Statistics and Applications
Volume 6, Nos.1 & 2, 2008 (New Series), pp. 235-256

Optimal design for event-related functional magnetic resonance imaging considering both individual stimulus effects and pairwise contrasts

Ming-Hung Kao, Abhyuday Mandal and John Stufken
Department of Statistics, The University of Georgia

Abstract

In this article, we study multi-objective optimal designs for event-related functional magnetic resonance imaging (ER-fMRI). The objectives considered are two common statistical goals, namely estimation and detection. We focus on the case when both individual stimulus effects and pairwise contrasts are of interest. Using a genetic algorithm, we search for optimal designs under different multi-objective design criteria. We study the designs that we obtain under different weights for individual stimulus effects and pairwise contrasts. We also study the performance of popular ER-fMRI designs currently in use by fMRI researchers.

Key words: Compound design criterion; Design efficiency; Individual stimulus effects; Pairwise contrasts; Discretization interval.

1 Introduction

ER-fMRI is one of the leading technologies for studying human brain activity in response to brief mental stimuli or tasks. Unlike the traditional fMRI, where long-period stimuli are used, ER-fMRI takes advantage of an ultra-fast MR scanner to allow the study of an effect due to a single, brief stimulus. ER-fMRI is a popular technique for brain mapping in both medical practice and scientific research and is arguably the most important advance in neuroscience (Rosen et al., 1998; Josephs and Henson, 1999; Culham, 2006).

A design for an ER-fMRI study consists of a sequence of stimuli of one or more types interlaced with a control condition (rest or fixation). Finding an optimal sequence of the stimuli best suited to the

researcher's need can be arduous. There are multiple reasons for this. First, an ER-fMRI design is a long sequence of finite numbers, typically consisting of hundreds of elements. The design space containing all possible arrangements is thus enormous and irregular (Buračas and Boynton, 2002; Liu, 2004). Second, the nature of ER-fMRI requires consideration of multiple objectives in a study. These objectives involve not only statistical goals but also psychological constraints. In addition, customized requirements can arise to make the problem even more complicated. To overcome these difficulties, an efficient search algorithm along with a well-defined multi-objective criterion (MO-criterion) is called for.

An efficient approach to search for multi-objective optimal designs for ER-fMRI is proposed by Kao et al. (2009). Their search algorithm is a genetic algorithm (GA) and their multi-objective design criteria are convex combinations of criteria for single objectives. Two popular statistical objectives in ER-fMRI are estimation and detection. Estimation refers to the estimation of the hemodynamic response function (HRF), a function of time describing the effect on the brain of a single, brief stimulus. Detection aims at investigating whether a region is activated by each stimulus type. This is accomplished by separately studying the amplitudes (or the peaks) of the HRFs evoked by different stimulus types. For both statistical objectives, a researcher may be interested in studying individual stimulus effects and pairwise contrasts of stimulus effects (Amaro and Barker, 2006; Liu and Frank, 2004; Donaldson and Buckner, 2002). Kao et al. (2009) consider optimal designs for each of these two interests. However, following Liu and Frank (2004) and Liu (2004), in this paper we study designs that are efficient if both individual effects and pairwise contrasts are of interest. In contrast to earlier work, our approach allows user-specified weights for individual stimulus effects and pairwise contrasts.

The search algorithm proposed by Kao et al. (2009) is adopted to search for optimal designs, but we define a different family of MO-criteria to meet the goal of this study. Each MO-criterion is a weighted sum of the criteria for estimation and detection, and each of the latter criteria is defined based on a convex combination of the criteria for individual stimulus effects and pairwise contrasts. We study efficient designs that we obtain by using different weighting schemes, and compare them to designs currently in use by fMRI researchers.

In the following section, we briefly introduce background informa-

tion regarding ER-fMRI designs. Section 3 describes our methodology including the statistical model, the design criteria and the search algorithm. Numerical results are presented in Section 4 and conclusions are provided in Section 5.

2 ER-fMRI designs

An ER-fMRI design is an alignment of events, including stimuli of different types and the control. For convenience, the symbols $0, 1, ..., Q$ are used to represent the events, where Q is the total number of stimulus types. The control is indicated by 0 and a type-i stimulus is denoted by i, $i = 1, ..., Q$. A design, denoted by ξ, looks like $\xi = \{101201210...1\}$. While being presented to the experimental subject, each event lasts for a short period of time relative to the inter-stimulus interval (ISI), the fixed time interval between event onsets. We note that 0s in a sequence are "pseudo-events". They may be thought of as periods of rest for the subject, even though the subject may still experience effects from previous events.

While the design sequence is presented, the MR scanner scans the subject's brain every few seconds; the time duration between two scans is referred to as TR or time to repetition. The blood oxygenation level dependent (BOLD) signals at each brain voxel (a small region of the brain) is collected every TR seconds to form a voxel-wise fMRI time series. A design issue for ER-fMRI is to allocate the stimuli so that statistical inference (related to estimation or detection) on these time series is most efficient, in some sense.

Depending on the study objectives, several well-known ER-fMRI designs currently in use by researchers are block designs, m-sequence-based designs, mixed designs, permuted block designs, and clustered m-sequences (Liu, 2004). In ER-fMRI, a block design is a sequence where stimuli of the same type are clustered into blocks. For example, a two-stimulus-type block design with a block size of four can consist of repetitions of $\{111122220000\}$. Repetitions of $\{1111000022220000\}$ and $\{11112222\}$ are other possible patterns. Block designs are good for detection because, at a region that is activated by a particular stimulus type, the lingering effects evoked by stimuli of that type will accumulate to create strong signals under block designs. The difference in the signal intensity between activation and non-activation increases, and this helps in identifying activation. Agreeing with this

intuition, block designs yield high design efficiencies when the detection problem is the only concern.

The m-sequence-based designs are m-sequences (Barker, 2004; Godfrey, 1993) and designs constructed from m-sequences. These sequences can be constructed from Galois fields or Reed-Muller codes (cf. MacWilliams and Sloane, 1977, Ch. 14), and look rather random with no clear pattern. They only exist if $Q + 1$ is a prime or prime power. The use of these designs for estimating the HRF is first proposed by Buračas and Boynton (2002). Liu and Frank (2004) and Liu (2004) also study these designs. The m-sequence-based designs have high efficiencies for estimation.

Mixed designs, permuted block designs, and clustered m-sequences are studied by Liu and Frank (2004) and Liu (2004) for the case when both estimation and detection are of interest. It is shown that there are designs in these classes that offer advantageous trade-offs between the two competing statistical objectives. A mixed design is formed by concatenating a fraction of a block design with a fraction of an m-sequence (or a random design). By changing the length of the "blocky" part, and hence that of the "random" part, the resulting designs can move toward having high efficiencies for estimation or high efficiencies for detection. Permuted block designs can be generated by, repeatedly, exchanging positions of two randomly chosen events in a block design. The efficiency for estimation is gradually increased, at the expense of the ability for detection. Clustered m-sequences are created by permuting events in an m-sequence so that the resulting design becomes more "blocky". The design gradually moves toward having a higher efficiency for detection.

3 Methodology

3.1 Statistical models

In this section, we specify the underlying model for the two primary statistical objectives, namely estimation and detection. As in Wager and Nichols (2003) and Liu and Frank (2004), two popular linear models are considered (see also Friston et al., 1995; Worsley and

Friston, 1995; Dale, 1999):

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{h} + \boldsymbol{S}\boldsymbol{\gamma} + \boldsymbol{e}, \text{ and} \tag{1}$$

$$\boldsymbol{Y} = \boldsymbol{Z}\boldsymbol{\theta} + \boldsymbol{S}\boldsymbol{\gamma} + \boldsymbol{\eta}, \tag{2}$$

where $\boldsymbol{Y}$ is the voxel-wise fMRI time series, $\boldsymbol{h} = (\boldsymbol{h}_1', ..., \boldsymbol{h}_Q')'$ is the parameter vector for the HRFs of the Q stimulus types, $\boldsymbol{X} = [\boldsymbol{X}_1 \cdots \boldsymbol{X}_Q]$ is the design matrix, $\boldsymbol{\theta} = (\theta_1, ..., \theta_Q)'$ represents the response amplitudes, $\boldsymbol{Z} = \boldsymbol{X}\boldsymbol{h}_0$ is the convolution of stimuli with an assumed basis, $\boldsymbol{h}_0$, for the HRF, $\boldsymbol{S}\boldsymbol{\gamma}$ is a nuisance term describing the trend or drift of $\boldsymbol{Y}$, and $\boldsymbol{e}$ and $\boldsymbol{\eta}$ are noise. Following Wager and Nichols (2003), we assume a known whitening matrix, $\boldsymbol{V}$, such that $\boldsymbol{V}\boldsymbol{e}$ and $\boldsymbol{V}\boldsymbol{\eta}$ are white noise. Model (1) is typically used for estimating the HRF and model (2) for detecting activation.

Our model formulation, explained in more detail in the remainder of this subsection, follows Kao et al. (2009). The major advantage of this model formulation lies in the use of the discretization interval (Dale, 1999) for parameterizing the HRF. The discretization enables the use of a finite set of interpretable parameters to capture the fluctuation of the continuous HRF over time. The length of the discretization interval, denoted by ΔT, is set to the greatest value dividing both the ISI and TR. The HRF parameters, captured in the vector $\boldsymbol{h}$, then represent the heights of the HRF for each stimulus after every ΔT seconds following the stimulus onset. This parametrization is explained in the following example.

Example 3.1. *In Figure 1, we consider one stimulus type ($Q = 1$). The time interval between two consecutive events is 2s ($ISI = 2s$), and that between two successive scans is 3s ($TR = 3s$). An illustrative design is $\xi = \{110100...\}$ with three stimuli taking place at 2s, 4s and 8s, respectively. These three stimuli are presented by vertical bars in the figure, and are followed by curves that represent the evoked HRF. The heights of these three overlapping HRFs accumulate to form the dash-dot curve. This curve represents the noise-free and trend-free BOLD responses induced by the three stimuli; it corresponds to $\boldsymbol{X}\boldsymbol{h}$ in model (1).*

The four vertical lines correspond to the first four MR scans at which the BOLD signal is observed. The heights of the HRFs, or equivalently the effects of the stimuli, that contribute to the observed signal are indicated by the dots on the lines. These heights are dif-

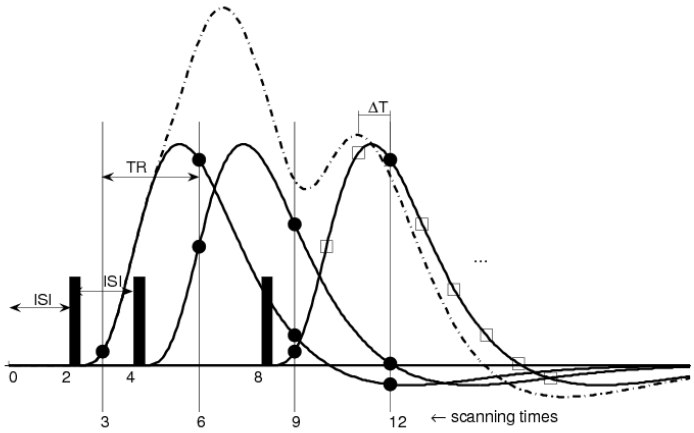


Figure 1: HRF parametrization with $Q = 1$, ISI = 2s, TR = 3s and $\Delta T = 1$s.

ferent. Therefore, we need different parameters to represent them as well as any other heights that could possibly contribute to a response.

Under the combination of ISI and TR in this example, the heights that could contribute to an observation occur every second on the HRF curve. They are shown as dots and squares on the third curve in Figure 1. The reason for the 1 second intervals is that a scan can occur 1s, 2s, 3s, ... after the onset of an event. In general, this time difference must be a multiple of the greatest number dividing both ISI and TR. Setting ΔT to this greatest common divisor, our HRF parameters then describe the discretized HRF, $h((j-1)\Delta T)$, $j = 1, 2, ...$ Here, $h(t)$ is the HRF at time t following the stimulus onset; $t = 0$ corresponds to the stimulus onset. All heights that could possibly make a contribution to an observation are represented with this parametrization. In addition, irrelevant heights that will never contribute to a response are left out.

The parameter vector $\boldsymbol{h}_i = (h_{1i}, ..., h_{ki})'$ represents the HRF, $h_i(t)$, for the type-i stimulus. With ΔT as defined in Example 3.1, we use h_{ji} to denote the height $h_i((j-1)\Delta T)$; $j = 1, ..., k$. Here, the length of $\boldsymbol{h}_i$ is $k = 1 + \lfloor K/\Delta T \rfloor$, where $\lfloor a \rfloor$ is the greatest integer less than or equal to a and K is the duration of the HRF, counting from the stimulus onset to the complete return to baseline.

The matrix $\boldsymbol{X}$ in model (1) is determined by both the design sequence and the HRF parametrization. The matrix corresponding to Example 3.1 is provided below as an illustration. If the duration of the HRF is 32 seconds then (since $Q = 1$) there are 33 columns. Each column is linked to an h_{j1} and is labeled by $t_j = (j-1)\Delta T$ ($\Delta T =$ 1s). Rows are labeled by scanning times, which are multiples of TR (= 3s).

$$\boldsymbol{X} = \boldsymbol{X}_1 = \begin{array}{c} \\ \\ 3s \rightarrow \\ 6s \rightarrow \\ 9s \rightarrow \\ 12s \rightarrow \\ \\ \end{array} \begin{array}{cccccccccccc} 0s & 1s & 2s & 3s & 4s & 5s & 6s & 7s & 8s & 9s & 10s & 11s \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \begin{array}{c} \ldots \\ \\ \ldots \\ \ldots \\ \ldots \\ \ldots \\ \\ \end{array} .$$

While using model (2), the same basis function, $\boldsymbol{h}_0$, for the HRFs is assumed for detection for each stimulus type. Throughout this article, we consider $\boldsymbol{h}_0$ to be the canonical HRF of the software SPM2 (http://www.fil.ion.ucl.ac.uk/spm/software/spm2/), which is a combination of two gamma distributions and is scaled to have a maximal value of one. In model (2), the matrix $\boldsymbol{Z} = \boldsymbol{X}\boldsymbol{h}_0$ represents the convolution of the stimuli with $\boldsymbol{h}_0$; see, e.g., Josephs et al. (1997) for details. The parameter $\boldsymbol{\theta}$ represents the maximal heights for the HRFs evoked by the stimuli. The larger θ_i the more the region is activated by the type-i stimulus.

As stated, the same basis function $\boldsymbol{h}_0$ is assumed for all the Q stimulus types. Only the amplitudes are allowed to vary. However, incorporating different basis functions for different stimuli is also possible in our approach. One can simply take $\boldsymbol{Z} = [\boldsymbol{X}_1\boldsymbol{h}_0^{(1)} \ldots \boldsymbol{X}_Q\boldsymbol{h}_0^{(Q)}]$, but this setting is beyond the scope of the current work.

3.2 Design criteria

For models (1) and (2), we consider the parametric functions $\boldsymbol{C}_x\boldsymbol{h}$ and $\boldsymbol{C}_z\boldsymbol{\theta}$, respectively, where

$$\boldsymbol{C}_x = \begin{bmatrix} (1-\delta_x)\boldsymbol{I}_{Qk} \\ \delta_x\boldsymbol{D}_x \end{bmatrix};\ \boldsymbol{C}_z = \begin{bmatrix} (1-\delta_z)\boldsymbol{I}_Q \\ \delta_z\boldsymbol{D}_z \end{bmatrix}.$$

Here, $\boldsymbol{I}_a$ is the a-by-a identity matrix, $\boldsymbol{D}_z$ (with elements of 0 and ± 1) is the $(Q(Q-1)/2)$-by-Q matrix in which the rows correspond

to the $Q(Q-1)/2$ pairwise contrasts, $\boldsymbol{D}_x = \boldsymbol{D}_z \otimes \boldsymbol{I}_k$ with $\otimes$ being the Kronecker product, and $\delta_x,\ \delta_z \in [0,1]$. When $\delta_x = \delta_z = 0$, the parametric functions correspond to individual stimulus effects only. When δ_x and δ_z increase, more weight is assigned to pairwise contrasts. With $\delta_x = \delta_z = 1$ we have the case when only pairwise contrasts are of interest. We note that Kao et al. (2009) study the two extreme cases of 0 and 1 separately. Liu and Frank (2004) and Liu (2004) investigated the case when $\delta_x = \delta_z = 1/2$. In this study, we allow the values of δ_x and δ_z to vary between 0 and 1. For simplicity, we consider $\delta_x = \delta_z$. However, the equality between the two coefficients is not required in our approach.

With these parametric functions of interest, we define the design criteria for estimation and detection. For estimation, the design criterion is

$$r_{cx}\left\{ trace\Big\{ \boldsymbol{C}_x \big[\boldsymbol{X}'\boldsymbol{V}'(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{VS}})\boldsymbol{V}\boldsymbol{X} \big]^{-1} \boldsymbol{C}_x' \Big\} \right\}^{-1}$$
$$= r_{cx}\left\{ (1-\delta_x)^2 trace[\boldsymbol{M}_x^{-1}] + \delta_x^2 trace[\boldsymbol{D}_x \boldsymbol{M}_x^{-1} \boldsymbol{D}_x'] \right\}^{-1},$$

where ξ is the design, $\boldsymbol{M}_x = \boldsymbol{M}_x(\xi) = \boldsymbol{X}'\boldsymbol{V}'(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{VS}})\boldsymbol{V}\boldsymbol{X}$ is the information matrix for $\boldsymbol{h}$, $P_{\boldsymbol{A}} = \boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'$ is the orthogonal projection onto the vector space spanned by the columns of $\boldsymbol{A}$, and r_{cx} is Q for $\delta_x = 0$, $Q(Q-1)/2 + Q$ for $\delta_x \in (0,1)$ and $Q(Q-1)/2$ for $\delta_x = 1$.

Similarly, we define the design criterion for detection as

$$r_{cz}\left\{ trace\Big\{ \boldsymbol{C}_z \big[\boldsymbol{Z}'\boldsymbol{V}'(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{VS}})\boldsymbol{V}\boldsymbol{Z} \big]^{-1} \boldsymbol{C}_z' \Big\} \right\}^{-1}$$
$$= r_{cz}\left\{ (1-\delta_z)^2 trace[\boldsymbol{M}_z^{-1}] + \delta_z^2 trace[\boldsymbol{D}_z \boldsymbol{M}_z^{-1} \boldsymbol{D}_z'] \right\}^{-1},$$

where $\boldsymbol{M}_z = \boldsymbol{M}_z(\xi) = \boldsymbol{Z}'\boldsymbol{V}'(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{VS}})\boldsymbol{V}\boldsymbol{Z}$, and r_{cz} is kQ for $\delta_z = 0$, $kQ(Q-1)/2 + kQ$ for $\delta_z \in (0,1)$ and $kQ(Q-1)/2$ for $\delta_z = 1$. The weights assigned to individual stimulus effects may be thought of as $\lambda_x = (1-\delta_x)^2/[(1-\delta_x)^2 + \delta_x^2]$ and $\lambda_z = ((1-\delta_z)^2/[(1-\delta_z)^2 + \delta_z^2]$, with $1-\lambda_x$ and $1-\lambda_z$ being the weights for pairwise contrasts. To explicitly present the dependence of the design criteria on the design and the

weights, we use $F_e(\xi;\lambda_x)$ to indicate the criterion for estimation and $F_d(\xi;\lambda_z)$ is for detection.

In general, we define the family of MO-criteria as

$$\left\{F^* = wF_e^*(\xi;\lambda_x)+(1-w)F_d^*(\xi;\lambda_z) : w \in [0,1], \lambda_x, \lambda_z \in [0,1]\right\}, \quad (3)$$

where (suppressing ξ, λ_x and λ_z)

$$F_j^* = \frac{F_j - \min(F_j)}{\max(F_j) - \min(F_j)}, j = e, d.$$

We use the extreme values of F_d and F_e to normalize these two criteria before combining them. This is done to ensure comparability between the two criteria; see also, Imhof and Wong (2000). The values of w, λ_x and λ_z can be assigned based on the researcher's discretion. After specifying these values, an MO-criterion is well-defined and the search algorithm of Kao et al. (2009) can be applied to search for optimal designs. The search algorithm is based on a genetic algorithm and is briefly introduced in the next section.

3.3 Search algorithm

GAs (Holland, 1975; 1992) are popular for solving optimization problems, in which good solutions (parents) are used to generate better ones (offsprings). The GA proposed by Kao et al. (2009) takes advantage of well-known results about good fMRI designs so that the search over the huge design space can be carried out more efficiently. The outline of the algorithm is as follows:

Step 1. (Initial designs) Generate G initial designs consisting of random designs, an m-sequence-based design, a block design and their combinations. Use the objective function to evaluate the fitness of each initial design.

Step 2. (Crossover) With probability proportional to fitness, draw with replacement $G/2$ pairs of designs to crossover — select a random cut-point and exchange the corresponding fractions of the paired designs.

Step 3. (Mutation) Randomly select $q\%$ of the events from the G offspring designs. Replace these events by randomly generated ones.

Step 4. (Immigration) Add to the population another I designs drawn from random designs, block designs and their combinations.

Step 5. (Fitness) Obtain the fitness scores of the offsprings and immigrants.

Step 6. (Natural selection) Keep the best G designs according to their fitness scores to form the parents of the next generation. Discard the others.

Step 7. (Stop) Repeat steps 2 through 6 until a stopping rule is met (e.g., after M generations). Keep track of the best design over generations.

Details about this algorithm can be found in Kao et al. (2009). It is worth noticing that the objective function used in Step 1 and Step 5 can be taken as the design criterion for estimation, or that for detection, or an MO-criterion. To use our MO-criterion, the extreme values of F_e and F_d are required. Theoretical values of $\max(F_e)$ and $\max(F_d)$ are generally not available. They are thus approximated numerically by the GA using the non-standardized function F_e (or F_d) as the objective function. The values of $\min(F_e)$ and $\min(F_d)$ are set to zero, corresponding to designs for which the parametric functions of interest are non-estimable.

When searching for optimal designs, we follow Kao et al. (2009) to use $G = 20$, $q = 1$, $I = 4$ and $M = 10,000$. A larger M does not seem to lead to significantly better designs.

4 Numerical results

We study optimal ER-fMRI designs through a series of numerical simulations. The focus is on investigating the impact of $\lambda = \lambda_x = \lambda_z$, for which we will consider the values of 1, 3/4, 2/3, 1/2, 1/3, 1/4, 1/8, 1/16 and 0. The case $\lambda = 1/2$ is also studied by Liu (2004). The number of stimulus types (Q) considered ranges from 2 to 4.

The length of the design is $L = 242$ for $Q = 2$, $L = 255$ for $Q = 3$ and $L = 624$ for $Q = 4$. Under these combinations of Q and L, an m-sequence exists. We will consider two cases for the model. Case I assumes that errors are white noise, i.e., $\boldsymbol{V} = \boldsymbol{I}$, and that $\boldsymbol{S}$ is a vector of ones. For Case II, we assume a stationary AR(1) error process with a correlation coefficient of 0.3, while $\boldsymbol{S\gamma}$ is taken to be a second-order Legendre polynomial drift. The resulting models for this case are closer to those that are used by fMRI researchers. The ISI and TR are both set to 2 seconds, so that ΔT is also 2 seconds. Under these conditions, we search for optimal designs using our genetic algorithm approach with specified optimality criteria.

4.1 Designs for detection

Here, we study the case when only detection is of interest; i.e., $w = 0$ in (3). As described in Section 2, block designs are expected to be optimal and our approach yields designs with a block structure. Figures 2 and 3 present these designs for $Q = 2$ and $Q = 3$, respectively. Different shades of grey indicate different stimulus types and white represents the control. The number above each block is the number of events contained in that block. For example, the top-left design in Figure 2 starts with seven controls (pseudo-events), followed by eight stimuli of one type and ends with a stimulus of the other type. We do not present the designs for $Q = 4$ since they provide little additional insight.

From Figures 2 and 3, the occurrence of the control condition decreases with λ. The range of the frequencies of the stimulus types in our designs is presented in Table 1. For both cases, the results agree with the approximated optimal stimulus frequencies that Liu and Frank (2004) derived for Case I.

In Table 2, we present relative efficiencies for the designs that we obtain. For example, for $Q = 3$ under Case I, the entry 93.9 for the row labeled $F_d^*(\xi; \lambda = 1)$ and the column labeled $\xi_{1/2}^d$ indicates that the optimal design for $\lambda = 1/2$, $Q = 3$ and Case I, say $\xi_{1/2}^d$, has an efficiency of 93.9% for detection if $\lambda = 1$. If we denote the optimal design for $\lambda = 1$ by ξ_1^d, this means that

$$\frac{F_d(\xi_{1/2}^d; 1)}{F_d(\xi_1^d; 1)} = 93.9\%$$

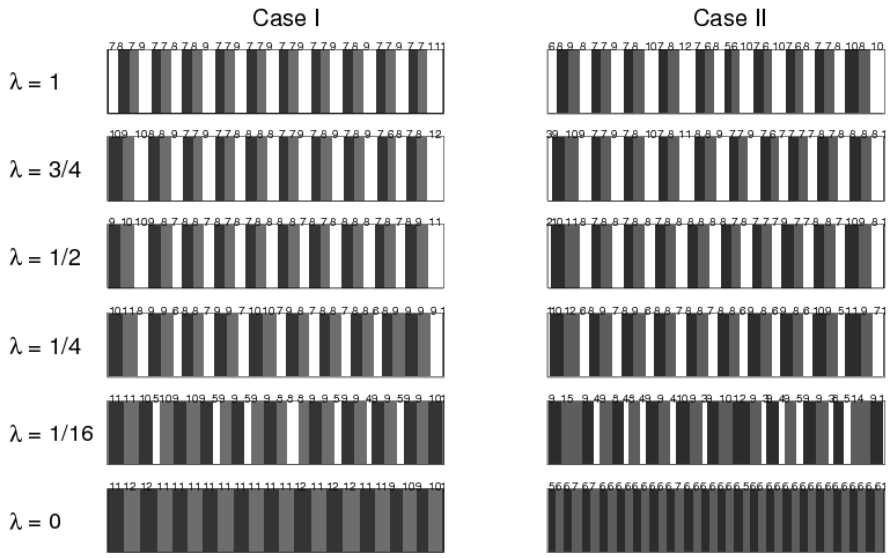


Figure 2: The designs obtained by our approach for detection with various λ when $Q = 2$. The rows corresponds to $\lambda = 1$, 3/4, 1/2, 1/4, 1/16, and 0, respectively. The first column is for Case I and the second column for Case II.

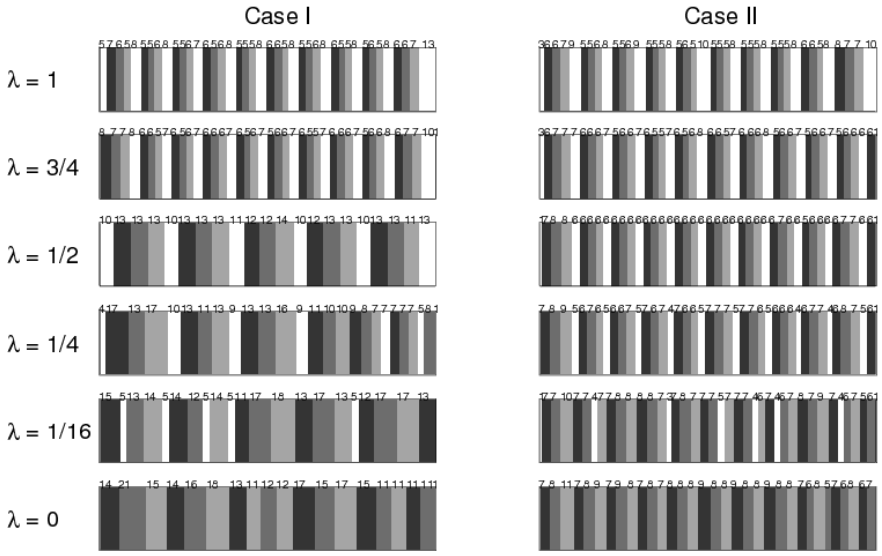


Figure 3: The designs obtained by our approach for detection when $Q = 3$. The rows corresponds to $\lambda = 1$, 3/4, 1/2, 1/4, 1/16, and 0, respectively. The first column is for Case I and the second column for Case II.

Table 1: Stimulus frequencies (range over different stimulus types) of our designs for detection[1]

Q	λ 1	3/4	2/3	1/2	1/3	1/4	1/8	1/16	0
				Case I					
2	0.29-0.30 (0.30)	0.31 (0.31)	0.32 (0.32)	0.33 (0.33)	0.35-0.36 (0.36)	0.36-0.37 (0.37)	0.39-0.40 (0.40)	0.42-0.43 (0.42)	0.50 (0.50)
3	0.21-0.22 (0.21)	0.24 (0.23)	0.24 (0.24)	0.25 (0.25)	0.26-0.27 (0.26)	0.27 (0.27)	0.29 (0.29)	0.30-0.31 (0.30)	0.33 (0.33)
4	0.17 (0.17)	0.18-0.19 (0.18)	0.19 (0.19)	0.20 (0.20)	0.21 (0.21)	0.21-0.22 (0.22)	0.22-0.23 (0.23)	0.23 (0.23)	0.25 (0.25)
				Case II					
2	0.29-0.30	0.31-0.32	0.31-0.32	0.33-0.34	0.36	0.37	0.40	0.42	0.50
3	0.22	0.23-0.24	0.24-0.25	0.25-0.26	0.26-0.27	0.27-0.28	0.29-0.30	0.29-0.31	0.33-0.35
4	0.17	0.18-0.19	0.19	0.20-0.21	0.21	0.21-0.22	0.22-0.23	0.23-0.24	0.25

[1] values in parentheses are the approximated optimal stimulus frequencies from Liu and Frank (2004)

Of course, $\xi^d_{1/2}$ has an efficiency of 100% if $\lambda = 1/2$, but this value is not shown in Table 2.

From Table 2, the optimal designs for pairwise contrasts ($\lambda = 0$) have an efficiency of less than 7% if interest is only in individual stimulus effects ($\lambda = 1$). On the other hand, optimal designs for individual stimulus effects are more robust, with an efficiency of at least 58.8% for pairwise contrasts. The table can also be used to find designs that achieve similar relative efficiencies for the two competing interests. For $Q = 2$, $\xi^d_{1/8}$ achieves similar $F^*_d(\xi; 1)$- and $F^*_d(\xi; 0)$-values. This holds for Case I and Case II. For $Q = 3$ and $Q = 4$, the same is true for $\xi^d_{1/4}$ and $\xi^d_{1/3}$, respectively. Moving away from these λs, we find designs that achieve a higher relative efficiency for one interest than for the other. Note that, equal weight for individual effects and pairwise contrasts (i.e., $\lambda = 1/2$) does not necessarily yield balanced relative efficiencies for the two interests.

4.2 Designs for estimation

In this simulation, we study the optimal designs when the focus is on estimation; i.e., $w = 1$ in (3). For this situation, m-sequence-based designs are often recommended. They are included as initial designs of our search algorithm, which is used again to search for optimal designs under the current setting. Table 3 presents the $F_e(\xi; \lambda)$-value of our

Table 2: Relative efficiency (percentage) for individual stimulus effects and pairwise contrasts of our designs for detection

Q	Criterion	Designs for Detection								
		ξ_1^d	$\xi_{3/4}^d$	$\xi_{2/3}^d$	$\xi_{1/2}^d$	$\xi_{1/3}^d$	$\xi_{1/4}^d$	$\xi_{1/8}^d$	$\xi_{1/16}^d$	ξ_0^d
		Case I								
2	$F_d^*(\xi;\lambda=1)$	100.0	99.5	99.0	97.3	94.0	89.7	78.3	64.9	4.9
	$F_d^*(\xi;\lambda=0)$	58.8	62.2	63.7	66.8	70.7	73.7	79.8	85.2	100.0
3	$F_d^*(\xi;\lambda=1)$	100.0	98.7	97.4	93.9	87.3	81.6	67.6	56.9	6.1
	$F_d^*(\xi;\lambda=0)$	63.6	69.8	71.8	75.6	79.7	82.9	88.6	90.1	100.0
4	$F_d^*(\xi;\lambda=1)$	100.0	97.3	95.5	90.7	83.2	77.6	61.9	47.9	3.0
	$F_d^*(\xi;\lambda=0)$	68.5	76.1	78.2	81.8	85.4	87.4	91.8	94.1	100.0
		Case II								
2	$F_d^*(\xi;\lambda=1)$	100.0	99.6	99.3	97.4	93.4	90.4	79.3	65.6	5.5
	$F_d^*(\xi;\lambda=0)$	59.6	63.1	64.3	67.5	71.6	73.7	79.5	84.8	100.0
3	$F_d^*(\xi;\lambda=1)$	100.0	98.5	97.5	93.7	88.1	82.7	64.8	55.9	6.7
	$F_d^*(\xi;\lambda=0)$	64.4	69.9	71.6	75.5	79.0	81.6	84.0	90.6	100.0
4	$F_d^*(\xi;\lambda=1)$	100.0	98.8	96.8	91.7	84.1	75.5	62.4	48.1	3.6
	$F_d^*(\xi;\lambda=0)$	68.1	74.2	75.9	80.2	83.7	84.4	88.7	91.6	100.0

Table 3: Efficiency (percentage) for estimation of our designs relative to the m-sequence-based designs

Q	Designs for Estimation								
	ξ_1^e	$\xi_{3/4}^e$	$\xi_{2/3}^e$	$\xi_{1/2}^e$	$\xi_{1/3}^e$	$\xi_{1/4}^e$	$\xi_{1/8}^e$	$\xi_{1/16}^e$	ξ_0^e
	Case I								
2	101.8	100.3	100.2	100.2	101.9	103.9	109.6	116.8	146.6
3	104.9	101.0	100.3	100.5	102.0	103.8	109.3	114.2	131.8
4	108.9	101.3	100.3	100.0	101.0	102.3	106.4	110.3	122.3
	Case II								
2	106.5	105.3	105.1	105.5	107.6	109.9	116.9	125.1	157.3
3	109.7	106.2	105.8	106.2	108.6	110.8	116.7	122.2	141.6
4	112.6	105.1	104.2	103.7	105.5	107.2	111.7	115.8	128.8

designs relative to the initial m-sequence-based design (in percentage). From the table, the efficiency of the m-sequence-based designs can be improved markedly when λ moves away from 1/2. When λ is close to 1/2, it is hard for Case I to find designs that are better than the m-sequence-based designs. In that particular situation, the stimulus frequencies of the m-sequence-based designs are close to the approximated optimal stimulus frequencies of Liu and Frank (2004). Our designs also yield similar frequencies in that situation. For other values of λ, our approach yields designs with higher efficiencies and the stimulus frequencies of the resulting designs are consistently in good agreement with the optimal stimulus frequencies.

It is known that the m-sequence-based designs can be suboptimal

under correlated noise and the appearance of drift; see, e.g., Kao et al. (2009) and Buračas and Boynton (2002). In contrast, our approach can adapt to these situations and lead to better designs. This is reflected in Table 3 for Case II. Even for $\lambda = 1/2$, when the stimulus frequency of our designs is similar to that of the m-sequence-based designs, the pattern of our designs beats that of the m-sequence-based designs.

Table 4, which is to be read in a similar way as Table 2, presents the relative efficiency achieved by our designs. Similar as for detection, ξ_1^e is more robust to a change in interests than ξ_0^e. Designs $\xi_{1/8}^e$ for $Q = 2$, $\xi_{1/4}^e$ for $Q = 3$ and $\xi_{1/16}^e$ for $Q = 4$ provide similar relative efficiencies for the two interests.

Table 4: Relative efficiency (percentage) for individual stimulus effects and pairwise contrasts of our designs for estimation

Q	Criterion	Designs for Estimation								
		ξ_1^e	$\xi_{3/4}^e$	$\xi_{2/3}^e$	$\xi_{1/2}^e$	$\xi_{1/3}^e$	$\xi_{1/4}^e$	$\xi_{1/8}^e$	$\xi_{1/16}^e$	ξ_0^e
		Case I								
2	$F_e^*(\xi;\lambda=1)$	100.0	99.1	98.9	97.2	94.4	89.9	77.4	67.3	10.6
	$F_e^*(\xi;\lambda=0)$	61.5	65.7	66.4	69.4	73.0	76.8	83.3	87.4	100.0
3	$F_e^*(\xi;\lambda=1)$	100.0	98.2	96.2	94.6	87.9	83.7	71.0	56.6	11.6
	$F_e^*(\xi;\lambda=0)$	66.8	72.5	75.2	77.2	81.5	83.6	88.6	92.1	100.0
4	$F_e^*(\xi;\lambda=1)$	100.0	97.3	95.1	91.5	83.0	76.5	62.4	48.4	11.6
	$F_e^*(\xi;\lambda=0)$	68.7	75.7	78.3	81.6	85.6	87.8	91.7	94.7	100.0
		Case II								
2	$F_e^*(\xi;\lambda=1)$	100.0	99.1	99.1	97.4	92.8	88.3	79.8	65.7	13.9
	$F_e^*(\xi;\lambda=0)$	61.4	65.5	65.5	69.2	74.5	78.1	82.6	88.2	100.0
3	$F_e^*(\xi;\lambda=1)$	100.0	98.3	97.3	93.6	86.6	82.0	68.4	57.7	12.9
	$F_e^*(\xi;\lambda=0)$	66.2	71.8	73.4	77.5	82.1	84.3	89.0	91.5	100.0
4	$F_e^*(\xi;\lambda=1)$	100.0	97.6	95.7	90.2	81.9	76.5	61.5	48.7	8.0
	$F_e^*(\xi;\lambda=0)$	68.5	75.2	77.4	81.4	85.8	87.8	91.9	94.4	100.0

4.3 Multi-objective designs

In this simulation, we allow the weight w to increase from 0 to 1 in steps of 0.05, thereby, gradually shifting emphasis from the detection problem to the estimation problem. For each Q and λ and for both cases, 21 optimal designs are obtained (one for each w-value). Figure 4 presents the $F_e^*(\xi;\lambda)$- versus $F_d^*(\xi;\lambda)$-values of the resulting designs for $Q = 2$ and Case I. Similar figures for other settings provide no further insight and are therefore omitted. The designs introduced in Section 2 are also presented in the figure. The block design in Figure 4 is the initial block design of our search algorithm. Mixed designs are

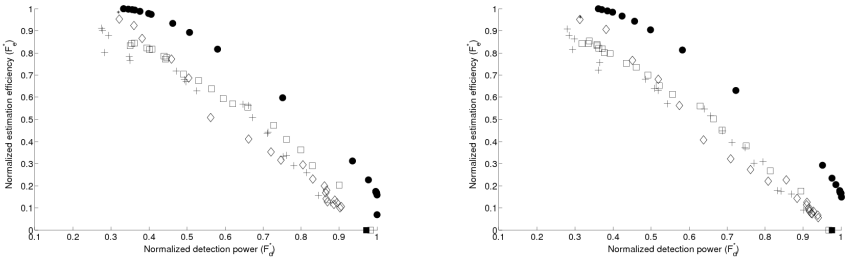
generated by combining a block design (our initial design) and an m-sequence-based design via crossover (Step 2 of the search algorithm in Subsection 3.3). This m-sequence-based design along with permuted block designs and clustered m-sequences are generated from the program provided by Liu (2004). For the permuted block designs, we choose a block design with a block size of eight since other block sizes do not seem to yield better results. As demonstrated in Liu (2004), for Case I and $\lambda = 1/2$, these designs provide an advantageous trade-off between the two statistical objectives, estimation and detection. Our designs provide a better trade-off, even for that special case. We note that the designs considered by Liu (2004) have a fixed stimulus frequency of $1/(Q+1)$. They can be good when $\lambda = 1/2$ because this is the optimal stimulus frequency for that case. However, when moving away from $\lambda = 1/2$, these designs are sub-optimal. Our approach finds much better designs as shown in Figure 4.

Table 5 presents means and standard deviations of the stimulus frequencies of our designs, which are computed over the 21 designs generated for each setting and the different stimulus types. Small standard deviations show that the stimulus frequencies vary little over the designs and the stimulus types. Again, these frequencies agree with the approximated optimal stimulus frequencies in Table 1.

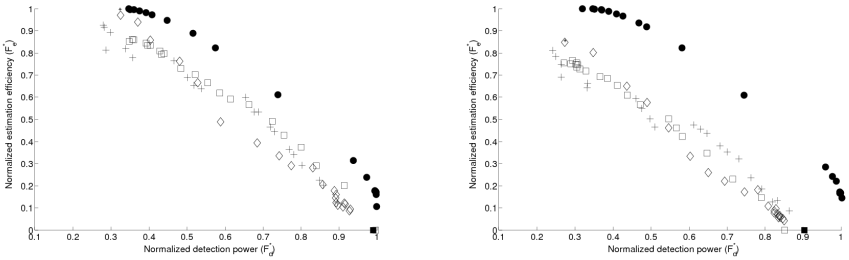
Table 5: Stimulus frequencies (mean and standard deviation) of our designs[1]

Q	λ 1	3/4	2/3	1/2	1/3	1/4	1/8	1/16	0
					Case I				
2	0.302	0.318	0.324	0.339	0.360	0.370	0.400	0.425	0.494
	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
3	0.218	0.234	0.242	0.253	0.267	0.276	0.292	0.302	0.330
	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)
4	0.169	0.187	0.192	0.201	0.210	0.216	0.226	0.232	0.249
	(0.002)	(0.002)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
					Case II				
2	0.298	0.316	0.323	0.338	0.360	0.373	0.400	0.426	0.495
	(0.005)	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)	(0.005)	(0.007)	(0.007)
3	0.218	0.234	0.242	0.253	0.267	0.276	0.292	0.303	0.330
	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)	(0.006)	(0.007)
4	0.169	0.187	0.192	0.202	0.211	0.216	0.226	0.233	0.249
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)

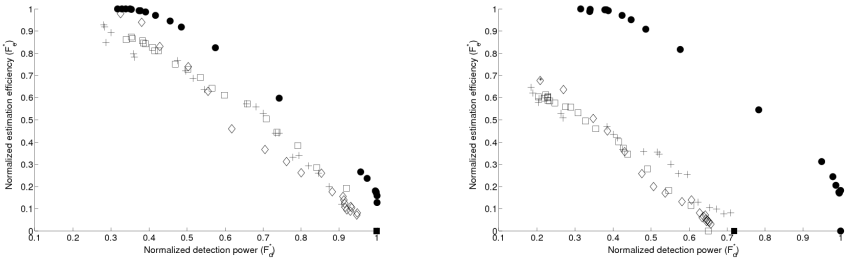
[1] the mean and standard deviation are taken over the Q stimulus types and the 21 designs



$\lambda = 1$ $\lambda = 1/4$



$\lambda = 3/4$ $\lambda = 1/16$



$\lambda = 1/2$ $\lambda = 0$

Figure 4: $F_e^*(\xi;\lambda)$-values versus $F_d^*(\xi;\lambda)$-values of designs obtained for Case I with $Q = 2$. $\bullet$: designs found by our approach; $*$: m-sequence; $\blacksquare$: block design; $\diamond$: clustered m-sequences; $\square$: permuted block designs; $+$: mixed designs;

5 Discussion and conclusions

In this article, we study optimal ER-fMRI designs when both individual stimulus effects and pairwise contrasts are of interest. These two interests are among the main concerns of fMRI researchers. Therefore, when planning ER-fMRI designs, it is crucial to find designs that are efficient for both interests.

Previous work either considers these two interests separately (Kao et al., 2009) or assigns equal weights to them (Liu and Frank, 2004; Liu, 2004). In contrast, we propose an approach for finding optimal designs allowing user-specified weights.

In our numerical results, we find (near-)optimal designs for each λ. The stimulus frequencies in these designs increase when λ decreases. This also means that the frequency of the control decreases when λ decreases. This phenomenon is expected from the approximated optimal stimulus frequencies derived by Liu and Frank (2004). Their approximation is derived for white noise and with neither drift nor trend, that is for our Case I. Our designs have stimulus frequencies that are in good agreement with the approximated optimum, not only for Case I but also for Case II.

Our numerical results show that the choice of $\lambda = 1/2$ does not necessarily yield a design that achieves similar relative efficiencies for the two interests. The value of λ that achieve this objective are, approximately, $\lambda = 1/8$ for $Q = 2$, $\lambda = 1/4$ for $Q = 3$, and $\lambda = 1/3$ for $Q = 4$.

We also observe that designs for individual stimulus effects retain a reasonable efficiency for pairwise contrasts and that they are relatively robust with respect to a change in interests. On the other hand, designs that are optimal for pairwise contrasts can have low efficiencies for estimating individual stimulus effects. These designs should not be used unless pairwise contrasts are the only concern.

Working with our defined MO-criteria, the search algorithm of Kao et al. (2009) can be applied for finding multi-objective optimal designs when both individual stimulus effects and pairwise contrasts are of interest. For detecting activation, the algorithm yields designs with a block structure. For estimating the HRF, we can find designs that work much better than the m-sequence-based designs. When considering these statistical objectives simultaneously, we find designs that provide advantageous trade-offs between the two competing objec-

tives. The algorithm can accommodate user-specified weights for the two objectives.

In addition to statistical objectives, Kao et al. (2009) also consider psychological constraints, and customized requirements when finding multi-objective optimal designs. It is straightforward to include these additional objectives in our family of MO-criteria and the search algorithm of Kao et al. (2009) can still be used for finding optimal designs. However, for the sake of clarity, we have only focused on statistical objectives in this study.

Acknowledgements : The research of John Stufken was in part supported by NSF Grant DMS-07-06917.

References

Amaro, E. J. and Barker, G. J. (2006). Study design in MRI: Basic principles, *Brain and Cognition* **60**, 220–232.

Barker, H. A. (2004). Primitive maximum-length sequences and pseudo-random signals, *Transactions of the Institute of Measurement & Control* **26**, 339–348.

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

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., 53–82.

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

Donaldson, D. I. and Buckner, R. L. (2002). Effective paradigm design in *Functional MRI: An Introdcution to Methods*, eds. Jezzard, P., Matthews, P. M., and Smith, S. M., New York: Oxford University Press, 177–195.

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.

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., 193–210.

Holland, J. H. (1975). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence,* Ann Arbor: University of Michigan Press.

Holland, J. H.(1992). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, Complex adaptive systems, Cambridge, Massachusetts: MIT Press.

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.

Josephs, O., Turner, R., and Friston, K. (1997). Event-related fMRI. *Human Brain Mapping* **5**, 243–248.

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.

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.

MacWilliams, F. J. and Sloane, N. J. A. (1977). *The Theory of Error Correcting Codes*, North-Holland Pub. Co.

Rosen, B. R., Buckner, R. L., and Dale, A. M. (1998). Event-related functional MRI: Past, present, and future. *PNAS* **95**, 773–780.

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

Worsley, K. J. and Friston, K. J. (1995). Analysis of fMRI time-series revisited–again. *NeuroImage* **2**, 173–181.

Ming-Hung Kao
Department of Statistics
The University of Georgia
E-mail: mhk@gmail.com; jasomkao@uga.edu

Abhyuday Mandal
Department of Statistics
The University of Georgia

John Stufken
Department of Statistics
The University of Georgia
E-mail: jstufken@uga.edu