

Robust Event-Related fMRI Designs Under A Nonlinear Model

Ming-Hung Kao, Dibyen Majumdar, Abhyuday Mandal, and John Stufken*

Abstract

Previous studies on event-related functional magnetic resonance imaging (ER-fMRI) 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 an approach for obtaining maximin efficient designs. Our approach involves a reduction in the parameter space and an efficient search algorithm. The designs that we obtain are much more robust against mis-specified HRF shapes than designs widely used by researchers.

KEY WORDS: A-optimality; Genetic algorithms; Hemodynamic response function; Information matrix; Maximin efficient designs

*Ming-Hung Kao is Assistant Professor, School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287 (E-mail: mhkao@math.asu.edu). Dibyen Majumdar is Associate Dean, College of Liberal Arts and Sciences, and Professor, Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL 60607. Abhyuday Mandal is Assistant Professor, and John Stufken is Professor and Department Head, Department of Statistics, University of Georgia, Athens, GA 30602. The research of Abhyuday Mandal was in part supported by NSF Grant DMS-09-05731, and that of John Stufken by NSF Grants DMS-07-06917 and DMS-10-07507.

1. INTRODUCTION

Functional magnetic resonance imaging (fMRI) is a pioneering, noninvasive brain mapping technology. This cutting-edge technology is widely used in cognitive neuroscience for studying brain functions (Culham 2006; D’Esposito, Zarahn, and Aguirre 1999). It has many important clinical potentials such as early identification of Alzheimer’s disease, pre-neurosurgical planning and post-neurosurgical evaluations (e.g. Bookheimer 2007; Wierenga and Bondi 2007); see also Lazar (2008) and Lindquist (2008) for overviews of fMRI.

In an fMRI experiment, a sequence of mental stimuli interlaced with a control (e.g., fixation or rest) is presented to a subject while an MR scanner scans the subject’s brain to collect a blood oxygenated level dependent (BOLD) time series from each brain voxel (e.g. a $3 \times 3 \times 5 \text{ mm}^3$ three-dimensional imaging unit). These BOLD time series help to make statistical inference about brain activity. To render precise and valid inference, high quality fMRI designs that help to collect informative data are crucially important. In this article, we aim at obtaining good event-related (ER) fMRI designs. These designs are sequences of brief stimuli ($\sim 1\text{s}$). Due to their flexibility, such designs are very popular (Josephs and Henson 1999), and are the main focus of existing researches on fMRI designs (Kao, Mandal, Lazar, and Stufken 2009a; Liu and Frank 2004; Liu 2004; Wager and Nichols 2003; Liu, Frank, Wong, and Buxton 2001).

Existing studies as well as current knowledge on fMRI designs are primarily based on the linear model framework (Friston et al. 1995; Worsley and Friston 1995). While popular, this framework is criticized by some researchers (Loh, Lindquist, and Wager 2008; Worsley and Taylor 2006; Handwerker, Ollinger, and D’Esposito 2004). A major

criticism is on the assumption of a fixed, known shape of the hemodynamic response function (HRF; a function of time describing the MR signal change due to a single stimulus). This assumption may not always be valid. Studies showed that the HRF shape can vary across subjects, scanning sessions and brain voxels. The shape is often uncertain. An a priori fixed shape can be wrong, leading to incorrect conclusions. To tackle this drawback, analysis methods such as the use of nonlinear models allowing uncertain HRF shapes are proposed; e.g, Lindquist and Wager (2007), Handwerker et al. (2004), and Miezin, Maccotta, Ollinger, Petersen, and Buckner (2000). However, to the best of our knowledge, this important issue has not been systematically studied at the design stage.

In the present study, we take the uncertainty of the HRF shape into account and obtain good designs for detecting activated brain voxels, a common objective of fMRI experiments. In contrast to previous studies, we consider at the design stage a nonlinear model that can accommodate a wide spectrum of feasible HRF shapes. The design problem becomes much more complicated because identifying good designs will depend on the unknown model parameters. One possible approach is to find a locally optimal design, which is the best for a specified value of the parameter vector (Chernoff 1953). Unfortunately, a good guess of the value of the parameter vector is almost always unavailable. More importantly, a good design should be efficient for various parameter vector values (or HRF shapes) associated with all the brain voxels of interest. Focusing on one, single value of the parameter vector is unsatisfactory. Here, we consider another approach, namely the maximin approach, that is much more appealing. We obtain designs that perform relatively well over a given range of each parameter. Specifically,

the maximin approach works on the worst case scenario and aims at designs maximizing the minimum precision of parameter estimates, the minimum is taken over the specified parameter space; see also Berger and Wong (2009) and Silvey (1980).

Unfortunately, the maximin ER-fMRI design problem is mathematically intractable and computationally difficult, if not infeasible. To make this problem manageable, we propose an approach to reduce the size of the parameter space, and then adopt the knowledge-based genetic algorithm (GA) of Kao et al. (2009a) to search for maximin efficient designs, which are highly efficient in terms of our defined maximin criterion. We demonstrate the usefulness of our proposed approach via case studies and a real example. Compared with designs widely used by researchers, our designs are much more robust against mis-specified HRF shapes.

The remainder of the article is organized as follows. In Section 2, we introduce our method including the underlying statistical model, design criterion, and an efficient maximin approach. 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 a control. During the presentation, each stimulus may last several milliseconds to a few seconds. Times between consecutive stimulus onsets are multiples of a pre-specified inter-stimulus-interval (ISI; e.g., 4s). The control fills in the time when no stimulus

is being presented. A design can be written as a sequence of finite numbers; e.g., $d = \{101210 \cdots 1\}$. An integer q ($\neq 0$) at the k th position of a sequence indicates an onset of a q th-type stimulus at time $(k - 1)ISI$. A ‘0’ means no stimulus onset at that time point.

During an fMRI experiment, such a sequence of stimuli is presented to the subject. Each stimulus evokes a change in the MR signal intensity at an activated brain voxel. The signal intensity takes about 25 to 30 seconds to rise and decay. This change is typically described by an HRF having an assumed shape with an unknown amplitude. The HRF shape is popularly assumed to be the double-gamma function of SPM (<http://www.fil.ion.ucl.ac.uk/spm/>), a widely used software for analyzing fMRI data.

When the next stimulus appears 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. 2s) to form the BOLD time series. Denoting the time series of a voxel by \mathbf{y} (T -by-1), we consider in this study the following model:

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

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 q th-type stimuli of a design d . The scalar θ_q is the unknown HRF amplitude for the q th-type stimuli. 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 $(ISI/\Delta T)$ and $(TR/\Delta T)$ integers. $\mathbf{X}_{d,q}$ is the zero-one design matrix with 1 indicating the heights of the HRF that contribute to each BOLD measurement (for more details, see Kao, Mandal, and

Stufken 2008). The nuisance term $\mathbf{S}\boldsymbol{\gamma}$ allows a drift/trend over time with an unknown parameter vector $\boldsymbol{\gamma}$. The correlated noise is denoted by \mathbf{e} . The focus of detection problems is on $\boldsymbol{\theta} = (\theta_1, \dots, \theta_Q)$; a large θ_q -value signals a voxel that is highly activated by the q th-type stimuli, $q = 1, \dots, Q$. In contrast to previous studies, the value of \mathbf{p} in model (1) is unknown. This allows an uncertain HRF shape.

Specifically, $\mathbf{h}(\mathbf{p})$ in model (1) is determined by a continuous function $g(t; \mathbf{p})$, where t represents time elapsed after a stimulus onset. More precisely, the j th element of $\mathbf{h}(\mathbf{p})$ is $g(\Delta T(j-1); \mathbf{p})$. 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 (see also, Wager, Vazquez, Hernandez, and Noll 2005):

$$g(t; \mathbf{p}) = \begin{cases} \frac{g_0(t; \mathbf{p})}{\max_s g_0(s; \mathbf{p})}, & \text{if } t \in [\max(0, p_6), p_7]; \\ 0, & \text{otherwise;} \end{cases}$$

$$g_0(t; \mathbf{p}) = f\left(t - p_6, \frac{p_1}{p_3}, p_3\right) - p_5 f\left(t - p_6, \frac{p_2}{p_4}, p_4\right); \quad f(x, \alpha, \beta) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha};$$

$\Gamma(\cdot)$ is the gamma function; and $f(x, \alpha, \beta)$ is the density function of the gamma distribution, $\text{gamma}(\alpha, \beta)$. We note that the double-gamma function of SPM has a fixed $(p_1, p_2, \dots, p_7) = (6, 16, 1, 1, 1/6, 0, 32)$. By contrast, we follow Wager et al. (2005) to treat two influential HRF parameters, namely p_1 , time-to-peak, and p_6 , time-to-onset, as free parameters, while keeping $(p_2, p_3, p_5, p_4, p_7)$ fixed at $(16, 1, 1/6, 1, 32)$. Without confusion, we will use \mathbf{p} to denote the unknown parameter vector (p_1, p_6) hereinafter.

2.2 The design criterion

We aim at a good design for detecting activation (or studying $\boldsymbol{\theta}$) using model (1). The goodness of a design will be evaluated by $Q/\text{trace}(\text{Cov}[\hat{\boldsymbol{\theta}}])$, the reciprocal of the average variance of the estimators $\hat{\boldsymbol{\theta}}$; i.e., the A -optimality criterion (e.g. Pukelsheim 1993). Following a popular technique (Box and Lucas 1959; Fedorov and Hackl 1997), we first linearize model (1) and then use the linearized model to approximate $\text{Cov}[\hat{\boldsymbol{\theta}}]$. The approximation is $\mathbf{M}^{-1}(d; \boldsymbol{\theta}, \mathbf{p})$, where

$$\begin{aligned}\mathbf{M}(d; \boldsymbol{\theta}, \mathbf{p}) &= \mathbf{E}_d(\mathbf{p})' [\mathbf{I}_T - w\{\mathbf{L}_d(\boldsymbol{\theta}, \mathbf{p})\}] \mathbf{E}_d(\mathbf{p}), \\ \mathbf{E}_d(\mathbf{p}) &= [\mathbf{I}_T - w\{\mathbf{V}\mathbf{S}\}] \mathbf{V}\mathbf{X}_d [\mathbf{I}_Q \otimes \mathbf{h}(\mathbf{p})], \\ \mathbf{L}_d(\boldsymbol{\theta}, \mathbf{p}) &= [\mathbf{L}_1, \mathbf{L}_6], \quad \mathbf{L}_i = [\mathbf{I}_T - w\{\mathbf{V}\mathbf{S}\}] \mathbf{V}\mathbf{X}_d \left[\mathbf{I}_Q \otimes \frac{\partial \mathbf{h}(\mathbf{p})}{\partial p_i} \right] \boldsymbol{\theta}, i = 1, 6,\end{aligned}$$

\mathbf{I}_a is the a -by- a identity matrix, $w\{\mathbf{A}\} = \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ is the orthogonal projection matrix onto the vector space spanned by the column vectors of \mathbf{A} , \mathbf{A}^- is a generalized inverse matrix of \mathbf{A} , \mathbf{V} is a known, whitening matrix so that $\mathbf{V}\mathbf{e}$ is white noise, $\mathbf{X}_d = [\mathbf{X}_{d,1}, \dots, \mathbf{X}_{d,Q}]$, and \otimes is the Kronecker product. 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 $\Phi_A(d; \boldsymbol{\theta}, \mathbf{p}) \equiv Q/\text{trace}(\mathbf{M}^{-1}(d; \boldsymbol{\theta}, \mathbf{p}))$. However, the answer will depend on the unknown $\boldsymbol{\theta}$ and \mathbf{p} . This makes such a nonlinear design problem notoriously difficult. To tackle this problem, we present a maximin approach in the next subsection. The maximin approach aims at a design maximizing

$$\min_{(\boldsymbol{\theta}, \mathbf{p}) \in \Theta \times \mathcal{P}} \Phi_A(d; \boldsymbol{\theta}, \mathbf{p}) \tag{2}$$

for a specified parameter space $\Theta \times \mathcal{P}$ for $(\boldsymbol{\theta}, \mathbf{p})$. To reflect an observation that the HRF typically increases in 0-2 seconds after the stimulus onset, reaches the peak in 5-8 seconds, and then falls back to baseline (Lindquist 2008; Rosen, Buckner, and Dale 1998), we set $\mathcal{P} = \{(p_1, p_6) \mid p_1 \in [6, 9], p_6 \in [0, 2]\}$. The range of p_1 is inspired by the fact that the mode of the gamma distribution $\text{gamma}(\alpha, 1)$ 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 A maximin approach

Obtaining a design maximizing the minimum of $\Phi_A(d; \boldsymbol{\theta}, \mathbf{p})$ over $\mathbb{R}^Q \times \mathcal{P}$ is arduous. We find the following observations useful in reducing the computational burden.

Lemma 1. $\mathbf{M}^{-1}(d; \mathbf{0}, \mathbf{p}) \leq \mathbf{M}^{-1}(d; \boldsymbol{\theta}, \mathbf{p})$ 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 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, p. 3300). It suggests that the Φ_A -value depends on the direction of $\boldsymbol{\theta}$, but not on its length. Therefore, 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 a larger Q , we can 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. In particular, 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 $\Theta = \{(\theta_1, \dots, \theta_Q)\}$, where

$$\begin{aligned}\theta_1 &= \cos \varphi_1; \theta_q = \cos \varphi_q \prod_{i=1}^{q-1} \sin \varphi_i, \quad q = 2, \dots, Q-1; \\ \theta_Q &= \prod_{i=1}^{Q-1} \sin \varphi_i; \quad \varphi_1, \dots, \varphi_{Q-1} \in (-\pi/2, \pi/2].\end{aligned}$$

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 $\mathcal{G} = \{\mathbf{G}_1, \dots, \mathbf{G}_G\}$ 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_0^* is a maximin design for $\Theta_0 \times \mathcal{P}$ and $\min_{\Theta_0 \times \mathcal{P}} \Phi_A(d_0^*; \boldsymbol{\theta}, \mathbf{p}) = \min_{\Theta_g \times \mathcal{P}} \Phi_A(d_0^*; \boldsymbol{\theta}, \mathbf{p})$ for any g , then d_0^* is also a maximin design for $\Theta \times \mathcal{P}$.

A proof of Lemma 3 can be found in Appendix A. This result motivates the following strategy: (1) identify a Θ_0 and \mathcal{G} ; (2) obtain a design d_0^* maximizing $\min_{\Theta_0 \times \mathcal{P}} \Phi_A(d; \boldsymbol{\theta}, \mathbf{p})$, for which the ratio

$$\mathcal{R}_g = \frac{\min_{\Theta_g \times \mathcal{P}} \Phi_A(d_0^*; \boldsymbol{\theta}, \mathbf{p})}{\min_{\Theta_0 \times \mathcal{P}} \Phi_A(d_0^*; \boldsymbol{\theta}, \mathbf{p})} \quad (3)$$

is 1 for any $g = 1, \dots, G$. If such a d_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 maximin efficiency of d_0^* . More precisely,

$$\min_{g \neq 0} \mathcal{R}_g \leq \min_{g \neq 0} \frac{\min_{\Theta_g \times \mathcal{P}} \Phi_A(d_0^*; \boldsymbol{\theta}, \boldsymbol{p})}{\min_{\Theta_0 \times \mathcal{P}} \Phi_A(d^*; \boldsymbol{\theta}, \boldsymbol{p})} \leq \min_{g \neq 0} \frac{\min_{\Theta_g \times \mathcal{P}} \Phi_A(d_0^*; \boldsymbol{\theta}, \boldsymbol{p})}{\min_{\Theta \times \mathcal{P}} \Phi_A(d^*; \boldsymbol{\theta}, \boldsymbol{p})} = \frac{\min_{\Theta \times \mathcal{P}} \Phi_A(d_0^*; \boldsymbol{\theta}, \boldsymbol{p})}{\min_{\Theta \times \mathcal{P}} \Phi_A(d^*; \boldsymbol{\theta}, \boldsymbol{p})}, \quad (4)$$

where d^* is a maximin design for $\Theta \times \mathcal{P}$. Note that the equality in (4) follows from Lemmas A.1 and A.2 in Appendix A. From (4), d_0^* is maximin efficient if the minimal \mathcal{R}_g is close to 1.

As demonstrated in the next section, our proposed strategy helps to obtain designs that are very efficient in terms of the maximin criterion. In addition, our strategy significantly reduces the computing time. To obtain maximin designs, we adopt the knowledge-based GA of Kao et al. (2009a), which is briefly described in Appendix B. A computer program implementing the adopted algorithm is available upon request from the first author.

3. CASE STUDIES AND A REAL EXAMPLE

3.1 Maximin efficient designs

We apply our proposed approach to obtain maximin efficient designs for $Q = 1, 2$, and 3. To facilitate comparisons with designs in the literature, the lengths of the designs are set to 255, 242, and 255, respectively, with an ISI of 4s and TR of 2s. Following Kao et al. (2009a) and Liu (2004), a second-order polynomial drift in the BOLD times series is assumed. The noise follows a stationary AR(1) process with a correlation coefficient $\rho = 0.3$. When searching for maximin designs, the minimal

Φ_A -value is evaluated over a fine grid of the parameter space. Specifically, the grid interval for \mathbf{p} is 0.1, and that for φ_i s (and thus $\boldsymbol{\theta}$) is 0.05π . Finer grid intervals of 0.05 and 0.01π for \mathbf{p} and $\boldsymbol{\theta}$, respectively, are used for evaluating the obtained designs and for design comparisons.

For $Q = 1$, we adopt the GA of Kao et al. (2009a) to search for a design d_0^* maximizing $\min_{\mathcal{P}} \Phi_A(d; 1, \mathbf{p})$. For $Q = 2$, and 3, we follow our proposed strategy to first identify a Θ_0 , and \mathcal{G} . To allow a small set Θ_0 , \mathcal{G} is selected to contain all the $Q!$ permutation matrices. With this \mathcal{G} , we set Θ_0 to $\{(\cos \varphi_1, \sin \varphi_1) \mid \varphi_1 \in [-\pi/4, \pi/4]\}$ for $Q = 2$. As for $Q = 3$, Θ_0 can be selected as $\{(\cos \varphi_1, \pm \sin \varphi_1 \cos \varphi_2, \pm \sin \varphi_1 \sin \varphi_2) \mid \varphi_1 \in [0, \arccos(1/\sqrt{3})], \varphi_2 \in [\kappa, \pi/4]\}$, where $\kappa = \arccos(\cos \varphi_1 / \sin \varphi_1)$ 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 (by using Lemma 2) that Lemma 3 still holds by replacing Θ_g with Θ_g^* . After defining Θ_0 , we then apply the GA to search for designs d_0^* maximizing $\min_{\Theta_0 \times \mathcal{P}} \Phi_A(d; \boldsymbol{\theta}, \mathbf{p})$, and check if they are maximin efficient for the entire parameter space using Lemma 3.

For $Q = 2$, the value of \mathcal{R}_1 in (3) is 99.32%. The obtained design d_0^* should thus be very efficient for the entire parameter space $\Theta \times \mathcal{P}$ in terms of the maximin criterion. This is verified by the fact that

$$\frac{\min_{\Theta \times \mathcal{P}} \Phi_A(d_0^*; \boldsymbol{\theta}, \mathbf{p})}{\min_{\Theta \times \mathcal{P}} \Phi_A(d^*; \boldsymbol{\theta}, \mathbf{p})} = 99.55\%,$$

where d^* is a maximin design for $\Theta \times \mathcal{P}$ obtained by the GA. We note that, with our method, d^* needs not be computed normally. It is obtained here only to demonstrate

that the efficiency of d_0^* is indeed as good as we claimed it to be based on (4). In addition, obtaining d^* requires a CPU time of about 4.42 hours, and computing d_0^* takes about 1.68 hours by using MATLAB Version 7.11 (R2010b) on a desktop computer with a 3.0 GHz Intel Pentium 4 quad-core processor.

A similar result is observed for $Q = 3$. The obtained design d_0^* for $\Theta_0 \times \mathcal{P}$ yields a minimal \mathcal{R}_g of 99.03%. Comparing with a maximin design d^* for $\Theta \times \mathcal{P}$, we have

$$\frac{\min_{\Theta \times \mathcal{P}} \Phi_A(d_0^*; \boldsymbol{\theta}, \mathbf{p})}{\min_{\Theta \times \mathcal{P}} \Phi_A(d^*; \boldsymbol{\theta}, \mathbf{p})} = 99.20\%.$$

For this case, the CPU time required for d^* is 109.71 hours. Our proposed strategy saved about 74% in CPU time spent and achieved a maximin efficient design d_0^* .

[Figure 1 about here]

The obtained maximin efficient designs are presented in Figure 1 (the last three rows). They do not seem to have a perceivable pattern. This is in contrast to current knowledge about ER-fMRI designs, which is built on linear models. Under linear models, block designs (e.g., first row of Figure 1) are usually recommended for detecting activation. Our results in the next subsection suggest that, when the HRF shape is uncertain, block designs can be significantly outperformed by our designs.

3.2 Design comparisons

We compare the obtained maximin efficient designs d_0^* with block designs, m -sequences, max- F_d , max- F_e , multi-objective (MO), 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 consists

of a fixed number of q 's. For the current settings, we consider block designs of the form $\{00001111 \cdots QQQQ\}$, corresponding to the 16s-off-16s-on pattern. Under linear models, these designs are known to yield high efficiencies in detecting activation (Henson 2007). An m -sequence can be generated from primitive polynomials for a Galois field (MacWilliams and Sloane 1977; Godfrey 1993; Buračas and Boynton 2002). These designs can be obtained from a computer program provided by Liu (2004) and are good for estimating the HRF. The max- F_d , max- F_e and MO designs are obtained by the GA of Kao et al. (2009a) with linear models. A max- F_d design maximizes the efficiency of detection, whereas a max- F_e design maximizes the HRF estimation efficiency. MO designs maximize the average of these two efficiencies. They attain good compromises between the two competing objectives of detection and estimation. We also generate 100 random designs. 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. (2009a) and Liu (2004).

[Table 1 about here]

Table 1 presents the $\min(\Phi_A)$ -values of the previously mentioned designs and the maximin efficient design d_0^* . The mean and standard deviation of $\min(\Phi_A)$ of the 100 random designs are also provided. As presented there, the block and max- F_d designs can be very inefficient in detecting brain activation under the nonlinear model that accounts for the HRF shape uncertainty. The m -sequences, max- F_e , MO (denoted by $\cdot 5F_d + \cdot 5F_e$ in Table 1), and random designs can achieve a better performance. This result is in contrast to that observed in previous studies with linear models (e.g., Kao et al. 2009a). Under linear models, block and max- F_d designs are commonly recom-

mended for detection. However, they typically have (near) zero estimation efficiencies, and do not allow the estimation of the HRF. We believe that such a deficiency makes these patterned designs bad choices when the HRF shape is uncertain. This is because, in contrast to linear models, the nonlinear model takes the uncertain HRF shape into account by accommodating both detection and estimation in a unified setting. While detection is the main focus, a reasonably precise estimate of the HRF parameters \mathbf{p} is intrinsically required.

On the other hand, with linear models, designs good for estimation usually maintain ‘moderate’ efficiencies for detection; e.g., the max- F_e design with $Q = 1$ achieves 70.67% efficiency for detection. Being able to handle both dimensions, these designs work better than the block and max- F_d designs for the nonlinear model. However, none of these widely considered designs perform nearly as well as the maximin efficient designs obtained by our approach.

Here, we demonstrate that our proposed approach facilitates the search for maximin ER-fMRI designs. For $Q = 2$ and 3, our empirical results suggest a possible direction for further decreasing the computational cost. Inspired by Imhof and Wong (2000), we select a subset $\Theta_D \times \mathcal{P}_D$ of $\Theta_0 \times \mathcal{P}$, that contains the 2^{Q-1} extreme points of Θ_0 and the four edges of \mathcal{P} . For $Q = 2$, we set $\Theta_D = \{(\cos \varphi_1, \sin \varphi_1) \mid \varphi_1 = -\pi/4, \pi/4\}$. For $Q = 3$, Θ_D is selected as $\{(1/\sqrt{3}, \pm 1/\sqrt{3}, \pm 1/\sqrt{3})\}$. A maximin design d_D^* is obtained for the even smaller parameter space $\Theta_D \times \mathcal{P}_D$. As presented in Table 1, the performance of d_D^* is very similar to (and slightly better than) that of d_0^* . In addition, calculating d_D^* only requires 3.1 and 13.34 minutes for $Q = 2$ and 3, respectively. While further investigations are needed, this empirically based approach seems promising.

3.3 An example

In addition to case studies, 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 2.5s ($ISI=2.5s$). The BOLD time series was acquired every 2.5s ($TR=2.5s$). 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 1.25s 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 a maximin efficient design. A simple modification is needed to accommodate the special requirement regarding that the study is conducted over two runs. Specifically, we replace the design matrix in model (1) by $diag(\mathbf{X}_{d,1}, \mathbf{X}_{d,1})$ 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 j th 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 1.25s 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 $\rho = 0.3$. The whitening matrix is thus of the form $(\mathbf{I}_2 \otimes \mathbf{V})$, where \mathbf{V} is a whitening matrix for each run; see also Kao, Mandal, and

Stufken (2009b).

[Figure 2 about here]

The achieved maximin design is presented in Figure 2. Similar to the designs for our case studies, the obtained design possesses both ‘blocky’ and ‘random’ components. It does not seem to have a perceivable pattern. For comparison, we also 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. With the current setting, this block design is closest to the commonly suggested 16s-off-16s-on pattern. 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 4.9s and 5.1s. Figure 2 presents the minimal Φ_A -value of each design. For the 100 randomly permuted designs, the mean and standard deviation of the minimal Φ_A -values are provided. Among the designs being compared, the design that we obtain is the most robust against HRF shape mis-specification.

4. CONCLUSION

In this study, we propose an approach for obtaining ER-fMRI designs for detecting brain activation by taking the uncertain HRF shape into account. We consider a nonlinear model that involves $\theta\mathbf{h}(\mathbf{p})$, 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 (Lindquist and Wager 2007; Handwerker et al. 2004; Miezin et al. 2000). While we focus on an $\mathbf{h}(\mathbf{p})$ having the same form as the popularly considered double-gamma function of SPM, it should be possible to extend our proposed design approach to $\mathbf{h}(\mathbf{p})$ s of other forms; e.g., the inverse logit function considered by Lindquist and Wager (2007).

Our proposed approach involves a large reduction in the parameter space and the adoption of the knowledge-based GA of Kao et al. (2009a) for searching for maximin efficient designs. The approach is demonstrated to be powerful via case studies and a real example. We also report an empirical result that can possibly lead to an even more efficient way in obtaining maximin efficient designs. Specifically, we focus on a judiciously selected subset of $\Theta_0 \times \mathcal{P}$ for $Q = 2$ and 3 , and achieve designs d_D^* that are rather robust (Subsection 3.2). This latter approach seems to offer a promising avenue for further investigation.

References

- Berger, M. P. F. and Wong, W. K. (2009), *An Introduction to Optimal Designs for Social and Biomedical Research*, Chichester, West Sussex, U.K.: Wiley.
- 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,” *Journal of Statistical Planning and Inference*, 137, 3291–3302.

- Box, G. E. P. and Lucas, H. L. (1959), “Design of Experiments in Non-Linear Situations,” *Biometrika*, 46, 77–90.
- Buračas, G. T. and Boynton, G. M. (2002), “Efficient Design of Event-Related fMRI Experiments Using M-Sequences,” *NeuroImage*, 16, 801–813.
- Chernoff, H. (1953), “Locally Optimal Designs for Estimating Parameters,” *Annals of Mathematical Statistics*, 24, 586–602.
- Culham, J. C. (2006), “Functional Neuroimaging: Experimental Design and Analysis,” in *Handbook of Functional Neuroimaging of Cognition*, eds. Cabeza, R. and Kingstone, A., Cambridge, Massachusetts: MIT Press, 2nd ed., pp. 53–82.
- D’Esposito, M., Zarahn, E., and Aguirre, G. K. (1999), “Event-Related Functional MRI: Implications for Cognitive Psychology,” *Psychological Bulletin*, 125, 155–164.
- Fedorov, V. V. and Hackl, P. (1997), *Model-Oriented Design of Experiments*, New York: Springer.
- Friston, K. J., Holmes, A. P., Poline, J. B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J., and Turner, R. (1995), “Analysis of fMRI Time-Series Revisited,” *NeuroImage*, 2, 45–53.
- Godfrey, K. (1993), *Perturbation Signals for System Identification*, New York: Prentice Hall.
- 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*, New York: Springer.
- Henson, R. N. A. (2007), "Efficient Experimental Design for fMRI," in *Statistical Parametric Mapping: The Analysis of Functional Brain Images*, eds. Friston, K. J., Ashburner, J. T., Kiebel, S. J., Nichols, T. E., and D., P. W., London: Academic, 1st ed., pp. 193–210.
- Imhof, L. and Wong, W. K. (2000), "A Graphical Method for Finding Maximin Efficiency Designs," *Biometrics*, 56, 113–117.
- 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), "Multi-Objective Optimal Experimental Designs for ER-fMRI Using Matlab," *Journal of Statistical Software*, 30, 1–13.
- Kao, M.-H., Mandal, A., Lazar, N., and Stufken, J. (2009a), "Multi-Objective Optimal Experimental Designs for Event-Related fMRI Studies," *NeuroImage*, 44, 849–856.
- Kao, M.-H., Mandal, A., and Stufken, J. (2008), "Optimal Design for Event-Related Functional Magnetic Resonance Imaging Considering Both Individual Stimulus Effects and Pairwise Contrasts," *Statistics and Applications*, 6, 225–241.
- (2009b), "Efficient Designs for Event-Related Functional Magnetic Resonance Imaging with Multiple Scanning Sessions," *Communications in Statistics-Theory and Methods*, 38, 3170–3182.

- Lazar, N. A. (2008), *The Statistical Analysis of Functional MRI Data*, Statistics for Biology and Health, New York: Springer.
- Lindquist, M. A. (2008), “The Statistical Analysis of fMRI Data,” *Statistical Science*, 23, 439–464.
- Lindquist, M. A. and Wager, T. D. (2007), “Validity and Power in Hemodynamic Response Modeling: A Comparison Study and a New Approach,” *Human Brain Mapping*, 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.
- Liu, T. T., Frank, L. R., Wong, E. C., and Buxton, R. B. (2001), “Detection Power, Estimation Efficiency, and Predictability in Event-Related fMRI,” *NeuroImage*, 13, 759–773.
- Loh, J. M., Lindquist, M. A., and Wager, T. D. (2008), “Residual Analysis for Detecting Mis-Modeling in fMRI,” *Statistica Sinica*, 18, 1421–1448.
- MacWilliams, F. J. and Sloane, N. J. A. (1977), *The Theory of Error Correcting Codes*, Amsterdam ; New York New York: North-Holland Pub. Co. ; sole distributors for the U.S.A. and Canada, Elsevier/North-Holland.
- 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.
- Pukelsheim, F. (1993), *Optimal Design of Experiments*, Wiley Series in Probability and Mathematical Statistics. Probability and Mathematical Statistics, New York: Wiley.
- Rosen, B. R., Buckner, R. L., and Dale, A. M. (1998), “Event-Related Functional MRI: Past, Present, and Future,” *PNAS*, 95, 773–780.
- Silvey, S. D. (1980), *Optimal Design: An Introduction to the Theory for Parameter Estimation*, Monographs on Applied Probability and Statistics., London ; New York: Chapman and Hall.
- 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 Friston, K. J. (1995), “Analysis of fMRI Time-Series Revisited—Again,” *NeuroImage*, 2, 173–181.
- Worsley, K. J. and Taylor, J. E. (2006), “Detecting fMRI Activation Allowing for Unknown Latency of the Hemodynamic Response,” *Neuroimage*, 29, 649–654.

APPENDIX A: A PROOF OF LEMMA 3

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

Lemma A.1. 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 $(1, 2, \dots, Q)'$, 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 $(\boldsymbol{\theta}, \mathbf{p}) \in \Theta_0 \times \mathcal{P}$.

Lemma A.2. 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_0^* satisfying the conditions of Lemma 3, $\min_{\Theta_g \times \mathcal{P}} \Phi_A(k_{G_g}(d); \boldsymbol{\theta}, \mathbf{p}) = \min_{\Theta_0 \times \mathcal{P}} \Phi_A(d_0^*; \boldsymbol{\theta}, \mathbf{p}) = \min_{\Theta_g \times \mathcal{P}} \Phi_A(d_0^*; \boldsymbol{\theta}, \mathbf{p})$ for any g . Therefore, d_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 $(\bigcup_{g=0}^G \Theta_g) \times \mathcal{P} = \Theta \times \mathcal{P}$.

APPENDIX B: A SEARCH ALGORITHM

The search algorithm that we consider is provided below. The objective function is $\min(\Phi_A)$, where the minimum is evaluated over a grid on a specified parameter space.

Step 1. Generate N (an even number) initial designs, including block, m -sequence(-based), and random designs, and mixed designs with partial block design and partial m -sequence(-based) or random design. Obtain the fitness (value of the objectives function) of each initial design.

Step 2. With probability proportional to fitness, select with replacement $N/2$ pairs of designs to generate $N/2$ pairs of offsprings via crossover and mutation. Specifically, the crossover operator exchanges the corresponding subsequences of the paired designs. The mutation operator perturbs a randomly selected portion α_m of elements of all offspring designs. Obtain the fitness of the resulting designs.

Step 3. Add to the population another $\alpha_i N$ (an integer) immigrants drawn from random and block designs, and their combinations. Obtain their fitness.

Step 4. Combine the N designs in the current generation, the N offspring designs and $\alpha_i N$ immigrants. According to their fitness, keep the best N designs to form the next generation, and discard the others.

Step 5. Repeat steps 2 through 4 until a stopping rule is met; e.g., no significant improvement. Keep track of the best design over generations.

When implementing the algorithm, we follow Kao et al. (2009a) to set $N = 20$, $\alpha_m = 1\%$ and $\alpha_i = 20\%$. The search is terminated if there is no significant improvement in the objective function. Specifically, the improvement is checked every 100 generations and is compared to that of the first 100 generations. We stop the search if the relative improvement is no more than 10^{-7} . In our experience, this stopping rule works well. See Kao (2009) for details.

Table 1: The $\min(\Phi_A)$ -values of widely used designs, and designs obtained by our approaches, including an alternative approach. The column labeled ‘random’ shows the the mean and standard deviation of $\min(\Phi_A)$ values for 100 random designs.

Q	block	widely used design				our approach		
		$\max-F_d$	m -seq.	$\max-F_e$	$.5F_d + .5F_e$	random	d_0^*	d_D^*
1	42.09	42.69	64.83	65.32	62.88	64.39 (1.58)	74.39	—
2	10.96	14.56	21.20	22.49	22.35	21.08 (0.70)	25.22	25.28
3	4.58	7.30	11.54	12.64	12.64	11.30 (0.44)	14.01	14.09

Figure 1. A block design for $Q = 1$ (formed by repetitions of $\{00001111\}$), and the maximin efficient designs, d_0^* , for $Q = 1, 2$ and 3 . Different shades of gray indicate different types of events (stimuli or control).

Figure 2. Upper: the obtained maximin efficient design. Lower: $\min(\Phi_A)$ -values of a block design, m -sequence-based design and a maximin efficient design, and the mean and standard deviation (error bar) of $\min(\Phi_A)$ of the 100 randomly permuted designs.

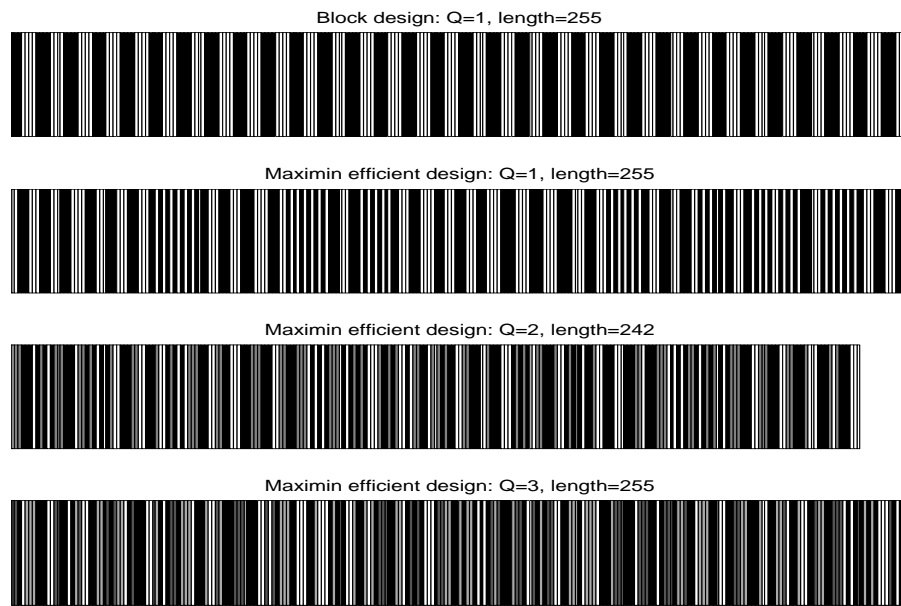


Figure 1: A block design for $Q = 1$ (formed by repetitions of $\{00001111\}$), and the maximin efficient designs, d_0^* , for $Q = 1, 2$ and 3 . Different shades of gray indicate different types of events (stimuli or control).

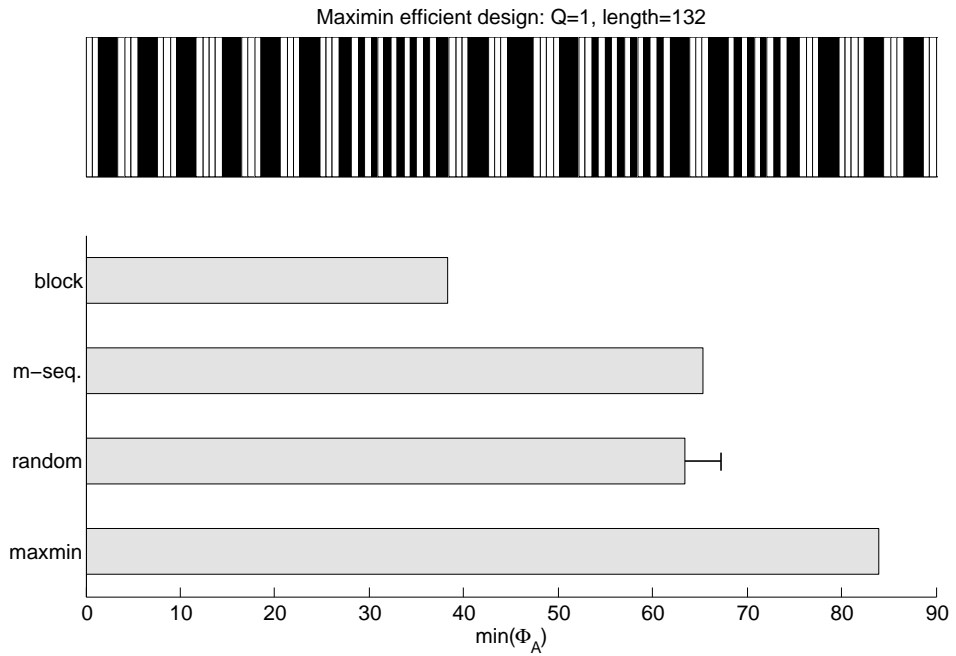


Figure 2: Upper: the obtained maximin efficient design. Lower: $\min(\Phi_A)$ -values of a block design, m -sequence-based design and a maximin efficient design, and the mean and standard deviation (error bar) of $\min(\Phi_A)$ of the 100 randomly permuted designs.