a larger Q, we have $\Theta = \{(\theta_1, \ldots, \theta_Q)\}$, where

$$\theta_1 = \cos\varphi_1; \qquad \theta_q = \cos\varphi_q \prod_{i=1}^{q-1} \sin\varphi_i, \qquad q = 2, \ldots, Q-1;$$

$$\theta_Q = \prod_{i=1}^{Q-1} \sin\varphi_i; \varphi_1, \ldots, \varphi_{Q-1} \in (-\pi/2, \pi/2].$$

The two lemmas allow for a large reduction in the parameter space and facilitate the search for maximin designs. To further decrease the computational cost, we propose an efficient strategy using the following result.

LEMMA 3. *Let* $\mathcal{G} = \{\mathbf{G}_1, \ldots, \mathbf{G}_G\}$ *be a set of* $Q \times Q$ *permutation matrices. Suppose* $\Theta_0 \subset \Theta$ *is such that* $\Theta = \bigcup_{g=0}^{G} \Theta_g$, *where* $\Theta_g = \{\mathbf{G}_g\boldsymbol{\theta} \mid \boldsymbol{\theta} \in \Theta_0\}$ *and* $\mathbf{G}_0 \equiv \mathbf{I}_Q$. *If* d_{Mm,Θ_0} *is a maximin design for* $\Theta_0 \times \mathcal{P}$ *and* $\min_{\Theta_0\times\mathcal{P}} \Phi_A(d_{\mathrm{Mm},\Theta_0}; \boldsymbol{\theta}, \mathbf{p}) = \min_{\Theta_g\times\mathcal{P}} \Phi_A(d_{\mathrm{Mm},\Theta_0}; \boldsymbol{\theta}, \mathbf{p})$ *for any* g, *then* d_{Mm,Θ_0} *is also a maximin design for* $\Theta \times \mathcal{P}$.

A proof of Lemma 3 can be found in the Appendix. It is noteworthy that, although we present Lemma 3 using Φ_A, this lemma can be applied to any optimality criterion that is invariant under simultaneous permutation of rows and columns of the information matrix $\mathbf{M}(d;\boldsymbol{\theta},\mathbf{p})$. Many commonly used optimality criteria, including A- and D-optimality, satisfy this invariance property; see also Cheng (1996). This lemma motivates the following strategy for obtaining maximin designs:

STRATEGY 1. (1) Identify a Θ_0 and $\mathcal{G}$; and (2) obtain a design d_{Mm,Θ_0} maximizing $\min_{\Theta_0\times\mathcal{P}} \Phi_A(d;\boldsymbol{\theta},\mathbf{p})$, for which the ratio

$$\mathcal{R}_g = \frac{\min_{\Theta_g\times\mathcal{P}} \Phi_A(d_{\mathrm{Mm},\Theta_0};\boldsymbol{\theta},\mathbf{p})}{\min_{\Theta_0\times\mathcal{P}} \Phi_A(d_{\mathrm{Mm},\Theta_0};\boldsymbol{\theta},\mathbf{p})} \tag{2.5}$$

is 1 for any $g = 1,\ldots,G$.

If such a d_{Mm,Θ_0} exists, then it is a maximin design for the entire parameter space. On the other hand, if $\mathcal{R}_g < 1$ for some g, calculating the minimal $\mathcal{R}_g$ still provides a lower bound for the efficiency of d_{Mm,Θ_0}. More precisely,

$$\begin{aligned}\min_{g\neq 0}\mathcal{R}_g &\le \min_{g\neq 0}\frac{\min_{\Theta_g\times\mathcal{P}} \Phi_A(d_{\mathrm{Mm},\Theta_0};\boldsymbol{\theta},\mathbf{p})}{\min_{\Theta_0\times\mathcal{P}} \Phi_A(d_{\mathrm{Mm}};\boldsymbol{\theta},\mathbf{p})}\\ &\le \min_{g\neq 0}\frac{\min_{\Theta_g\times\mathcal{P}} \Phi_A(d_{\mathrm{Mm},\Theta_0};\boldsymbol{\theta},\mathbf{p})}{\min_{\Theta\times\mathcal{P}} \Phi_A(d_{\mathrm{Mm}};\boldsymbol{\theta},\mathbf{p})} \qquad (2.6)\\ &= \frac{\min_{\Theta\times\mathcal{P}} \Phi_A(d_{\mathrm{Mm},\Theta_0};\boldsymbol{\theta},\mathbf{p})}{\min_{\Theta\times\mathcal{P}} \Phi_A(d_{\mathrm{Mm}};\boldsymbol{\theta},\mathbf{p})},\end{aligned}$$

where d_{Mm} is a maximin design for $\Theta\times\mathcal{P}$. Note that the equality in (2.6) follows from the fact that $\min_{\Theta_0\times\mathcal{P}} \Phi_A(d_{\mathrm{Mm},\Theta_0};\boldsymbol{\theta},\mathbf{p}) \ge \min_{\Theta_g\times\mathcal{P}} \Phi_A(d_{\mathrm{Mm},\Theta_0};\boldsymbol{\theta},\mathbf{p})$ for any g, which can be proved by using Lemmas 5 and 6 in the Appendix. If the minimal $\mathcal{R}_g$ is close to 1, d_{Mm,Θ_0} will perform well in terms of the maximin criterion (2.3).

We now turn to results that help to obtain maximin-efficient designs. To compute the RE-value in (2.4), we need locally optimal designs for all $(\boldsymbol{\theta},\mathbf{p})$ in the parameter space. Obtaining these locally optimal designs is computationally demanding (or infeasible), especially when the parameter space is large. The following results partly relieve this computational burden.

COROLLARY 1. *A locally optimal design $d^*_{\boldsymbol{\theta},\mathbf{p}}$ for $(\boldsymbol{\theta},\mathbf{p})$ is also a locally optimum design for $(c\boldsymbol{\theta},\mathbf{p})$ for any $c\neq 0$.*

Corollary 1 follows from Lemma 2. We also have the following corollary that allows for a reduction in the parameter space when obtaining maximin-efficient designs.

COROLLARY 2. $\mathrm{RE}(d;\boldsymbol{\theta},\mathbf{p}) = \mathrm{RE}(d;c\boldsymbol{\theta},\mathbf{p})$ *for any design d and $c\neq 0$.*

With this corollary, we may now reduce the parameter space to $\{\{\mathbf{0}\}\cup\Theta\}\times\mathcal{P}$ with Θ being the surface of the Q-dimensional unit hemisphere centered at the origin. Similarly to Lemma 3, we make the following observation to help further reduce the parameter space.

LEMMA 4. *A maximin-efficient design* $d_{\mathrm{MmE},\Theta_0}$ *for* $\{\{\mathbf{0}\}\cup\Theta_0\}\times\mathcal{P}$ *is also a maximin-efficient design for* $\Theta\times\mathcal{P}$ *if, for any* g,

$$\min_{\{\{\mathbf{0}\}\cup\Theta_0\}\times\mathcal{P}} \mathrm{RE}(d_{\mathrm{MmE},\Theta_0};\boldsymbol{\theta},\mathbf{p}) = \min_{\{\{\mathbf{0}\}\cup\Theta_g\}\times\mathcal{P}} \mathrm{RE}(d_{\mathrm{MmE},\Theta_0};\boldsymbol{\theta},\mathbf{p}).$$

With Lemma 4, a strategy similar to Strategy 1 can be considered for obtaining maximin-efficient designs. Specifically, we may find a maximin-efficient design for the reduced parameter space $\{\{\mathbf{0}\}\cup\Theta_0\}\times\mathcal{P}$, and check if the minimal RE-value of the obtained design is similar across all $\{\{\mathbf{0}\}\cup\Theta_g\}\times\mathcal{P}$, $g=0,1,\ldots,G$. Unfortunately, this strategy, which works well for the maximin approach, may fail to provide good maximin-efficient designs; see Section 3. A closer look reveals that the RE-values and min-RE can be greatly changed after a permutation of the coordinates of $\boldsymbol{\theta}$. This motivates us to consider another strategy suggested by Lemmas 3 and 4. The idea is to search for maximin-type designs over a restricted class of designs for which the Φ_A-values are (nearly) invariant to permutations of the elements of $\boldsymbol{\theta}$. The restricted design class Ξ_0 that we consider is described below.

With $Q(>1)$ stimulus types and a design length L, each design in the restricted design class Ξ_0 is formed by a "short design" of length $\lceil L/Q\rceil$; $\lceil a\rceil$ is the smallest integer $\geq a$. The labels of stimulus types of the initial short design are cyclically permuted to generate additional $Q-1$ short designs. In particular, the label q in the current short design is replaced by $q+1$ in the next short design, $q=1,\ldots,Q-1$; the label Q becomes 1, and 0's are kept intact. A design of length L is then achieved by concatenating the Q short designs and leaving out the last $(Q\lceil L/Q\rceil - L)$ elements. The design class Ξ_0 has also been considered by Kao, Mandal and Stufken (2009). Here, we are able to show in the supplementary document [Kao et al. (2013)] that, with a simplified model and two stimulus types ($Q=2$), the Φ_A-values of designs in Ξ_0 are quite insensitive to permutations of the elements of $\boldsymbol{\theta}$. Based on our empirical results, this observation tends to remain true for more realistic situations. We now describe our proposed second strategy.

STRATEGY 2. (1) Identify a Θ_0 and $\mathcal{G}$; and (2) obtain a design $d_{\mathrm{MmE},\Theta_0}$ that maximizes $\min_{\{\{\mathbf{0}\}\cup\Theta_0\}\times\mathcal{P}} \mathrm{RE}(d;\boldsymbol{\theta},\mathbf{p})$ in the subclass Ξ_0.

The results of our case studies indicate that good maximin-efficient designs can be found by Strategy 2 with a greatly reduced computing time. We note that, by considering the maximin criterion of (2.3) in Strategy 2, maximin designs can be obtained over the subclass Ξ_0. In the next subsection, we apply these strategies for illustration.

3. Case studies and a real example.

3.1. *Maximin designs.* We consider three cases with $(Q,L)=(1,255)$, $(2,242)$, and $(3,255)$. For each case, ISI is set to 4 s and TR is 2 s. Following

Kao et al. (2009) and Liu (2004), a second-order polynomial drift in the BOLD times series and an AR(1) noise with an autocorrelation coefficient $\rho = 0.3$ are assumed. With these settings, we adapt the knowledge-based genetic algorithm (GA) of Kao et al. (2009) to search for maximin designs; see the supplementary document [Kao et al. (2013)] for the details of this GA. During the search, the minima of Φ_A of candidate designs are evaluated over a grid on the specific parameter space. The grid interval is 0.2 for $\mathbf{p}$, and 0.1π for φ_i's (and thus $\boldsymbol{\theta}$). When comparing the obtained designs, finer grid intervals of 0.1 and 0.05π are considered for $\mathbf{p}$ and $\boldsymbol{\theta}$, respectively. We implement our MATLAB programs on a desktop computer with a 3.4 GHz Core i7-2600 processor. These programs are available from the authors.

For $Q = 1$, our GA search first targets a design maximizing $\min_{\mathcal{P}} \Phi_A(d; 1, \mathbf{p})$. Although we focus only on $\{1\} \times \mathcal{P}$, Lemmas 1 and 2 warrant that the obtained maximin designs are for the entire parameter space $\mathbb{R}^1 \times \mathcal{P}$. For $Q = 2$ and 3, we follow the proposed strategies to first identify a subset Θ_0 of Θ and a class of permutation matrices $\mathcal{G}$. We recommend to include all the Q-by-Q permutation matrices, except for the identity matrix, in $\mathcal{G}$ to allow a small Θ_0. We take Θ_0 as $\{(\cos\varphi_1, \sin\varphi_1) \mid \varphi_1 \in [-\pi/4, \pi/4]\}$ for $Q = 2$, and as $\{(\cos\varphi_1, \pm\sin\varphi_1\cos\varphi_2, \pm\sin\varphi_1\sin\varphi_2) \mid \varphi_1 \in [0, \arccos(1/\sqrt{3})],\ \varphi_2 \in [\kappa, \pi/4]\}$ for $Q = 3$; $\kappa = \arccos(\cos\varphi_1/\sin\varphi_1)$ if $\varphi_1 > \pi/4$, and $\kappa = 0$, otherwise. We note that, for these two cases, the Θ defined after Lemma 2 can be written as $\Theta = \bigcup_{g=0}^{Q!-1} \Theta_g^*$, where $\Theta_g^* = \{\tau_{g,\theta}\mathbf{G}_g\boldsymbol{\theta} \mid \boldsymbol{\theta} \in \Theta_0\}$, and $\tau_{g,\theta}$ is the sign of $((\mathbf{G}_g\boldsymbol{\theta}))_1$, the first element of $\mathbf{G}_g\boldsymbol{\theta}$; we set $\tau_{g,\theta}$ to 1 when $((\mathbf{G}_g\boldsymbol{\theta}))_1 = 0$. It is easy to see that, by using Lemma 2, Lemma 3 still holds after replacing Θ_g with Θ_g^*.

With the selected Θ_0 for $Q > 1$, the GA is applied to search for d_{Mm,Θ_0} maximizing $\min_{\Theta_0\times\mathcal{P}} \Phi_A(d; \boldsymbol{\theta}, \mathbf{p})$. Both Strategies 1 and 2 are considered to reduce computational burden. Specifically, following Strategy 1, we apply the GA to find d_{Mm,Θ_0} over the space Ξ of all designs, and obtain the minimal $\mathcal{R}_g$ in (2.5) as a lower bound of the efficiency of the obtained design. We also use the GA to find such a design over the restricted design class Ξ_0 (i.e., Strategy 2). For each case, we generate ten designs by using different random seeds in the GA.

Table 1 presents the maximum, mean, and standard error of min-Φ_A of the ten GA-generated designs. The mean CPU time for obtaining these designs is also reported. As in Table 1, our two strategies yield designs with similar min-Φ_A values. In addition, the minimal $\mathcal{R}_g$ for the designs obtained with Strategy 1 is at least 98.78% for $Q = 2$ and at least 97.99% for $Q = 3$, indicating that our obtained designs are very efficient compared with a maximin design d_{Mm} for $\Theta \times \mathcal{P}$. We also note that a direct search for d_{Mm} can be very time consuming for $Q > 1$. By focusing on the reduced parameter space $\Theta_0 \times \mathcal{P}$, our proposed methods can efficiently generate high quality designs. In addition, obtaining maximin designs over the subclass Ξ_0 of designs can further reduce the computational burden without having a negative effect on the design efficiencies. These results provide compelling evidence for the efficiency and effectiveness of Strategy 2.

TABLE 1
The performance and mean CPU time (in minutes) for d_{Mm} *obtained over* Ξ *(all designs) for* $Q = 1$; *and* d_{Mm,Θ_0} *obtained over* Ξ *and the subclass* Ξ_0 *for* $Q = 2$ *and* 3

	$Q=1$	$Q=2$		$Q=3$	
	Ξ	Ξ	Ξ_0	Ξ	Ξ_0
min-Φ_A					
Maximum	75.67	25.67	25.64	14.19	14.41
Mean	75.34	25.48	25.54	14.06	14.29
Std. err.	0.08	0.05	0.04	0.03	0.03
Mean CPU time	0.85	9.29	4.16	165.96	56.54

We also compare the obtained designs with some traditional designs that are widely used in practice. Figure 2 presents the boxplots of the average estimation precision (Φ_A) over $\Theta \times \mathcal{P}$ for the competing designs. In that figure, the selected maximin design for $Q = 1$ is the design maximizing min-Φ_A over the ten d_{Mm} designs obtained by the GA; for $Q > 1$, the selected designs maximize min-Φ_A over the ten d_{Mm,Θ_0} designs obtained via Strategy 2. The traditional designs include block designs, m-sequences, max-F_d, max-F_e, bi-objective, and random designs. Block designs for fMRI are sequences formed by repetitions of $\{B_0 B_1 B_2 \cdots B_Q\}$, where B_q is a sequence of q's (i.e., $\{qq \cdots q\}$) of a given size. Here, we consider block designs of size four that are formed by $\{00001111 \cdots QQQQ\}$. Under lin-

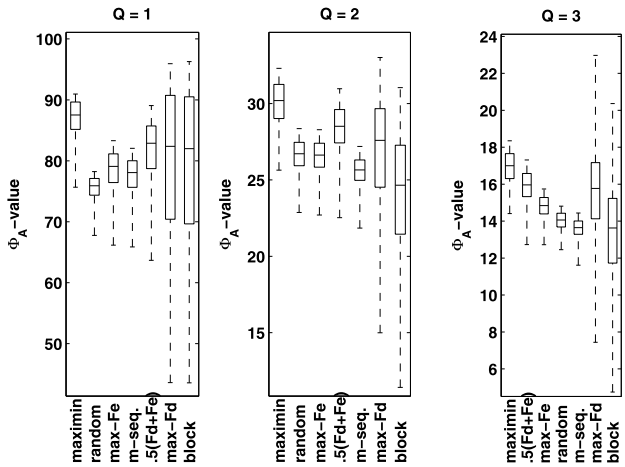

FIG. 2. *Boxplots of the* Φ_A*-values over* $\Theta \times \mathcal{P}$ *of the competing designs, including maximin designs, designs obtained with linear models for estimation* (max-F_e), *detection* (max-F_d) *and both* ($0.5(F_d + F_e)$), *a block design, an* m*-sequence, and the design maximizing* min-Φ_A *selected from* 100 *randomly generated designs. Designs are ordered by their* min-Φ_A *values.*

ear models, these designs can yield high performance for detecting brain activation [Maus et al. (2011, 2010a), Henson (2007)]. An m-sequence can be generated from primitive polynomials for a Galois field [Buračas and Boynton (2002), Godfrey (1993), MacWilliams and Sloane (1977)]. These designs can be obtained from a MATLAB program provided by Liu (2004) and are good for estimating the HRF. The max-F_d, max-F_e, and bi-objective ($0.5F_d + 0.5F_e$) designs are obtained by the GA of Kao et al. (2009) with linear models. A max-F_d design maximizes the efficiency of detection, whereas a max-F_e design maximizes the HRF estimation efficiency. The bi-objective designs maximize the average of theses two efficiencies. They offer a compromise between the two competing objectives of detection and estimation. We also generate 100 random designs and select the one yielding the maximal min-Φ_A. When lacking design tools for sophisticated experimental settings, as considered here, random designs are not uncommon in practice. More details about these designs can be found in Kao et al. (2009) and Liu (2004).

The designs in Figure 2 are ordered by their min-Φ_A values. Clearly, the maximin designs are much better than the other designs and have relatively small dispersions in Φ_A-values across the parameter space. This indicates that the estimation precisions yielded by the maximin designs are quite robust against a misspecified parameter vector value. We also observe that, while the block and max-F_d designs are recommended for detecting activation under linear models, they do not perform well for detection under the nonlinear model. The Φ_A-value of these two types of designs can vary greatly over the parameter space, and, at the worst cases, their Φ_A-values can be very low, indicating poor precisions in estimating $\boldsymbol{\theta}$.

3.2. *Maximin-efficient designs.* Our proposed methods are also applied to obtain maximin-efficient designs. For $Q = 1$, we first use the GA to search for the required locally optimal designs for each grid point on $\{0, 1\} \times \mathcal{P}$, and then a design maximizing $\min_{\{0,1\}\times\mathcal{P}} \mathrm{RE}(d; \boldsymbol{\theta}, \mathbf{p})$. Based on Corollary 2, the GA actually yields a maximin-efficient design d_{MmE} for $\mathbb{R}^1 \times \mathcal{P}$ even though the reduced parameter space is considered. For $Q = 2$ and 3, we consider the Θ_0 presented in the previous subsection. We then apply the GA to search for (1) locally optimal designs over $\{\{\mathbf{0}\} \cup \Theta_0\} \times \mathcal{P}$; and (2) a maximin-efficient designs $d_{\mathrm{MmE},\Theta_0}$ optimizing $\min_{\{\{\mathbf{0}\}\cup\Theta_0\}\times\mathcal{P}} \mathrm{RE}(d; \boldsymbol{\theta}, \mathbf{p})$. The $d_{\mathrm{MmE},\Theta_0}$ designs are obtained over the entire design space Ξ (Strategy 1) and over the subclass Ξ_0 (Strategy 2).

Table 2, to be read as Table 1, presents a comparison among the obtained maximin-efficient designs. By omitting the time needed for obtaining locally optimal designs, the CPU times in Table 2 for obtaining the maximin-efficient designs are similar to those for maximin designs in Table 1. However, obtaining maximin-efficient designs requires locally optimal designs. This unfortunately makes maximin-efficient designs computationally much more expensive than maximin designs, especially when Q becomes large. Specifically, for $Q = 1$, we use the GA to obtain 352 locally optimal designs, each requiring about 2.88 s. The GA takes about 4 hours to find 1232 locally optimal designs for $Q = 2$, and about 46

TABLE 2
The performance and mean CPU time (in minutes) for d_{MmE} *obtained over* Ξ *(all designs) for* $Q = 1$; *and* $d_{\mathrm{MmE},\Theta_0}$ *obtained over* Ξ *and the subclass* Ξ_0 *for* $Q = 2$ *and* 3

	$Q=1$	$Q=2$		$Q=3$	
	Ξ	Ξ	Ξ_0	Ξ	Ξ_0
min-RE					
Maximum	0.835	0.790	0.829	0.797	0.829
Mean	0.830	0.783	0.820	0.783	0.823
Std. err.	0.001	0.002	0.002	0.003	0.001
Mean CPU time	0.88[1]	11.69[2]	5.88[2]	207.51[3]	52.04[3]

[1] Additional 17 minutes are needed for finding the required locally optimal designs.

[2] Additional 4 hours are needed for finding the required locally optimal designs.

[3] Additional 46 hours are needed for finding the required locally optimal designs.

hours to generate 5984 locally optimal designs for $Q = 3$. Results in Table 2 also indicate that Strategy 2 outperforms Strategy 1 in terms of the achieved design efficiency and required CPU time. Strategy 2 is thus recommended.

In Figure 3 we compare the RE-values over $\Theta \times \mathcal{P}$ of the maximin-efficient designs and the traditional designs introduced in the previous subsection. The selected maximin-efficient design for $Q = 1$ maximizes min-RE over the ten d_{MmE} designs; the maximin-efficient designs for $Q > 1$ are selected from the ten

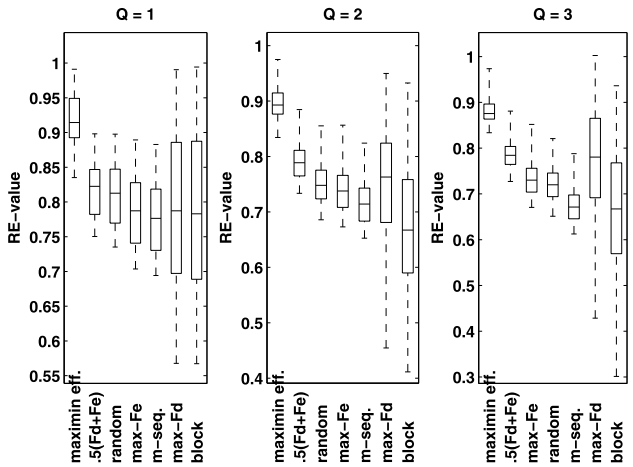

FIG. 3. *Boxplots of the* RE*-values over* $\Theta \times \mathcal{P}$ *of maximin-efficient designs*, max-F_e, max-F_d, *and bi-objective* $(0.5(F_d + F_e))$ *designs*, *a block design*, *an m-sequence*, *and the design maximizing* min-RE *selected from* 100 *randomly generated designs*. *Designs are ordered by their* min-RE *values*.

$d_{\mathrm{MmE},\Theta_0}$ designs obtained by Strategy 2. As presented in the figure, our designs significantly outperform the traditional designs.

3.3. *An example*. In this subsection we consider an experimental setting employed by Miezin et al. (2000), in which a 1.5-s 8-Hz flickering checkerboard (stimulus) is presented interlaced with a visual fixation (control). Upon the onset of each checkerboard, subjects responded by pressing a key with their right hands. The minimal time between consecutive stimulus onsets was 2.5 s (ISI = 2.5 s). The BOLD time series was acquired every 2.5 s (TR = 2.5 s). The experimenters presented the same design twice to a subject with a 2-minute rest period in between the two runs. Each run lasted about 5.5 minutes. To allow an effective sampling rate of the hemodynamic response, stimulus onsets were synchronized with MR scans in the first run and were shifted 1.25 s in the second run.

Miezin and colleagues demonstrated that the time-to-peak (p_1) and time-to-onset (p_6) of the HRF can vary across brain voxels. Taking this uncertainty into account, we apply our proposed approach to obtain maximin and maximin-efficient designs. A simple modification is needed to accommodate the special requirement that the study is conducted over two runs. Specifically, we replace the design matrix in model (2.1) by $\mathrm{diag}(\mathbf{X}_{d,1}, \mathbf{X}_{d,1})$ since the same sequence of stimuli is presented twice. In addition, $\mathbf{h}(\mathbf{p})$ is now $(\mathbf{h}_1(\mathbf{p})', \mathbf{h}_2(\mathbf{p})')'$, where the jth element of $\mathbf{h}_1(\mathbf{p})$ is $g(2.5(j-1); \mathbf{p})$ and that of $\mathbf{h}_2(\mathbf{p})$ is $g(1.25+2.5(j-1); \mathbf{p})$. This accounts for the difference of 1.25 s in the HRF sampling time points between the two runs. We also consider the nuisance term $[(\mathbf{S}\boldsymbol{\gamma}_1)', (\mathbf{S}\boldsymbol{\gamma}_2)']'$ that allows run effects, where $\mathbf{S}\boldsymbol{\gamma}_i$ corresponds to a second-order polynomial drift, $i = 1, 2$. The noise of the two runs are assumed to be two independent AR(1) processes with autocorrelation coefficient $\rho = 0.3$. The whitening matrix thus has the form $(\mathbf{I}_2 \otimes \mathbf{V})$, where $\mathbf{V}$ is a whitening matrix for each run [see also Kao, Mandal and Stufken (2009)]. We also investigated the performance of our obtained designs when $\rho = 0$ or 0.5, and found that our designs are still quite efficient with a different ρ-value.

In addition to maximin-type designs, we generate a block design, m-sequence-based design, and 100 designs that are random permutations of a design consisting of 50% "0"s and 50% "1"s. The block design is formed by repeating {000000111111}, which has the 15s-off-15s-on pattern that is often recommended for detecting brain activation. An m-sequence does not exist in this case. We thus follow Liu (2004) to generate an m-sequence-based design by concatenating an m-sequence of length 127 with its first 5 elements. The 100 randomly permuted designs are constructed to mimic the design considered by Miezin et al. (2000). Each of these 100 designs are selected so that the average time between consecutive stimulus onsets is within 4.9 s and 5.1 s. Among these 100 designs, we select the one yielding the maximal min-Φ_A value when comparing with the maximin design, and the one maximizing min-RE when comparing with the maximin-efficient design. Figure 4 provides summaries of the performances of these designs. As

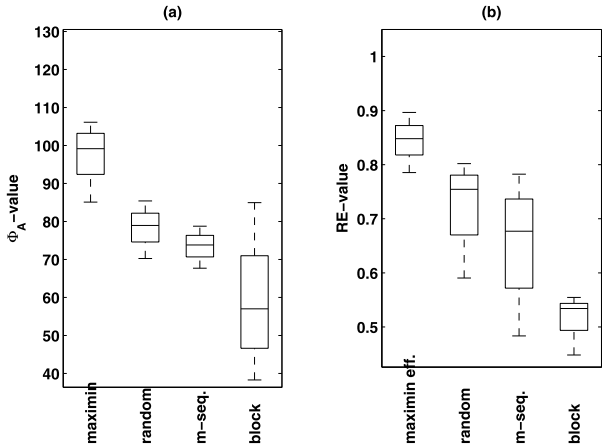

FIG. 4. *Boxplots of the* (a) Φ_A*-values, and* (b) RE*-values of the competing designs over* $\{1\} \times \mathcal{P}$.

shown in the figure, our proposed methods consistently generate high-quality designs that significantly outperform the traditional designs.

We also explore the performance of the designs with different ρ-values. When $\rho = 0$, the maximin design for $\rho = 0.3$ attains 95.1% of the min-Φ_A value of the maximin designs for $\rho = 0.0$. The min-RE value of maximin-efficient design for $\rho = 0.3$ is 98.1% of that of the maximin-efficient design for $\rho = 0.0$. For cases where ρ is as large as $\rho = 0.5$, the relative min-Φ_A of the maximin design for $\rho = 0.3$ to the maximin design for $\rho = 0.5$ is 97.2%. For this same condition, the maximin-efficient design for $\rho = 0.3$ retains 92.0% of the min-RE of the maximin-efficient design for $\rho = 0.5$. Our obtained designs, especially the maximin designs, perform relatively well when comparing with the best design for a ρ-value that is as small as 0 and as high as 0.5.

4. Conclusions. We obtain high-quality experimental designs for fMRI experiments to help to render efficient statistical inference on brain activity with a nonlinear model. In contrast to linear models, the nonlinear model allows us to detect brain voxels activated by the mental stimuli while the uncertain HRF shape is taken into account. However, optimal designs for the nonlinear model depend on unknown model parameters, making the design problem notoriously difficult. To tackle this problem, we consider maximin and maximin-efficient designs and propose efficient approaches for obtaining these designs. Our approaches involve a large reduction in the parameter space, a restricted class of ER-fMRI designs, and the use of the knowledge-based GA of Kao et al. (2009) for searching for maximin-type designs. These approaches, especially Strategy 2, are demonstrated to be powerful via case studies and a real example.

Maximin and maximin-efficient designs are widely accepted, although obtaining them is almost always difficult. Pronzato and Walter (1988) studied both types

of designs and concluded that maximin designs have definite advantages over maximin-efficient designs when reducing the worst possible uncertainty for estimating the parameter is of concern. On the other hand, Dette and Biedermann (2003) considered the maximin-efficient criterion because it tends to avoid placing too much attention on a certain parameter vector value. Huang and Lin (2006) suggested that both maximin-type criteria deserve consideration; the selection may thus be guided by the need and preference of the experimenter. Our approaches allow us to efficiently find both maximin and maximin-efficient designs, although the latter designs are computationally more expensive than the former designs.

We also observe that, while blocked designs are recommended for detecting brain activation under linear models, they are very poor for the same objective under nonlinear models with uncertain HRF shapes. We believe that the inferiority of block designs is mainly due to their low efficiencies in estimating the HRF [e.g., Liu and Frank (2004)]. The estimation of the HRF shape, although not the main concern, is needed when the HRF shape is uncertain. A good design should thus allow for a reasonable efficiency in performing this task. Designs with random components tend to serve this purpose well.

The designs that we found are for a nonlinear model, in which the HRF is approximated by $\theta\mathbf{h}(\mathbf{p})$, the product of the unknown HRF amplitude θ and the uncertain HRF shape $\mathbf{h}(\mathbf{p})$. Such models are not uncommon in the literature [Handwerker, Ollinger and D'Esposito (2004), Lindquist and Wager (2007), Miezin et al. (2000)]. While we focus on an $\mathbf{h}(\mathbf{p})$ having the same form as the popular double-gamma function of SPM, the proposed approaches can be extended to $\mathbf{h}(\mathbf{p})$'s of other forms, for example, the inverse logit function considered by Lindquist and Wager (2007).

When implementing our approaches in the case studies, we consider an AR(1) noise with known constant autocorrelation coefficient $\rho = 0.3$. This selection is guided mainly by previous studies, for example, Lenoski et al. (2008) and Worsley et al. (2002). From these studies, the use of AR(1) noise tends to provide satisfactory analysis results. With AR(1), the results of Maus et al. (2010b) for linear models suggest that the obtained designs for $\rho = 0.3$ do not suffer a significant loss in design efficiency under other values of $\rho \in [0, 0.5]$. We also observe a similar outcome under the nonlinear model.

We also note that the assumed AR(1) model with $\rho = 0.3$ could be idealistic. First, other models for autocorrelated noise might be more appropriate for some data [e.g., Lindquist (2008)]. In addition, whether for an AR(1) model or another model, knowledge about the unknown, possibly nonconstant parameter(s) may not always be available at the design stage. For a selected model, our methods could then be combined with the approach of Maus et al. (2010b) to search for optimal designs (using the maximin or maximin-efficient criterion) by taking uncertain HRF shape and autocorrelation parameters into account. While such a maximin-type approach can be generalized to accommodate different models for autocorrelated noise, the design problem can become very challenging. Developing an

efficient method for cases where both HRF shape and correlation are uncertain is a topic of future research.

APPENDIX: A PROOF OF LEMMA 3

The following two lemmas are straightforward. Their proofs are thus omitted. The notation used is as in Lemma 3.

LEMMA 5. *For a permutation matrix* $\mathbf{G}_g$, *let* $k_{G_g}(d)$ *be the design obtained by relabeling the stimulus types, the same way as* $\mathbf{G}_g$ *permutes* $(1, 2, \ldots, Q)'$, *of a design* d. *We have* $\mathbf{M}(k_{G_g}(d); \mathbf{G}_g\boldsymbol{\theta}, \mathbf{p}) = \mathbf{G}_g'\mathbf{M}(d; \boldsymbol{\theta}, \mathbf{p})\mathbf{G}_g$ *and, thus,* $\Phi_A(k_{G_g}(d); \mathbf{G}_g\boldsymbol{\theta}, \mathbf{p}) = \Phi_A(d; \boldsymbol{\theta}, \mathbf{p})$ *for any* $(\theta, \mathbf{p}) \in \Theta_0 \times \mathcal{P}$.

LEMMA 6. *The following two conditions are equivalent*: (1) d_0^* *is a maximin design for* $\Theta_0 \times \mathcal{P}$; *and* (2) $k_{G_g}(d_0^*)$ *is a maximin design for* $\Theta_g \times \mathcal{P}$ *for any* g.

PROOF OF LEMMA 3. For a d_{Mm,Θ_0} satisfying the conditions of Lemma 3, we have $\min_{\Theta_g \times \mathcal{P}} \Phi_A(k_{G_g}(d_{\mathrm{Mm},\Theta_0}); \boldsymbol{\theta}, \mathbf{p}) = \min_{\Theta_0 \times \mathcal{P}} \Phi_A(d_{\mathrm{Mm},\Theta_0}; \boldsymbol{\theta}, \mathbf{p}) = \min_{\Theta_g \times \mathcal{P}} \Phi_A(d_{\mathrm{Mm},\Theta_0}; \boldsymbol{\theta}, \mathbf{p})$ for any g. Therefore, d_{Mm,Θ_0}, which is a maximin design for $\Theta_0 \times \mathcal{P}$, is also a maximin design for $\Theta_g \times \mathcal{P}$ for any g, and for $(\bigcup_{g=0}^{G} \Theta_g) \times \mathcal{P} = \Theta \times \mathcal{P}$. □

Acknowledgments. We thank anonymous referees for raising questions that resulted in an improvement of this article.

SUPPLEMENTARY MATERIAL

Supplement to "Maximin and maximin-efficient event-related FMRI designs under a nonlinear model" (DOI: 10.1214/13-AOAS658SUPP; .pdf). We provide (1) a proof that the Φ_A-value of a design in the restricted design class Ξ_0 is insensitive to permutations of the elements of $\boldsymbol{\theta}$; and (2) a genetic algorithm for obtaining ER-fMRI designs.

REFERENCES

BERGER, M. P. F., KING, C. Y. J. and WONG, W. K. (2000). Minimax D-optimal designs for item response theory models. *Psychometrika* **65** 377–390. MR1792702

BERGER, M. P. F. and WONG, W. K. (2009). *An Introduction to Optimal Designs for Social and Biomedical Research*. Wiley, Chichester.

BOOKHEIMER, S. (2007). Pre-surgical language mapping with functional magnetic resonance imaging. *Neuropsychology Review* **17** 145–155.

BOSE, M. and STUFKEN, J. (2007). Optimal crossover designs when carryover effects are proportional to direct effects. *J. Statist. Plann. Inference* **137** 3291–3302. MR2363257

BOX, G. E. P. and LUCAS, H. L. (1959). Design of experiments in non-linear situations. *Biometrika* **46** 77–90. MR0102155

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

CHEN, R.-B., WONG, W. K. and LI, K.-Y. (2008). Optimal minimax designs over a prespecified interval in a heteroscedastic polynomial model. *Statist. Probab. Lett.* **78** 1914–1921. MR2528561

CHEN, R. B., CHANG, S. P., WANG, W. and WONG, W. K. (2011). Optimal experimental designs via particle swarm optimization methods. Technical report. Available at http://www.math.ntu.edu.tw/~mathlib/preprint/2011-03.pdf.

CHENG, C.-S. (1996). Optimal design: Exact theory. In *Design and Analysis of Experiments. Handbook of Statistics* **13** 977–1006. North-Holland, Amsterdam. MR1492588

CHERNOFF, H. (1953). Locally optimal designs for estimating parameters. *Ann. Math. Statist.* **24** 586–602. MR0058932

CULHAM, J. C. (2006). Functional neuroimaging: Experimental design and analysis. In *Handbook of Functional Neuroimaging of Cognition*, 2nd ed. (R. Cabeza and A. Kingstone, eds.) 53–82. MIT Press, Cambridge, MA.

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

D'ESPOSITO, M., ZARAHN, E. and AGUIRRE, G. K. (1999). Event-related functional MRI: Implications for cognitive psychology. *Psychol. Bull.* **125** 155–164.

DETTE, H. and BIEDERMANN, S. (2003). Robust and efficient designs for the Michaelis–Menten model. *J. Amer. Statist. Assoc.* **98** 679–686. MR2011681

DETTE, H., HAINES, L. M. and IMHOF, L. A. (2007). Maximin and Bayesian optimal designs for regression models. *Statist. Sinica* **17** 463–480. MR2408676

FEDOROV, V. V. and HACKL, P. (1997). *Model-Oriented Design of Experiments. Lecture Notes in Statistics* **125**. Springer, New York. MR1454123

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.

GODFREY, K. (1993). *Perturbation Signals for System Identification*. Prentice Hall, New York.

HANDWERKER, D. A., OLLINGER, J. M. and D'ESPOSITO, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* **21** 1639–1651.

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

HENSON, R. N. A. (2007). Efficient experimental design for fMRI. In *Statistical Parametric Mapping*: *The Analysis of Functional Brain Images* (K. J. Friston, J. T. Ashburner, S. J. Kiebel, T. E. Nichols and W. D. Penny, eds.) 193–210. Academic Press, London.

HUANG, M.-N. L. and LIN, C.-S. (2006). Minimax and maximin efficient designs for estimating the location-shift parameter of parallel models with dual responses. *J. Multivariate Anal.* **97** 198–210. MR2208849

HUETTEL, S. A. (2012). Event-related fMRI in cognition. *Neuroimage* **62** 1152–1156.

JOSEPHS, O. and HENSON, R. N. A. (1999). Event-related functional magnetic resonance imaging: Modelling, inference and optimization. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **354** 1215–1228.

KAO, M. H. (2009). Optimal experimental designs for event-related functional magnetic resonance imaging. Ph.D. thesis, Univ. Georgia, Athens, GA.

KAO, M.-H., MANDAL, A. and STUFKEN, J. (2009). Efficient designs for event-related functional magnetic resonance imaging with multiple scanning sessions. *Comm. Statist. Theory Methods* **38** 3170–3182. MR2568211

KAO, M.-H., MANDAL, A. and STUFKEN, J. (2012). Constrained multiobjective designs for functional magnetic resonance imaging experiments via a modified non-dominated sorting genetic algorithm. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **61** 515–534. MR2960736

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.

KAO, M. H., MAJUMDAR, D., MANDAL, A. and STUFKEN, J. (2013). Supplement to "Maximin and maximin-efficient event-related FMRI designs under a nonlinear model." DOI:10.1214/13-AOAS658SUPP.

KING, J. and WONG, W.-K. (2000). Minimax D-optimal designs for the logistic model. *Biometrics* **56** 1263–1267. MR1816369

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

LENOSKI, B., BAXTER, L. C., KARAM, L. J., MAISOG, J. and DEBBINS, J. (2008). On the performance of autocorrelation estimation algorithms for fMRI analysis. *IEEE Journal of Selected Topics in Signal Processing* **2** 828–838.

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

LINDQUIST, M. A. and WAGER, T. D. (2007). Validity and power in hemodynamic response modeling: A comparison study and a new approach. *Hum. Brain Mapp.* **28** 764–784.

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.

LOH, J. M., LINDQUIST, M. A. and WAGER, T. D. (2008). Residual analysis for detecting mismodeling in fMRI. *Statist. Sinica* **18** 1421–1448. MR2468275

MACWILLIAMS, F. J. and SLOANE, N. J. A. (1977). *The Theory of Error Correcting Codes*. Elsevier, New York.

MAUS, B., VAN BREUKELEN, G. J. P., GOEBEL, R. and BERGER, M. P. F. (2010a). Optimization of blocked designs in fMRI studies. *Psychometrika* **75** 373–390. MR2719934

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

MAUS, B., VAN BREUKELEN, G. J. P., GOEBEL, R. and BERGER, M. P. F. (2011). Optimal design of multi-subject blocked fMRI experiments. *NeuroImage* **56** 1338–1352.

MAUS, B., VAN BREUKELEN, G. J. P., GOEBEL, R. and BERGER, M. P. F. (2012). Optimal design for nonlinear estimation of the hemodynamic response function. *Hum. Brain Mapp.* **33** 1253–1267.

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.

MÜLLER, C. H. (1995). Maximin efficient designs for estimating nonlinear aspects in linear models. *J. Statist. Plann. Inference* **44** 117–132. MR1323074

PRONZATO, L. and WALTER, E. (1988). Robust experiment design via maximin optimization. *Math. Biosci.* **89** 161–176. MR0942855

ROSEN, B. R., BUCKNER, R. L. and DALE, A. M. (1998). Event-related functional MRI: Past, present, and future. *Proc. Natl. Acad. Sci. USA* **95** 773–780.

SILVEY, S. D. (1980). *Optimal Design*: *An Introduction to the Theory for Parameter Estimation*. Chapman & Hall, London. MR0606742

SITTER, R. R. (1992). Robust designs for binary data. *Biometrics* **48** 1145–1155. MR1212858

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.

WAGER, T. D., VAZQUEZ, A., HERNANDEZ, L. and NOLL, D. C. (2005). Accounting for nonlinear BOLD effects in fMRI: Parameter estimates and a model for prediction in rapid event-related studies. *NeuroImage* **25** 206–218.

WIERENGA, C. and BONDI, M. (2007). Use of functional magnetic resonance imaging in the early identification of Alzheimer's disease. *Neuropsychology Review* **17** 127–143.

WORSLEY, K. J. and TAYLOR, J. E. (2006). Detecting fMRI activation allowing for unknown latency of the hemodynamic response. *Neuroimage* **29** 649–654.

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.

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

D. MAJUMDAR
DEPARTMENT OF MATHEMATICS, STATISTICS
AND COMPUTER SCIENCE
UNIVERSITY OF ILLINOIS AT CHICAGO
CHICAGO, ILLINOIS 60607
USA
E-MAIL: dibyen@uic.edu

A. MANDAL
J. STUFKEN
DEPARTMENT OF STATISTICS
UNIVERSITY OF GEORGIA
ATHENS, GEORGIA 30602
USA
E-MAIL: amandal@stat.uga.edu
jstufken@stat.uga.edu