NeuroImage 44 (2009) 849–856

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Technical Note

Multi-objective optimal experimental designs for event-related fMRI studies

Ming-Hung Kao *, Abhyuday Mandal, Nicole Lazar, John Stufken

Department of Statistics, University of Georgia, 101 Cedar Street, Athens, GA 30602, USA

ARTICLE INFO

Article history:
Received 24 February 2008
Revised 16 September 2008
Accepted 22 September 2008
Available online 4 October 2008

Keywords:
Compound design criterion
Design efficiency
Genetic algorithms

ABSTRACT

In this article, we propose an efficient approach to find optimal experimental designs for event-related functional magnetic resonance imaging (ER-fMRI). We consider multiple objectives, including estimating the hemodynamic response function (HRF), detecting activation, circumventing psychological confounds and fulfilling customized requirements. Taking into account these goals, we formulate a family of multi-objective design criteria and develop a genetic-algorithm-based technique to search for optimal designs. Our proposed technique incorporates existing knowledge about the performance of fMRI designs, and its usefulness is shown through simulations. Although our approach also works for other linear combinations of parameters, we primarily focus on the case when the interest lies either in the individual stimulus effects or in pairwise contrasts between stimulus types. Under either of these popular cases, our algorithm outperforms the previous approaches. We also find designs yielding higher estimation efficiencies than *m*-sequences. When the underlying model is with white noise and a constant nuisance parameter, the stimulus frequencies of the designs we obtained are in good agreement with the optimal stimulus frequencies derived by Liu and Frank, 2004, *NeuroImage* 21: 387-400. In addition, our approach is built upon a rigorous model formulation.

Introduction

ER-fMRI is one of the leading technologies for studying human brain activity in response to mental stimuli (Josephs et al., 1997; Rosen et al., 1998; Dale, 1999; Bandettini and Cox, 2000). Before conducting an ER-fMRI experiment, a design sequence consisting of stimuli of one or more types interlaced with rests is prepared. This sequence is presented to an experimental subject, while the MR scanner measures changes in the subject's blood oxygenation level dependent (BOLD) response for the end purpose of statistical inference. The design issue here is to best allocate the stimuli so that inference is precise and valid.

Two common statistical goals in ER-fMRI are to estimate the HRF (the noise-free BOLD time series triggered by a single, brief stimulus), and to detect brain activation; see also Buxton et al. (2000) and Birn et al. (2002). Considering both goals in one experiment is not uncommon, but it requires a good multi-objective design that simultaneously achieves high efficiencies on both dimensions. However, statistical efficiency is not the only concern for planning ER-fMRI design sequences. Psychology plays an important, even crucial, role. When a design sequence is patterned or easy to predict, psychological effects such as habituation or anticipation may occur to confound stimulus effects (Dale, 1999). Therefore, a good design should provide safeguards against the psychological confounds while retaining a high efficiency for statistical inference. Moreover, customized requirements such as a required frequency for each stimulus type might also arise to further complicate the design problem. As a consequence, the search for a good, multi-objective design is inevitable and a well-defined multi-objective design criterion (or MO-criterion for short) is needed to evaluate competing designs. In addition, the design space, consisting of all possible ER-fMRI designs, is enormous and irregular (Buračas and Boynton, 2002; Liu, 2004). Searching over this huge space for an optimal design is an arduous task, thus an efficient search algorithm is as well crucial.

Wager and Nichols (2003), referred to as WN henceforward, propose a framework for finding multi-objective optimal ER-fMRI designs. They formulate the MO-criterion as a weighted average of the design criteria for the individual objectives of interest. A modified genetic algorithm (or WN's GA) is also introduced to search for optimal or near-optimal multi-objective designs. This trailblazing work has been applied in many studies over the last few years (e.g., Callan et al., 2006; Ramautar et al., 2006; Summerfield et al., 2006; Wang et al., 2007).

Inspired by WN's pioneering work, we develop an efficient approach to search for optimal multi-objective designs. Our approach has two major advantages. First, we incorporate well-known fMRI designs in our algorithm to facilitate the search. Second, we define a family of MO-criteria that allows consistent design comparisons. While crucial to the success of a search algorithm, WN's criteria do not always achieve this. Furthermore, our algorithm is simple and easy to implement, yet effective.

* Corresponding author. Fax: +1 706 542 3391.
E-mail address: jasonkao@uga.edu (M.-H. Kao).

1053-8119/$ – see front matter
doi:10.1016/j.neuroimage.2008.09.025

The efficiency and effectiveness of our approach are demonstrated through simulations under two popular cases, one focuses on individual stimulus effects and the other on pairwise contrasts. We also discuss the situation when both cases are simultaneously of interest. While taking less computation time than WN's approach, our algorithm achieves designs with significantly higher efficiencies. We also demonstrate that our designs form an advantageous trade-off between estimation efficiency and detection power, and we find designs yielding higher estimation efficiencies than *m*-sequences. Moreover, under the model with white noise and a constant nuisance parameter, the stimulus frequencies of the designs we obtained are in good agreement with the optimal stimulus frequencies derived by Liu and Frank (2004).

In this technical note, our proposed algorithm is introduced, and its performance is demonstrated via simulations. Other details and additional simulations are presented in (Kao et al., 2007). The rest of the article is organized as follows. Section 2 presents our proposed approach. Simulations are provided in Section 3. Conclusions and a discussion are in Section 4.

Methodology

We propose an efficient and effective approach to search for optimal multi-objective designs for ER-fMRI. Four objectives are considered: 1) estimating the HRF, 2) detecting brain activation, 3) avoiding psychological confounds, and 4) maintaining the desired stimulus frequency in the design sequence. By assigning weights to these objectives based on the researcher's discretion, our algorithm finds a design best suited to the researcher's needs. We briefly introduce our approach in this section. The approach is general enough that other objectives, beside the four listed above, could be accommodated as well.

Underlying model and design criteria

To find an optimal design, we need to specify the underlying model for the two primary statistical objectives, namely estimation and detection. As in WN and Liu and Frank (2004), two popular linear models are considered (Friston et al., 1995; Worsley and Friston, 1995; Dale, 1999):

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

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

where $\boldsymbol{Y}$ is the voxel-wise BOLD time series, $\boldsymbol{h}=(\boldsymbol{h}_1',\ldots\boldsymbol{h}_Q')'$ 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, $\theta=(\theta_1,\ldots,\theta_Q)'$ represents the response amplitudes, $\boldsymbol{Z}=\boldsymbol{X}\boldsymbol{h}_0$ is the convolution of stimuli with an assumed basis, $\boldsymbol{h}_0$, of the HRF, $\boldsymbol{S}\gamma$ is a nuisance term describing the trend or drift of $\boldsymbol{Y}$, and $\boldsymbol{e}$ and $\boldsymbol{\eta}$ are noise. Following WN, 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. Under these models, the *A*- or *D*-optimal design criteria can be applied to evaluate competing designs with respect to the objectives of estimation and detection. Both of these criteria are widely accepted and the choice between *A*- and *D*-optimality depends on individual preference. *A*-optimality aims at minimizing the average variance of estimators of parametric functions. In our simulations, these will be individual stimulus effects, or pairwise contrasts. On the other hand, a *D*-optimal design minimizes the generalized variance of estimators of linearly independent parametric functions, or, under normality, it minimizes the volume of simultaneous elliptical confidence regions for these parametric functions at any specified confidence level. For our problem, these parametric functions will be either individual stimulus effects, or (Q-1) linearly independent pairwise contrasts. For further details, see Atkinson et al. (2007).

For technical reasons, we formulate these design criteria as "larger-the-better" criteria, and designs maximizing them help to optimize statistically meaningful functions of the parameter estimators as previously described. The value of the design criterion for estimation, referred to as "estimation efficiency", is denoted by F_e. Likewise, the term "detection power" and the notation F_d are used to indicate the value of the design criterion for detection. These two criteria are defined to have one of the following two forms:

$$F_i = \begin{cases} r_c/trace(\boldsymbol{M}), & for\ A\text{−}optimality; \\ det(\boldsymbol{M})^{-1/r_c}, & for\ D\text{−}optimality, \end{cases} \tag{3}$$

where $\boldsymbol{M}=\boldsymbol{C}[\boldsymbol{W}'\boldsymbol{V}'(\boldsymbol{I}\text{-}P_{VS})\boldsymbol{V}\boldsymbol{W}]^{-}\boldsymbol{C}'$, $\boldsymbol{W}\equiv\boldsymbol{X}$ for F_e, $\boldsymbol{W}\equiv\boldsymbol{Z}$ for F_d, $\boldsymbol{I}$ is an identity matrix, $P_{\boldsymbol{A}}=\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-}\boldsymbol{A}'$ is the orthogonal projection on the vector space spanned by the column vectors of $\boldsymbol{A}$, $\boldsymbol{A}^{-}$ is a generalized inverse matrix of $\boldsymbol{A}$, $\boldsymbol{C}$ is a matrix of linear combinations of the parameters, and r_c is the number of rows of $\boldsymbol{C}$.

The third objective is to avoid psychological confounds. We would like a sequence that makes it difficult for a subject to anticipate future stimuli based on past stimuli. To achieve this, the *R*th order counter-balancing property of WN is considered, where *R* is a given integer. This property is defined on a sub-design of the original design obtained by keeping only the stimuli but deleting all rests. For any $r\in\{1,\ldots,R\}$, we count the pairs of stimuli that appear in positions $(t,t+r)$ in the sub-design, $t=1,\ldots,(n\text{-}r)$; n is the length of the sub-design. The *R*th order counterbalancing aims at having each pair appear a number of times that is proportional to the product of the specified proportions for the stimuli. The corresponding design criterion can be written as:

$$F_c = \sum_{r=1}^{R}\sum_{i=1}^{Q}\sum_{j=1}^{Q}\lfloor |n_{ij}^{(r)}-(n-r)P_iP_j| \rfloor,$$

where $n_{ij}^{(r)}$ is the number of occurrences of a type-i stimulus being the tth element and a type-j stimulus being the $(t+r)$th element, $t=1,\ldots,(n\text{-}r)$, P_i is the specified proportion for the type-i stimulus in the sub-design which may be taken as $1/Q$ if there is no preference, and $\lfloor |a| \rfloor$ is the integer part of the absolute value of a. This criterion measures the departure from counterbalancing and is a "smaller-the-better" criterion.

The fourth design criterion is also defined on the sub-design. It is $F_f=\sum_{i=1}^{Q}\lfloor |n_i-nP_i| \rfloor$, where n_i is the number of the type-i stimulus in the sub-design. This criterion helps to maintain the desired stimulus frequency and is a "smaller-the-better" criterion.

An MO-criterion is defined as a convex combination of the above four individual criteria. To ensure comparability, they are standardized before combining. We use the following standardization:

$$F_i^* = \begin{cases} \dfrac{F_i-\min(F_i)}{\max(F_i)-\min(F_i)}, & i=d,e; \\ 1-\dfrac{F_i-\min(F_i)}{\max(F_i)-\min(F_i)}, & i=c,f. \end{cases}$$

Our family of MO-criteria is then defined as $\{F^*=w_cF_c^*+w_dF_d^*+w_eF_e^*+w_fF_f^*:w_i\geq 0,i=c,d,e,f;\sum_i w_i=1\}$; w_is are weights selected based on the researcher's emphasis in a given study.

By contrast, WN standardize each F_i by its mean and standard deviation over designs within the current generation of their GA. Since designs change with successive generations, so do these means and standard deviations. The resulting MO-criteria are moving targets during the search. Thus, fair, consistent design comparisons may not be achieved. Our MO-criteria are free from this drawback.

With the MO-criterion for evaluating the "goodness" of competing designs, we propose a GA-based algorithm to search for the optimal ER-fMRI design.

Search algorithm

GAs (Holland, 1975, 1992) are popular for solving optimization problems, in which good solutions (parents) are used to generate better ones (offsprings). To efficiently apply this technique, we take advantage of well-known results about good fMRI designs so that our search over the huge design space can be carried out more efficiently. The outline of our 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 "genetic material" in paired designs. See (Wager and Nichols, 2003) for a nice graphical presentation. Here, the "genetic material" is the design sequence.
Step 3. (Mutation) Randomly select q% of the events from the G offspring designs. Replace these events by randomly generated ones. Here, an event is a stimulus or a rest.
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_g generations). Keep track of the best design over generations.

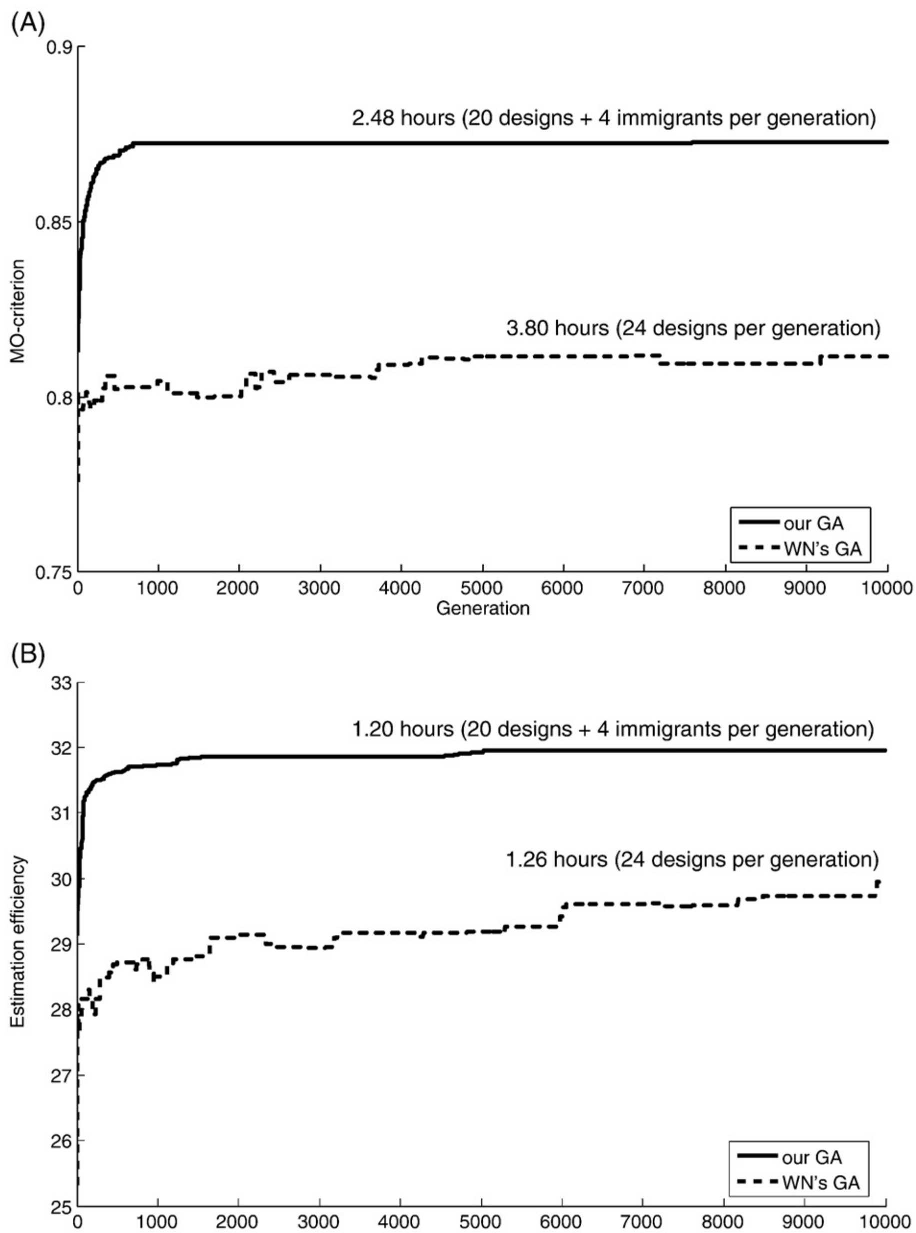

Fig. 1. Achieved values of design efficiency vs. generation for (A) $w_c=w_d=w_e=w_f$, (B) $w_e=1$, and (C) $w_d=1$. CPU times for completing 10,000 generations are presented.

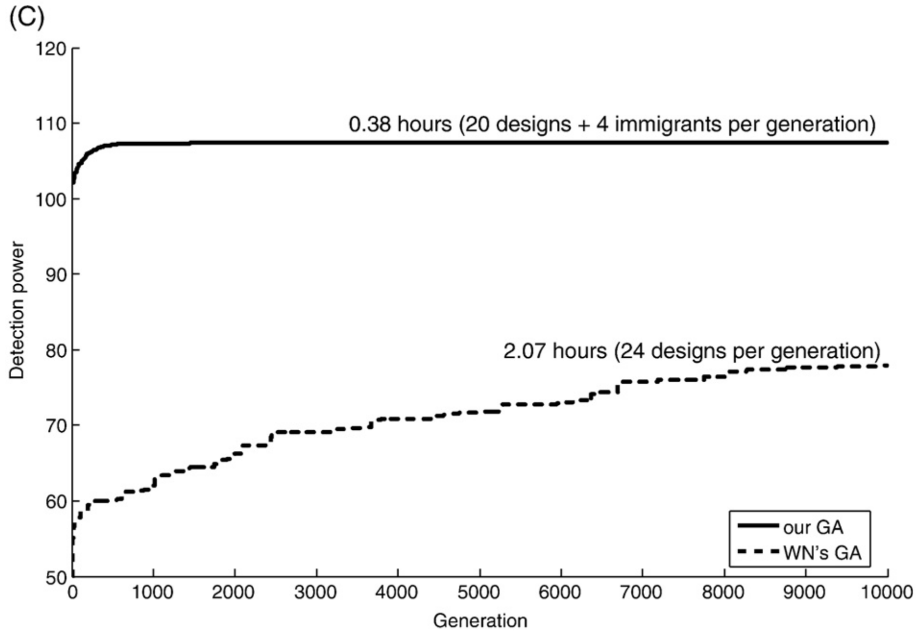

Fig. 1 (*continued*).

We describe below some details of our GA. MATLAB code implementing this algorithm can be found at http://www.stat.uga.edu/~amandal.

Initial designs and immigrants

In Step 1, *m*-sequence-based designs or *m*-sequences are generated following Liu (2004); see also Buračas and Boynton (2002). These designs are well-known for their high estimation efficiencies. Since they are not always available, concatenations or truncations of the existing ones are also considered. We include the one yielding the highest estimation efficiency as one of the initial designs.

The initial block design has the highest detection power among designs of differing numbers of blocks and of two different patterns. In this pool of candidate block designs, the number of blocks for each stimulus type ranges among one to five, 10, 15, 20, 25, 30, and 40. The two patterns include repetitions of NABC and NANBNC, where N is a block of rests and A, B and C represent blocks of stimuli of different types. In addition to the initial block design, immigrants in Step 4 ensure a steady supply of blocks of different sizes.

The combination of a block design with an *m*-sequence-based design or a random design is obtained through crossover. These mixed designs constitute a portion, e.g., one-third, of the initial designs. The remaining initial designs are formed by random designs.

Objective function

The objective function used in Step 1 and Step 5 of our GA evaluates the fitness or "goodness" of the designs. Based on the goal of the search, the objective function can be taken as a single F_i or as an MO-criterion with weights selected by the researcher's interest. Note that the extreme values of the F_is are required to use our MO-criteria.

Theoretical values of max(F_e) and max(F_d) are generally not available. They can be approximated by performing a "pre-run" of our 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 parameters of interest are non-estimable. Both min(F_c) and min(F_f) are zero. Their maximal values are attained by the design containing only the stimulus type with the smallest specified proportion P_i. With these extreme values and given weights, an MO-criterion F^* is well-defined and serves as the objective function for finding optimal multi-objective designs.

Simulations

In the following illustrative simulations, we consider designs with three stimulus types (Q=3) and L=255 events. The ISI (inter-stimulus interval, time between consecutive event onsets) and the TR (time to repetition, or sampling rate) are both set to two seconds.

For F_e and F_d, we use the *A*-optimality criterion and consider two popular situations, namely individual stimulus effects and pairwise contrasts. For the former situation, the $\boldsymbol{C}$ matrix described after (3) is the identity matrix. For the latter case, the rows of $\boldsymbol{C}$ correspond to the $Q(Q\text{-}1)/2$ pairwise contrasts between stimulus types. The canonical HRF, a combination of two Gamma distributions (SPM2, http://www.fil.ion.ucl.ac.uk/spm), is used as $\boldsymbol{h}_0$ in model (2). In the first two simulations, the drift, described by $\boldsymbol{S}\gamma$, is assumed to be a second-order Legendre polynomial, and the noise follows a stationary AR(1) process with a correlation coefficient of 0.3. In the last simulation, white noise is assumed and $\boldsymbol{S}$ is taken to be a vector of ones. As for F_c and F_f, we require a third-order counterbalancing property (R=3) and equal frequencies for the three stimulus types; i.e., P_i=1/3, i=1,2,3.

Unless otherwise specified, the algorithmic parameters are G (size of population)=20, q (percentage of mutation)=1%, I (number of immigrants)=4 and M_g (number of generations)=10,000. A larger value of M_g does not seem to lead to significantly better designs. The simulations are performed on a Pentium Dual 3.20/3.19 GHz computer with 3.5 Gb of RAM.

Simulation 1

We first consider three weighting schemes, namely (A) w_c=w_d=w_e=w_f=0.25, (B) w_e=1, and (C) w_d=1, with the $\boldsymbol{C}$ matrix being the identity matrix. The first weighting scheme finds a multi-objective design, whereas the latter two schemes search for the best designs for estimation and detection, respectively. The achieved values of the design criterion over the 10,000 GA generations are presented in Fig. 1. For weighting scheme (A), the value of the MO-criterion is presented. The estimation efficiency, F_e, and detection power, F_d, are reported for weighting schemes (B) and (C), respectively.

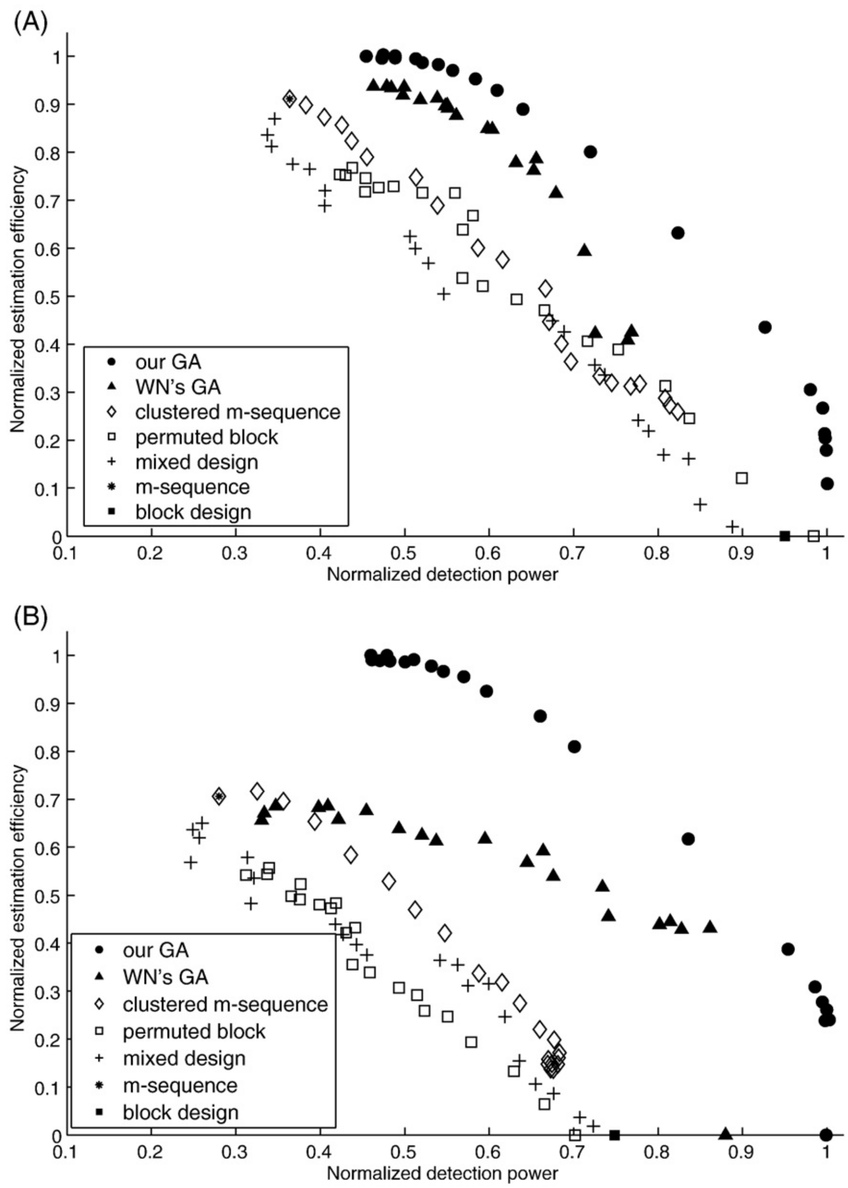

Fig. 2. Normalized estimation efficiency vs. detection power for different designs: (A) individual stimulus effects; (B) pairwise contrasts.

Our GA is compared to WN's GA. For comparison, we include 24 designs in each generation of their GA since they do not allow immigration. As shown in Fig. 1A, our GA achieves a value of the MO-criterion of 0.873 while WN's GA attains 0.812. In addition, our algorithm uses less CPU time than their GA. Significant improvements made by our GA are also observed in Fig. 1 for the other two weighting schemes. Note that the efficiency curve for the MO-criterion in WN's GA is not monotone, a result of the inconsistency of their normalization method that was pointed out in Subsection 2.1.

Under weighting scheme (B), our GA finds a design yielding a higher estimation efficiency than the *m*-sequence-based design. The estimation efficiency is 31.96 for our design compared to 29.12 for the *m*-sequence-based design. Our design, featuring small off-diagonal elements in the information matrix (not shown), possesses a property similar to the "decorrelation" property described in Buračas and Boynton (2002). In their paper, random designs with this property are observed to yield higher estimation efficiencies than *m*-sequence-based designs when correlated noise is assumed. In addition to

Table 1
The F_e-values and the proportions of the stimuli: individual stimulus effects

Number of types (Q)	2	3	4	6	7	8	10	12
Length of design (L)	242	255	624	342	511	728	1330	2196
F_e-value								
our GA	41.17	33.34	68.39	26.76	36.20	46.68	72.73	105.18
m-sequence	40.43	31.80	63.08	24.33	31.94	40.72	61.38	85.39
Stimulus proportion								
our GA (min-max)	0.29	0.21-0.23	0.17	0.12-0.13	0.10-0.11	0.09-0.10	0.08	0.06-0.07
approximated optimum	0.29	0.21	0.17	0.12	0.10	0.09	0.08	0.06
CPU time (hours)	0.07	0.11	0.46	0.37	0.68	1.35	4.51	13.09

Table 2
The F_e-values and the proportions of the stimuli: pairwise contrasts

Number of types (Q)	2	3	4	6	7	8	10	12
Length of design (L)	242	255	624	342	511	728	1330	2196
F_e-value								
our GA	56.52	39.04	75.19	25.59	33.38	42.02	62.43	86.56
m-sequence	38.74	30.23	61.70	23.57	31.42	39.95	60.00	84.14
Stimulus proportion								
our GA (min-max)	0.49	0.32-0.33	0.24-0.25	0.16-0.17	0.14	0.12-0.13	0.10	0.08-0.09
approximated optimum	0.50	0.33	0.25	0.17	0.14	0.13	0.10	0.08
CPU time (hours)	0.06	0.11	0.47	0.43	0.87	1.70	5.70	16.49

correlated noise, our algorithm can also take into account the second-order polynomial drift, $\boldsymbol{S}\gamma$.

Another feature held by our design pertains to the stimulus frequency. While the stimulus proportion of the *m*-sequence-based design is always $1/(Q+1)$, that of our design concurs with the approximated optimal proportion of Liu and Frank (2004). The relative frequencies of the three stimulus types in our design are 0.21, 0.22, and 0.22, and the approximated optimal proportion is 0.21. As shown in Simulation 3, this agreement is reached consistently.

When focusing on weighting scheme (C), our GA finds a design close to a block design; see Fig. 4A in the Appendix. Although this design looks similar to the initial block design, our algorithm does not always yield designs that are similar to the initial ones. For example, when considering the pairwise contrasts between stimulus types, the design found by our GA contains only blocks of stimuli while the initial block design includes also rests; see Fig. 4B in the Appendix. Our GA tends to converge to a block design when detecting activation is the only concern. The design parameters, including block sizes, number of blocks and the design pattern, are tuned to yield a high efficiency along the evolution of our GA.

Simulation 2

This simulation focuses on the two statistical objectives — detection and estimation. By letting w_d increase from 0 to 1 in steps of 0.05 and keeping $w_c=w_f=0$ (w_e decreases accordingly), our GA finds designs providing an advantageous trade-off between estimation efficiency and detection power. Fig. 2A provides the result for individual stimulus effects and Fig. 2B for pairwise contrasts. Again, we compare our designs to WN's designs. Our designs significantly outperform theirs.

In addition, design efficiencies of mixed designs, clustered *m*-sequences and permuted block designs obtained from Liu (2004) are also presented. We also show in Fig. 2 the initial *m*-sequence-based designs of our algorithm, presented as *, and the initial block designs, denoted by ■.

As demonstrated by (Liu, 2004), mixed designs, clustered *m*-sequences and permuted block designs can offer advantageous trade-offs between estimation efficiency and detection power when individual stimulus effects and pairwise contrasts are simultaneously of interest. Results for this case are presented in Section 4.

Simulation 3

In this simulation, we follow Buračas and Boynton (2002) to work on white noise and set $\boldsymbol{S}$ in model (1) to a vector of ones, accounting for the overall mean of the fMRI time series. We focus on two separate cases, namely estimating $\boldsymbol{h}$ and estimating $\boldsymbol{h}_i$-$\boldsymbol{h}_j$ for $1\leq i<j\leq Q$. Different combinations of Q and L used by Liu (2004) are considered. Our GA then finds designs optimizing the estimation efficiency; i.e., $w_e=1$. For this comparison, we include only random designs as initial designs in our GA. Due to the computation time, here we let the algorithm run for only 2,000 generations at each combination.

We compare our designs to *m*-sequence-based designs, which are demonstrated by Buračas and Boynton (2002) to have high estimation efficiencies. The values of F_e achieved by our designs and by *m*-sequence-based designs are presented in Tables 1 and 2. The CPU time spent by our GA is also provided. Even without the help of the *m*-sequence-based design, our GA consistently finds better designs. As shown in Tables 1 and 2, the stimulus proportions of our designs are again in good agreement with those optimal values approximated by Liu and Frank (2004).

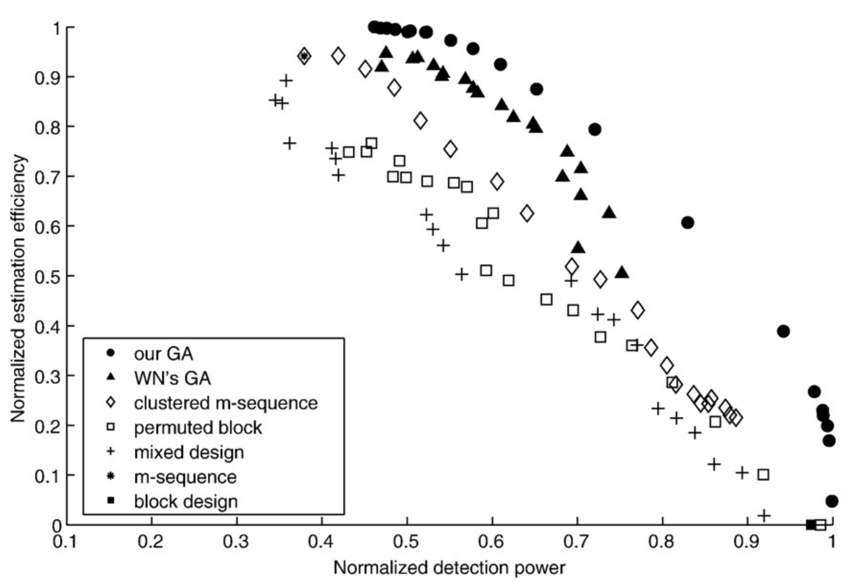

Fig. 3. Normalized estimation efficiency vs. detection power for different designs when both individual stimulus effects and pairwise contrasts are of interest.

Conclusions and discussion

In this technical note, we propose an algorithm to search for optimal ER-fMRI designs. Our proposed algorithm works for any combination of the four popular objectives in ER-fMRI, but is flexible enough to accommodate other goals as well. Through simulations, we show that our algorithm outperforms others currently in use by researchers when either the individual stimulus effects or pairwise contrasts are of interest.

Conceptually, our algorithm follows Holland's (1975) notion of building blocks; see also (Goldberg, 1989). Rooted in the fundamental theorem of GAs, also known as the schema theorem, the building block hypothesis views these constructs as the driving engine for GAs (Goldberg, 1989). Ensuring a good supply of these building blocks is thus one of the key steps for developing good GAs (Goldberg, 2002; Ahn, 2006). The inclusion of good ER-fMRI designs as both initial designs and immigrants follows this concept. Furthermore, using good ER-fMRI designs as initial designs also means that our algorithm starts from a good position.

The m-sequence-based design is not included as an initial design in Simulation 3 because our design is compared to this design. For Simulation 3 and quite a few other situations, we can find designs yielding higher estimation efficiencies than m-sequence-based designs without the benefit of an m-sequence-based design among the initial designs. However, this can be hard when both $\boldsymbol{h}$ and pairwise contrasts between $\boldsymbol{h}_i$s are of interest, and the model is with white noise but with neither drift nor trend. For that particular situation, the optimal stimulus proportion is $1/(Q+1)$ and the m-sequence-based design is known to be (near-)optimal (Liu and Frank, 2004; Liu, 2004). Note that the m-sequence-based design are known to exist only when $Q+1$ is a prime or a prime power. In contrast, our GA is flexible enough to accommodate any number of stimulus types.

While good initial fMRI designs help to expedite the search, the well-defined design criterion ensures that our GA, when it evolves, finds a better design. As pointed out previously, WN's design criterion is a moving target during the search. Achieving a better design is thus not guaranteed. By contrast, our MO-criterion provides a stable, clear target for the search algorithm.

Our algorithm approximates $\max(F_e)$ and $\max(F_d)$ that are needed for our MO-criterion. A possible alternative is to follow Liu and Frank (2004) to find analytical approximations. For the special cases of the Simulation 2, we apply their approach to find the bound for F_e. It is 34.16 when focusing on individual stimulus effects and is 42.50 for pairwise contrasts. These analytical approximations are larger than our numerical ones, which are 31.96 and 38.21, respectively. However, it is unknown whether their approximated $\max(F_e)$ can actually be achieved by any design. Also, the analytical approximation to $\max(F_d)$ depends on a parameter θ_{min}; see Liu and Frank (2004) for details. Deciding the value of θ_{min} suitable for each situation can be hard. Furthermore, the requisite bounds should adapt to a wide range of conditions, such as different correlation structures and nuisance terms. While these situations can easily be accommodated in our approach, it can be difficult to analytically derive bounds best suited to each circumstance.

Our algorithm can also be applied when individual stimulus effects and pairwise contrasts are simultaneously of interest. For illustration, consider the same conditions as in Simulation 2 of Section 3, where a second-order polynomial drift and AR(1) noise are assumed; such assumptions are closer to reality, compared to the model with white noise and without drift or trend. Fig. 3 presents the F_e*-value versus the F_d*-value achieved by our designs, WN's designs, and the designs studied by (Liu, 2004). Note that, in the case of detection, the matrix $\boldsymbol{C}$ after (3) is the identity matrix for Fig. 2A, and the rows of $\boldsymbol{C}$ for Fig. 2B represent the pairwise contrasts. Following (Liu, 2004), the matrix $\boldsymbol{C}$ for Fig. 3 combines all of these rows into one matrix. Similar comments apply for the estimation problem. Again, our algorithm yields better designs.

We note that it should be possible to find clustered m-sequences and permuted block designs to reach efficiencies as high as those of our designs. However, unlike our algorithm, the procedures to generate these designs do not attempt to maximize a design optimality criterion, so that finding good designs of these types using existing algorithms depends on luck. One may be able to develop an effective algorithm that uses the concepts on which these designs are based for finding efficient designs, but that would also require some procedures to hone in on the optimal stimulus frequencies. Pursuing this is beyond the scope of the current work.

One additional advantage of our GA that is not elaborated here, but in Kao et al. (2007), is the formulation of the statistical model when ISI≠mTR for any integer m. Our approach applies the discretization interval of Dale (1999) for the HRF parametrization. Denoting the length of this interval as ΔT, we set ΔT to the greatest value dividing both the ISI and TR. The resulting linear models agree with those of WN when ISI=mTR for some integer m, but our parameters remain interpretable when ISI≠mTR for any integer m. Specifically, the ith HRF parameter in WN's model corresponds to the height of the HRF at the ith scan after the stimulus onset. Each parameter in their models may simultaneously represent more than one height of the HRF when ISI≠mTR. By contrast, our underlying model faithfully reflects the fluctuation in the HRF, and thus results in a more rigorous model formulation.

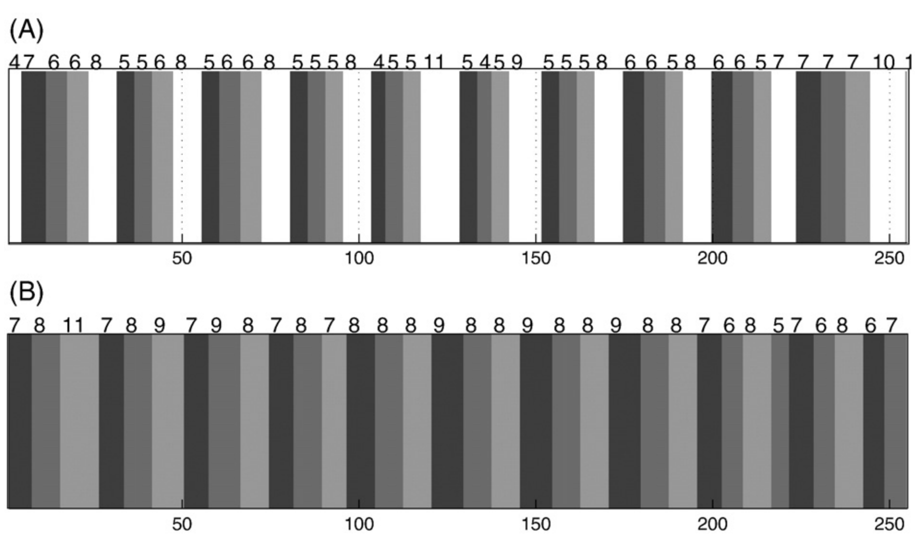

Fig. 4. Best designs for detection found by our GA: (A) individual stimulus effects; (B) pairwise contrasts.

Acknowledgments

The research of Nicole Lazar was in part supported by NSF Grant DMS-07-06192, and that of John Stufken by NSF Grant DMS-07-06917. The authors are thankful to the anonymous referees for their comments and suggestions, which resulted in an improvement of this work.

Appendix A

We provide in Fig. 4 the best designs for detecting activation found by our GA, assuming a second-order Legendre polynomial drift and a stationary AR(1) noise with a correlation coefficient of 0.3; for details see Section 3. Fig. 4A shows the design when the interest lies in individual stimulus effects and Fig. 4B is for pairwise contrasts. Different shades indicate different stimulus types with white representing rest. The number above each shaded bar presents the number of stimulus types included in that block. Both designs look like block designs. While rest is included in the first design, it is expelled by our GA when the interest lies only in pairwise contrasts. Note that the initial block designs for both searches contain rests.

References

Ahn, C.W., 2006. Advances in Evolutionary Algorithms: Theory, Design and Practice, Studies in computational intelligence. Springer, Berlin, New York.
Atkinson, A.C., Donev, A.N., Tobias, R.D., 2007. Optimum Experimental Designs, with SAS. Oxford University Press, Great Britain.
Bandettini, P.A., Cox, R.W., 2000. Event-Related fMRI Contrast When Using Constant Interstimulus Interval: Theory and Experiment. Magn. Reson. Med. 43, 540–548.
Birn, R.M., Cox, R.W., Bandettini, P.A., 2002. Detection versus Estimation in Event-Related fMRI: Choosing the Optimal Stimulus Timing. NeuroImage 15, 252–264.
Buračas, G.T., Boynton, G.M., 2002. Efficient Design of Event-Related fMRI Experiments Using M-Sequences. NeuroImage 16, 801–813.
Buxton, R.B., Liu, T.T., Martinez, A., Frank, L.R., Luh, W.M., Wong, E.C., 2000. Sorting Out Event-Related Paradigms in fMRI: The Distinction Between Detecting an Activation and Estimating the Hemodynamic Response. NeuroImage 11, S457.
Callan, A.M., Callan, D.E., Tajima, K., Akahane-Yamada, R., 2006. Neural Processes Involved with Perception of Non-Native Durational Contrasts. Neuroreport 17, 1353–1357.
Dale, A.M., 1999. Optimal Experimental Design for Event-Related fMRI. Hum. Brain Mapp. 8, 109–114.
Friston, K.J., Holmes, A.P., Poline, J.B., Grasby, P.J., Williams, S.C.R., Frackowiak, R.S.J., Turner, R., 1995. Analysis of fMRI Time-Series Revisited. NeuroImage 2, 45–53.
Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, Massachusetts.
Goldberg, D.E., 2002. The Design of Innovation: Lessons from and for Competent Genetic Algorithms. Kluwer Academic Publishers, Boston.
Holland, J.H., 1975. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications To Biology, Control, and Artificial Intelligence. University of Michigan Press, Ann Arbor.
Holland, J.H., 1992. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence, Complex adaptive systems, first ed., MIT Press, Cambridge, Massachusetts.
Josephs, O., Turner, R., Friston, K., 1997. Event-Related fMRI. Hum. Brain Mapp. 5, 243–248.
Kao, M.-H., Mandal, A., Lazar, N., Stufken, J., 2007. Multi-Objective Optimal Experimental Designs for Event-Related fMRI Studies. Tech. rep., Department of Statistics, University of Georgia, http://www.stat.uga.edu/~amandal/MO-fMRI.pdf.
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., Frank, L.R., 2004. Efficiency, Power, and Entropy in Event-Related fMRI with Multiple Trial Types: Part I: Theory. NeuroImage 21, 387–400.
Ramautar, J.R., Slagter, H.A., Kok, A., Ridderinkhof, K.R., 2006. Probability Effects in the Stop-Signal Paradigm: The Insula and the Significance of Failed Inhibition. Brain Res. 1105, 143–154.
Rosen, B.R., Buckner, R.L., Dale, A.M., 1998. Event-Related functional MRI: Past, Present, and Future. PNAS 95, 773–780.
Summerfield, C., Egner, T., Mangels, J., Hirsch, J., 2006. Mistaking a House for a Face: Neural Correlates of Misperception in Healthy Humans. Cerebral Cortex 16, 500–508.
Wager, T.D., Nichols, T.E., 2003. Optimization of Experimental Design in fMRI: A General Framework Using A Genetic Algorithm. NeuroImage 18, 293–309.
Wang, Y.P., Xue, G., Chen, C.S., Xue, F., Dong, Q., 2007. Neural Bases of Asymmetric Language Switching in Second-Language Learners: An ER-fMRI Study. Neuroimage 35, 862–870.
Worsley, K.J., Friston, K.J., 1995. Analysis of fMRI Time-Series Revisited–Again. NeuroImage 2, 173–181.