Robust Experimental Designs for fMRI with an Uncertain Design Matrix

by

Lin Zhou

A Thesis Presented in Partial Fulfillment
of the Requirement for the Degree
Master of Science

Approved October 2014 by the
Graduate Supervisory Committee:

Ming-Hung Kao, Chair
Mark Reiser
John Stufken
Bruno Welfert

ARIZONA STATE UNIVERSITY

December 2014

ABSTRACT

Obtaining high-quality experimental designs to optimize statistical efficiency and data quality is quite challenging for functional magnetic resonance imaging (fMRI). The primary fMRI design issue is on the selection of the best sequence of stimuli based on a statistically meaningful optimality criterion. Some previous studies have provided some guidance and powerful computational tools for obtaining good fMRI designs. However, these results are mainly for basic experimental settings with simple statistical models. In this work, a type of modern fMRI experiments is considered, in which the design matrix of the statistical model depends not only on the selected design, but also on the experimental subject's probabilistic behavior during the experiment. The design matrix is thus uncertain at the design stage, making it difficult to select good designs. By taking this uncertainty into account, a very efficient approach for obtaining high-quality fMRI designs is developed in this study. The proposed approach is built upon an analytical result, and an efficient computer algorithm. It is shown through case studies that the proposed approach can outperform an existing method in terms of computing time, and the quality of the obtained designs.

DEDICATION

To Mother and Father

ACKNOWLEDGEMENTS

I would like to sincerely thank my committee members. A special thanks to my supervisor, Dr. Ming-hung Kao, for his guidance, support and encouragement and most of all patience throughout the study. Thank you Dr. Mark Reiser, Dr. John Stufken and Dr. Bruno Welfert for serving on my committee.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

Functional neuroimaging experiments utilizing the pioneering functional magnetic resonance imaging (fMRI) technology help to provide insights into the way that the brain works. Such experiments are widely conducted in various research fields such as psychology, neuroscience, and education for studying brain functions in response to some mental stimuli such as pictures or sounds; see also Lindquist (2008). Researchers also use fMRI as one of the powerful tools for studying some diseases related to brain functions as highlighted in a special issue on clinical applications of fMRI in Neuropsychology Review, Vol. 17 et al. (2007). The use of fMRI is arguably an important advance in neuroscience, and it has many practical applications.

In an fMRI experiment, it is not uncommon that an experimental subject is presented with a sequence of stimuli of one or more types, possibly interlaced with rest periods. For such studies, an fMRI design determines the onset times of each stimulus. For example, a stimulus can be a 1-second picture of smiling face that appears at multiple time points in the experiment. During the periods when there is no stimulus presentation, the subject is asked to rest or gaze at a visual fixation. Each stimulus may evoke neuronal activity at some regions of the subject's brain. This neuronal activity leads to a rise or fall in the concentration of the oxygenated blood in the cerebral vessels. This leads to a change in the strength of the local magnetic field around the activated brain regions. The fluctuation in the strength of the magnetic field is picked up by an fMRI scanner. Specifically, the fMRI scanner collects the blood oxygenation level dependent (BOLD) measurements at regular time points from each of the, say, $64 \times 64 \times 30$ brain voxels (3D imaging units). We will thus have an fMRI

time series from each brain voxel that reflects the change in the strength of the local magnetic field. These time series are analyzed to make statistical inference about the brain activity at the corresponding brain voxels. Such an inference is typically made based on (some characteristics of) the hemodynamic response function (HRF), a function of time modeling the change in the strength of the magnetic field following a stimulus onset (see Chapter 2 for further details).

In traditional fMRI studies, the design matrix of the statistical model for analyzing fMRI data normally can be completely determined by the selected design. However, this no longer holds true for some modern experiments that aim at investigating the brain activity evoked by the subject reactions to the stimuli (e.g., the subject answers to the presented questions). Cordes *et al.* (2012) reported an experiment of this sort. For such experiments, the design matrix will depend not only on the selected design, but also on the subject's reaction to each of these stimuli. As the subject's reactions are uncertain at the design stage (before the experiment starts), it is unlikely to have an accurate evaluation of the quality of designs. This makes it very challenging to select good designs suited to this type of modern experiments.

The goal of this project is to develop an efficient and effective approach for finding high-quality designs to improve the quality of fMRI experiments when the design matrix is uncertain. Our target is at fMRI designs that are robust to possible reactions of the subject. To that end, we build our design selection approach on an analytical result and an efficient computer algorithm. Specifically, we derive an analytical form for our design selection criterion that allows us to evaluate the robustness of designs without much computational effort. We then adapt the genetic algorithm considered by Kao *et al.* (2009) to search for an fMRI design optimizing this criterion. As demonstrated in our case studies, our criterion can serve as a cheap, but good surrogate of the design selection criterion proposed by Cordes *et al.* (2012). For these cases where

the latter criterion is considered, our designs can perform similarly to or slightly better than the designs obtained by using the method proposed by Cordes *et al.* (2012). More importantly, our approach is much faster than that of Cordes *et al.* (2012).

The following chapters are organized as follows. In Chapter 2, we provide some background information about fMRI designs and a motivation example. We then present our methodology in Chapter 3. Chapter 4 illustrates some case studies to demonstrate the usefulness of our approach. A discussion can be found in Chapter 5.

Chapter 2

BACKGROUND INFORMATION AND A MOTIVATION EXAMPLE

2.1 Background Information

In an (event-related) fMRI study, there might be tens or hundreds brief stimuli of one or more types presented to the subject at different time points. Each stimulus may last several milliseconds or a few seconds, immediately followed by a period of 'control' such as a rest period or a presentation of a visual fixation. For example, an experiment might involve 1-second pictures of familiar faces, which form the first stimulus type, and 1-second pictures of unfamiliar faces that form the second stimulus type. Each stimulus can possibly appear every τ_{ISI} seconds, where τ_{ISI} is a pre-specified time (e.g., 4 seconds), and is sometimes termed as the inter-stimulus interval. During the period from the offset of a stimulus to the onset of the next one, the subject is exposed to the control (e.g., rest or visual fixation). An experiment can have a duration of several minutes (e.g., 10 minutes). A design for such an experiment is often represented as an ordered sequence of N elements; i.e. $\boldsymbol{d} = (d_1, ..., d_N)$, where N is typically tens or hundreds. With Q stimulus types, each element d_n in a design can take a value from $\{0, 1, ..., Q\}$. For example, a design with $Q = 2$ may look like $\boldsymbol{d} = (1, 0, 2, 1, ..., 0)$. The n^{th} position of $\boldsymbol{d}$ corresponds to time $(n-1)\tau_{ISI}$, $n = 1, ..., N$. Time 0 is typically set to the time point when the first valid MR measurement is acquired by an fMRI scanner. When $d_n = q > 0$, there is an onset of the q^{th}-type stimulus at the n^{th} time point. With $d_n = 0$, there is no stimulus onset at that time point. For example, a '1' in $\boldsymbol{d}$ may indicate an appearance of a familiar face, a '2' is for an unfamiliar face, and a '0' means that none of these pictures occurs.

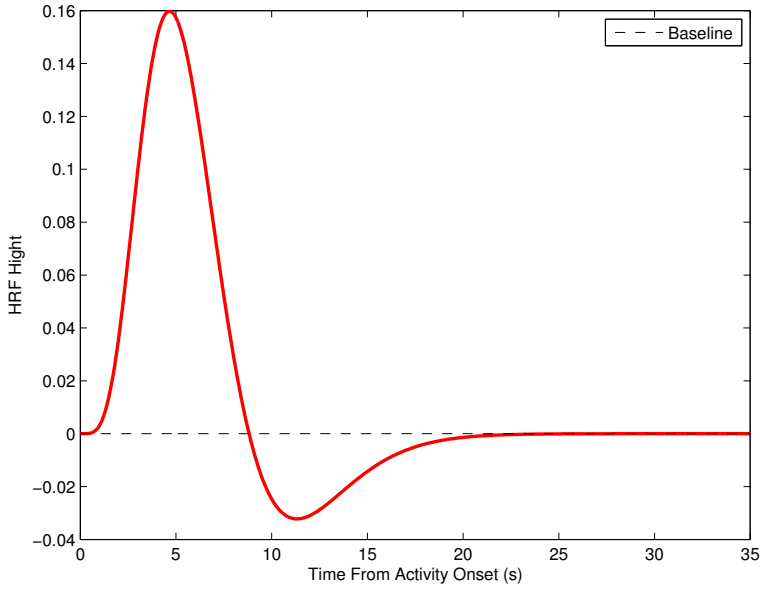

Figure 2.1: An example Hemodynamic Response Function, HRF.

An fMRI design described previously determines the onset times and the presentation order of the stimuli. When a selected design is presented to the subject, an fMRI scanner repeatedly scans the subject's brain to collect data for making statistical inference about the brain activity evoked by the stimuli. In particular, there will be, say, $64 \times 64 \times 30$ three-dimensional image elements, called voxels, each having a size of, say, $3 \times 3 \times 5\ mm^3$. The fMRI scanner repeatedly scans through these voxels. From each voxel, an MR measurement is collected every τ_{TR} seconds (e.g., $\tau_{TR} = 2$ seconds) to form an fMRI time series. These collected time series are then analyzed for making inference about the brain activity in response to the stimuli. Such an inference is typically made via studying the hemodynamic response function, HRF, which is a function of time modeling the stimulus-induced change in the concentrations of the oxy- and deoxy-blood in the cerebral blood vessels at a voxel. The HRF typically has a long duration (e.g., 30 seconds) relative to τ_{ISI} and τ_{TR}. It may look like the curve presented in **Figure** 2.1. It is commonly assumed that, at each voxel,

the stimuli of the same type have the same HRF, whereas stimuli of different types may have different HRFs. Consequently, there will be Q possibly different HRFs in cases with Q stimulus types. An study objective of interest is to estimate these Q HRFs.

A commonly considered linear model for the estimation of the HRFs is (Dale, 1999; Liu and Frank, 2004):

$$\boldsymbol{y} = \sum_{q=1}^{Q} \boldsymbol{X}_{d,q}\boldsymbol{h}_q + \boldsymbol{S}\boldsymbol{\gamma} + \boldsymbol{\epsilon}. \tag{2.1}$$

Here, $\boldsymbol{y} = (y_1, ..., y_T)'$ is the MR measurements obtained from a voxel every τ_{TR} seconds. The unknown parameter vector $\boldsymbol{h}_q = (h_{q,1}, ..., h_{q,K})'$ represents K heights of the HRF of the q^{th}-type stimulus that can contribute to $\boldsymbol{y}$, where the pre-specified integer K is sufficiently large so that the height of the HRF is negligible after $h_{q,K}$. The T-by-K matrix $\boldsymbol{X}_{d,q}$ is the 0-1 design matrix for the q^{th}-type stimulus. $\boldsymbol{S}\boldsymbol{\gamma}$ is a nuisance term modeling the possible drift/trend of $\boldsymbol{y}$ with $\boldsymbol{S}$ being a specified matrix and $\boldsymbol{\gamma}$ being the corresponding parameter vector. The vector $\boldsymbol{\epsilon}$ consists of the T correlated error terms. We note that $\boldsymbol{X}_{d,q}$ is determined by the selected design $\boldsymbol{d}$ as presented in the following example. In that example, we only discuss a case where $\tau_{ISI} = \tau_{TR}$ and $N = T$. For simplicity, we lie the focus of the current study on such cases, but note that a construction of $\boldsymbol{X}_{q,K}$ for cases where $\tau_{ISI} \neq \tau_{TR}$ can be found in Kao *et al.* (2013).

Example 1. Let the design with two types ($Q = 2$) be $\boldsymbol{d} = \{12102\ldots0\}$. To construct $\boldsymbol{X}_{d,1}$ and $\boldsymbol{X}_{d,2}$, we first construct two indicator vectors $\boldsymbol{\delta}_1$ and $\boldsymbol{\delta}_2$. The lengths of these indicator vectors are the same as the length of $\boldsymbol{d}$. The j^{th} element of $\boldsymbol{\delta}_q$ is $\delta_{q,j} = 1$ when $d_j = q$, where d_j denotes the j^{th} entry of the design. Thus we have $\boldsymbol{\delta}_1 = \{10100\ldots0\}$ and $\boldsymbol{\delta}_2 = \{01001\ldots0\}$ respectively. Since $\tau_{ISI} = \tau_{TR}$, $\boldsymbol{X}_{d,1}$

and $\boldsymbol{X}_{d,2}$ are,

$$\boldsymbol{X}_{d,1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \end{pmatrix}; \boldsymbol{X}_{d,2} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \end{pmatrix}.$$

For the q^{th}-type stimulus, the matrix $\boldsymbol{X}_{d,q}$ is a T-by-K Toeplitz matrix where the first column is $\boldsymbol{\delta}_q$ and the first row is defined as $[(\boldsymbol{\delta}_q)_1, 0, \ldots, 0]$.

□

An important design goal is to select an fMRI design $\boldsymbol{d}$ that yields the most precise generalized least square estimate(GLSE) of $\boldsymbol{h} = (\boldsymbol{h}_1', ..., \boldsymbol{h}_Q')'$. This often is to find a $\boldsymbol{d}$ that optimizes some statistically meaningful function of the information matrix of $\boldsymbol{h}$. With Model (2.1), the information matrix of $\boldsymbol{h}$ with a given design $\boldsymbol{d}$ can be written as:

$$\boldsymbol{M}_d = \boldsymbol{X}_d' \boldsymbol{V}' (\boldsymbol{I}_T - \boldsymbol{P}_{VS}) \boldsymbol{V} \boldsymbol{X}_d,$$

where $\boldsymbol{X}_d = [\boldsymbol{X}_{d,1}, ..., \boldsymbol{X}_{d,Q}]$, $\boldsymbol{V}$ is a whitening matrix that such that $cov(\boldsymbol{V}\boldsymbol{\epsilon}) = \sigma^2 \boldsymbol{I}_T$, $\sigma^2 > 0$ is the error variance, $\boldsymbol{I}_T$ is the identify matrix of order T, and $\boldsymbol{P}_{VS} = \boldsymbol{V}\boldsymbol{S}(\boldsymbol{S}'\boldsymbol{V}'\boldsymbol{V}\boldsymbol{S})^{-}\boldsymbol{S}'\boldsymbol{V}'$ is the orthogonal projection onto the space spanned by the columns of $\boldsymbol{V}\boldsymbol{S}$ with $\boldsymbol{A}^{-}$ denoting a generalized inverse of a matrix $\boldsymbol{A}$. It is noteworthy that $\boldsymbol{M}_d$ is inversely proportional to the covariance of the GLSE of $\boldsymbol{h}$, and it depends on the design $\boldsymbol{d}$ through the design matrix $\boldsymbol{X}_d$. We would like a $\boldsymbol{d}$ that 'maximizes' the information matrix $\boldsymbol{M}_d$. However, this goal is often not achievable. A common

strategy is thus to find a $\boldsymbol{d}$ that optimizes $\phi(\boldsymbol{M}_d)$ for some real function ϕ; see Chapter 3 for a further discussion on popularly used ϕ in fMRI. With a selected ϕ, one may utilize a computer algorithm such as the genetic algorithm (GA) of Kao *et al.* (2009) to search for an optimal $\boldsymbol{d}$; see the Appendix for a description about this genetic algorithm.

The previous described design selection method works for Model (2.1), and some other models where the design matrix $\boldsymbol{X}$ can be determined once a design $\boldsymbol{d}$, and the relevant design parameters, including τ_{ISI}, τ_{TR}, K, and T, are selected. However, it does not help to construct designs for some modern experiments where $\boldsymbol{X}$ is uncertain at the design stage. We describe an example of such an experiment in the next section.

2.2 A Motivation Example

The design selection method described in the previous section is for traditional settings of fMRI studies, where the design matrix $\boldsymbol{X}_d$ can be determined at the design stage. Here, our focus is on a modern experimental setting, where the design matrix $\boldsymbol{X}_d$ in the model is uncertain. A study of this sort is reported by Cordes *et al.* (2012). Specifically, Cordes and coauthors reported an fMRI experiment for studying brain activity evoked by the subject's reactions to the pictures (stimuli) presented to her/him during the experiment. In the pilot study that they conducted, a list of pictures was presented asymmetrically about the vertical axis for the subjects to study. These subjects are then enrolled in an fMRI experiment, in which some of the pictures that the subjects have studied were presented with the same or opposite left/right orientation, interlaced with some other new pictures. Consequently, there are three stimulus types, namely

1. studied pictures with the 'same' orientation;

2. studied pictures with the 'different' orientation; and

3. 'new' pictures.

During the fMRI scanning session, each picture was presented to the subject for 3 seconds. The subjects were asked to select an answer among 'same', 'different' and 'new' for each picture. The main objective is to study the brain activity with respect to the stimulus-response pairs. As described in the next chapter, the model parameters of interest are linked to both the stimuli, and the subject's answers, which are unknown at the design stage. Consequently, the design matrix, the information matrix, and the value of the optimality criterion $\phi(\boldsymbol{M}_d)$ for any given $\boldsymbol{d}$ are uncertain since they all depend on the subject's answers. This makes it vary challenging to obtain a good designs for such experiments.Cordes *et al.* (2012) proposed a method for tackle this challenging design issue. Briefly speaking, they first approximate the probabilities of the subject's answers to each type of pictures. For each candidate design $\boldsymbol{d}$, they then simulate, say, 100 realizations of the subject's answers to obtain 100 realizations of the design matrix, and thus, 100 realizations of the ϕ-values for $\boldsymbol{d}$. A summary statistic such as the median or mean of these 100 ϕ-values is obtained to evaluate the goodness of $\boldsymbol{d}$. Conceptually, this is similar to use $E\{\phi(\boldsymbol{M}_d)\}$ to evaluate the goodness of $\boldsymbol{d}$, where the expectation $E\{.\}$ is taken over the probability of the subject's answer to each given stimulus type. Since a closed form of $E\{\phi(\boldsymbol{M}_d)\}$ is in general unavailable, an Monte Carlo simulation is considered to approximate this criterion. Unfortunately, the Monte Carlo simulation is time consuming, and it needs to be repeated for every candidate design. The procedure thus requires much computational effort as reported in Cordes *et al.* (2012). An efficient approach is called for.

Chapter 3

METHODOLOGY

3.1 The General Linear Model

For addressing the design issue described in Section 2.2, we consider an extension of Model (2.1). For simplicity, we assume that the subject can have R possible responses for every stimulus. In the motivating example, $R = 3$. Our approach can be easily extended to the case where the number of possible responses can vary across stimulus types. With Q stimulus types, the model that we consider is:

$$\boldsymbol{y} = \sum_{r=1}^{R}\sum_{q=1}^{Q}\sum_{k=1}^{K} \boldsymbol{x}_{r,q,k} h_{r,q,k} + \boldsymbol{S\gamma} + \boldsymbol{\epsilon} = \boldsymbol{X}_d \boldsymbol{h} + \boldsymbol{S\gamma} + \boldsymbol{\epsilon} \tag{3.1}$$

Here, $\boldsymbol{y}$ is defined as in Model (2.1). $\boldsymbol{X}_d$ is the design matrix whose columns are the 0-1 vectors $\boldsymbol{x}_{r,q,k}$'s, The $\boldsymbol{x}_{r,q,k}$ vector indicates the contribution of $h_{r,q,k}$ to $\boldsymbol{y}$. In particular, its n^{th} element $(\boldsymbol{x}_{r,q,k})_n = 1$ if $h_{r,q,k}$ contributes to y_n, the n^{th} MR measurement. Each parameter $h_{r,q,k}$ represents the k^{th} height of the (discretized) HRF evoked by the event that the subject selected the r^{th} answer to a stimulus of the q^{th} type, $k = 1, ..., K,\ r = 1, ..., R,\ q = 1, ..., Q$. The discretization interval for discretizing each HRF is ΔT, and is the greatest real value that makes both $(\tau_{ISI}/\Delta T)$ and $(\tau_{TR}/\Delta T)$ integers; see Kao *et al.* (2009). With this ΔT, $h_{r,q,k}$ is the HRF height evaluated at $(k-1)\Delta T$ seconds after an onset of the corresponding event, where $k = 1, \ldots, K$, and $K = \lfloor 1 + (32/\Delta T) \rfloor$ for a 32-second HRF; $\lfloor a \rfloor$ is the integer part of a. The remaining terms in Model (3.1) are as in Model (2.1).

3.2 Design Criterion

With Model (3.1), we would like to find a design that helps to render the most precise GLSE $\hat{\boldsymbol{h}}$ of the HRFs, $\boldsymbol{h}$. The goodness of a design will be evaluated by some optimality criterion ϕ of the information matrix. The information matrix is:

$$\boldsymbol{M}_d = \boldsymbol{X}_d'\boldsymbol{V}'(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{VS}})\boldsymbol{V}\boldsymbol{X}_d = \boldsymbol{X}_d'[\boldsymbol{V}'\boldsymbol{V} - \boldsymbol{V}'\boldsymbol{V}\boldsymbol{S}(\boldsymbol{S}'\boldsymbol{V}'\boldsymbol{V}\boldsymbol{S})^{-}\boldsymbol{S}'\boldsymbol{V}'\boldsymbol{V}]\boldsymbol{X}_d. \quad (3.2)$$

It is not uncommon to assume that the error term $\boldsymbol{\epsilon}$ follows a stationary first-order autoregressive (AR1) process. Under this assumption, the $\boldsymbol{V}'\boldsymbol{V}$ that is needed for calculating $\boldsymbol{M}_d$ is

$$\begin{pmatrix} 1 & -\rho & 0 & \cdots & \cdots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & \vdots \\ 0 & -\rho & 1+\rho^2 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & \cdots & 0 & -\rho & 1 \end{pmatrix}$$

for some $\rho \in (-1, 1)$. Other correlation structures for $\boldsymbol{\epsilon}$ may also be considered.

For the optimality criterion ϕ, we will consider the A- and D-optimality criteria since they are common in fMRI design studies (Dale, 1999; Wager and Nichols, 2003; Liu and Frank, 2004; Kao *et al.*, 2009; Maus *et al.*, 2010). Extending our method to other optimality criteria should be straightforward. For the A-optimality criterion, we write:

$$\phi\left(\boldsymbol{M}_d\right) = \frac{RQK}{trace\left[\boldsymbol{M}_d^{-1}\right]},$$

The A-value is set to 0 when $\boldsymbol{M}_d$ is singular. In such a case, $\boldsymbol{h}$ is non-estimable. As for the D-optimality criterion, we have

$$\phi\left(\boldsymbol{M}_d\right) = det(\boldsymbol{M}_d)^{-\frac{1}{RQK}}.$$

Unfortunately, the design matrix $\boldsymbol{X}$ in Model (3.1), and thus the information matrix $\boldsymbol{M}_d$, will depend on the subject's reaction during the experiment. Consequently, $\phi(\boldsymbol{M}_d)$ is unavailable at the design stage. To tackle this issue, we propose to consider the expectation $E(\boldsymbol{M}_d)$ of the information matrix $\boldsymbol{M}_d$. The expectation is taken over the probability of the subject's answer to the presented stimulus. This probability can be approximated from, say, a pilot study. We then obtain a 'robust' design that maximizes $\phi_1(\boldsymbol{d}) \equiv \phi(E(\boldsymbol{M}_d))$, where ϕ can be set to the A- or D-optimality criterion. Another possibility is by considering $\phi_2(\boldsymbol{d}) \equiv E\{\phi(\boldsymbol{M}_d)\}$. Both criteria, which are sometimes viewed as the Bayesian versions of the optimality criteria, have been considered in the design literature; see Ch.18 of Atkinson *et al.* (2007).

As described in Section 2.2, the design selection criterion considered by Cordes *et al.* (2012) is linked to $\phi_2(\boldsymbol{d})$. A major disadvantage for considering this criterion is that the ϕ_2-value is in general unavailable for an fMRI design $\boldsymbol{d}$. For evaluating the goodness of $\boldsymbol{d}$'s using ϕ_2, we may follow Cordes *et al.* (2012) to conduct a Monte Carlo simulation to generate m, say 100, realizations of $\phi(\boldsymbol{M}_d)$, and then approximate $\phi_2(\boldsymbol{d})$ by a summary statistic such as the mean/median of the m realizations of the ϕ-value. This process unfortunately is computationally very expensive. By contrast, our proposed criterion value $\phi_1(\boldsymbol{d})$ is very easy to compute. This is because a closed form of $E(\boldsymbol{M}_d)$ can be analytically derived (see the next Section). For comparison purposes, we consider not only $\phi_1(\boldsymbol{d})$, but also $\phi_2(\boldsymbol{d})$ for design evaluations; the ϕ_2-value of each $\boldsymbol{d}$ will be approximated by the mean of m realizations of $\phi(\boldsymbol{M}_d)$. Hereinafter, this approximation of $\phi_2(\boldsymbol{d})$ is denoted by $\phi_2(\boldsymbol{d}; m)$. With a selected criterion ($\phi_1(\boldsymbol{d})$ or $\phi_2(\boldsymbol{d}; m)$), we then adapt the genetic algorithm of Kao *et al.* (2009) to search for an $\boldsymbol{d}$ that optimize the criterion. We show, through case studies, that designs optimizing $\phi_1(\boldsymbol{d})$ also tend to perform very well with respect to $\phi_2(\boldsymbol{d}; m)$. While $\phi_1(\boldsymbol{d})$ is of interest on its own right, this criterion is also demonstrated to be a

cheap, but good, surrogate for $\phi_2(\boldsymbol{d})$.

In the next section, we derive a closed form for $E(\boldsymbol{M}_d)$. The genetic algorithm used to search for designs optimizing a selected criterion is described in the Appendix.

3.3 Formulations

Without loss of generality, we assume that $\boldsymbol{X}_d$ in (3.2) has the form of $\boldsymbol{X} = [\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K]$, where for $k = 1, ..., K$, and $q = 1, ..., Q$, $\boldsymbol{X}_k = [\boldsymbol{X}_{1,k}, \ldots, \boldsymbol{X}_{Q,k}]$ with dimension $N \times RQ$, and $\boldsymbol{X}_{q,k} = [\boldsymbol{x}_{1,q,k}, \ldots, \boldsymbol{x}_{R,q,k}]$, and $\boldsymbol{x}_{1,q,k}$ is defined as in Model (3.1). Since most optimality criteria ϕ are invariant to a simultaneous permutation of rows and columns, a re-arrangement of the columns of $\boldsymbol{X}_d$ will not change the value of $\phi(E\{\boldsymbol{M}_d\})$. We have also derived a closed form for $E\{\boldsymbol{M}_d\}$ by setting $\boldsymbol{X}_d = [\boldsymbol{X}_{d,1}, ..., \boldsymbol{X}_{d,Q}]$, where $\boldsymbol{X}_{d,q}$ contains all the QK vectors $\boldsymbol{x}_{r,q,k}$ of the same q. We omit this latter result because, comparing with the former arrangement of $\boldsymbol{X}_d$, it tends to take more CPU time to calculate $E\{\boldsymbol{M}_d\}$ when the closed form derived by the latter choice of $\boldsymbol{X}_d$ is considered. For simplicity, the focus of the current study is on cases with $\tau_{ISI} = \tau_{TR}$. We will extend the result to cases where $\tau_{ISI} \neq \tau_{TR}$ in a future study. With $\tau_{ISI} = \tau_{TR}$, $((\boldsymbol{x}_{r,q,1}))_n = 1$ when the n^{th} stimulus is of the q^{th} type, and the subject selected the r^{th} answer for that stimulus. In addition, $\boldsymbol{x}_{r,q,k} = \boldsymbol{L}^{k-1}\boldsymbol{x}_{r,q,1}$, where $\mathbf{L}$ is an N-by-N matrix with

$$\boldsymbol{L} = \begin{pmatrix} \mathbf{0}^T & 0 \\ \mathbf{I}_{N-1} & \mathbf{0} \end{pmatrix},$$

where $\boldsymbol{I}_a$ is the a-by-a identity matrix. We now derive the expectation of the information matrix $E\{\boldsymbol{M}_d\}$. The expectation is taken over $p(r \mid q)$, the conditional probability when subject selects the r^{th} answer for a stimulus of the q^{th} type. Here, we assume that $p(r \mid q)$ remains the same throughout the experiment, and the subject's answer only depends on the current stimulus, and is independent of his/her answers

to the previous stimuli; we also assume that, for each stimulus, the subject selects one answer from the R possible answers, and if there is no stimulus, the subject does not respond, and $P(r = 0 \mid q = 0) = 1$. Our results can easily be extended to a more general case such as $P(r \mid 0) > 1$ for $r = 1, \ldots, R$, and/or $P(0 \mid q) > 1$ for $q = 1, \ldots, Q$. For convenience, we also use $\boldsymbol{A}$ to represent $\boldsymbol{V}'(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{VS}})\boldsymbol{V}$. Consequently, $\boldsymbol{M}_d = \boldsymbol{X}'\boldsymbol{A}\boldsymbol{X}$. The main idea is then to make use of the formula for the expectation of a quadratic form as presented in Ch.5 of Rencher and Schaalje (2008). We now present some details of our derivations of $E\{\boldsymbol{M}_d\}$. We note that all the expectations (and covariances) are conditional on the design $\boldsymbol{d}$. For simplicity, we write $E\{\cdot\}$ (and $cov\{\cdot\}$) instead of $E\{\cdot \mid \boldsymbol{d}\}$ (and $cov\{\cdot \mid \boldsymbol{d}\}$).

First, the expectation of $\boldsymbol{M}_d$ can be written as

$$\begin{aligned} E\{\boldsymbol{M}_d\} &= (E(\boldsymbol{X}_i'\boldsymbol{A}\boldsymbol{X}_j))_{i,j=1,\ldots,K} \\ &= (E([\boldsymbol{L}^{i-1}\boldsymbol{X}_1]'\boldsymbol{A}[\boldsymbol{L}^{j-1}\boldsymbol{X}_1]))_{i,j=1,\ldots,K} \\ &= (E(\boldsymbol{X}_1'(\boldsymbol{L}^{i-1})'\boldsymbol{A}\boldsymbol{L}^{j-1}\boldsymbol{X}_1))_{i,j=1,\ldots,K}. \end{aligned}$$

Let $\boldsymbol{A}_{i,j} = (\boldsymbol{L}^{i-1})'\boldsymbol{A}\boldsymbol{L}^{j-1}$. We then have

$$E\{\boldsymbol{M}_d\} = (E(\boldsymbol{X}_1'\boldsymbol{A}_{i,j}\boldsymbol{X}_1))_{i,j=1,\ldots,K}.$$

Note that

$$E(\boldsymbol{X}_1'\boldsymbol{A}_{i,j}\boldsymbol{X}_1) = (E(\boldsymbol{X}_{p,1}'\boldsymbol{A}_{i,j}\boldsymbol{X}_{q,1}))_{p,q=1,\ldots,Q}.$$

Here, $\boldsymbol{X}_{p,1}'\boldsymbol{A}_{i,j}\boldsymbol{X}_{q,1}$ can be written as:

$$\begin{pmatrix} \boldsymbol{x}_{1,p,1}'\boldsymbol{A}_{i,j}\boldsymbol{x}_{1,q,1} & \boldsymbol{x}_{1,p,1}'\boldsymbol{A}_{i,j}\boldsymbol{x}_{2,q,1} & \cdots & \boldsymbol{x}_{1,p,1}'\boldsymbol{A}_{i,j}\boldsymbol{x}_{R,q,1} \\ \boldsymbol{x}_{2,p,1}'\boldsymbol{A}_{i,j}\boldsymbol{x}_{1,q,1} & \boldsymbol{x}_{2,p,1}'\boldsymbol{A}_{i,j}\boldsymbol{x}_{2,q,1} & \cdots & \boldsymbol{x}_{2,p,1}'\boldsymbol{A}_{i,j}\boldsymbol{x}_{R,q,1} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{x}_{R,p,1}'\boldsymbol{A}_{i,j}\boldsymbol{x}_{1,q,1} & \boldsymbol{x}_{R,p,1}'\boldsymbol{A}_{i,j}\boldsymbol{x}_{2,q,1} & \cdots & \boldsymbol{x}_{r,p,1}'\boldsymbol{A}_{i,j}\boldsymbol{x}_{R,q,1} \end{pmatrix}_{R\times R}.$$

We now present the expectation of quadratic forms $\boldsymbol{x}'_{u,p,1}\boldsymbol{A}_{i,j}\boldsymbol{x}_{r,q,1}$ in the previously described matrix. First, we define $\boldsymbol{\delta}_q$ as the 0-1 indicator vector for the q^{th}-type stimulus. Specifically, the n^{th} element $(\boldsymbol{\delta}_q)_n$ of $\boldsymbol{\delta}_q$ is 1 if the corresponding d_n in the design $\boldsymbol{d} = (d_1, \ldots, d_N)$ is q (the q^{th} stimulus type); otherwise, $(\boldsymbol{\delta}_q)_n = 0$. We have the following results. There, $diag(\boldsymbol{\delta}_q)$ indicates the diagonal matrix whose diagonal elements are the elements of $\boldsymbol{\delta}_q$.

1. For $p = q$:

Case 1: $u = r$,

$$E(\boldsymbol{x}_{r,q,1}) = E(\boldsymbol{x}_{r,p,1}) = E(\boldsymbol{x}_{u,q,1}) = E(\boldsymbol{x}_{u,p,1}) = p(r|q)\boldsymbol{\delta}_q;$$

$$cov(\boldsymbol{x}_{u,p,1}, \boldsymbol{x}_{r,q,1}) = cov(\boldsymbol{x}_{r,q,1}, \boldsymbol{x}_{r,q,1}) = p(r|q)(1 - p(r|q))diag(\boldsymbol{\delta}_q);$$

$$\begin{aligned} tr[\boldsymbol{A}_{i,j}cov(\boldsymbol{x}_{u,p,1}, \boldsymbol{x}_{r,q,1})] &= tr[\boldsymbol{A}_{i,j}cov(\boldsymbol{x}_{r,q,1}, \boldsymbol{x}_{r,q,1})] \\ &= tr[\boldsymbol{A}_{i,j}p(r|q)(1 - p(r|q))diag(\boldsymbol{\delta}_q)] \\ &= p(r|q)(1 - p(r|q))tr[\boldsymbol{A}_{i,j}diag(\boldsymbol{\delta}_q)] \\ &= (p(r|q) - p(r|q)p(r|q))tr[\boldsymbol{A}_{i,j}diag(\boldsymbol{\delta}_q)]; \end{aligned}$$

$$\begin{aligned} &E(\boldsymbol{x}'_{u,p,1}\boldsymbol{A}_{i,j}\boldsymbol{x}_{r,q,1}) \\ &= tr[\boldsymbol{A}_{i,j}cov(\boldsymbol{x}_{u,p,1}, \boldsymbol{x}_{r,q,1})] + E(\boldsymbol{x}'_{u,p,1})\boldsymbol{A}_{i,j}E(\boldsymbol{x}_{r,q,1}) \\ &= tr[\boldsymbol{A}_{i,j}cov(\boldsymbol{x}_{r,q,1}, \boldsymbol{x}_{r,q,1})] + E(\boldsymbol{x}'_{r,q,1})\boldsymbol{A}_{i,j}E(\boldsymbol{x}_{r,q,1}) \\ &= (p(r|q) - p(r|q)^2)tr[\boldsymbol{A}_{i,j}diag(\boldsymbol{\delta}_q)] + p(r|q)^2\boldsymbol{\delta}_q^{'}\boldsymbol{A}_{i,j}\boldsymbol{\delta}_q. \end{aligned}$$

Case 2: $u \neq r$,

$$\begin{cases} E(\boldsymbol{x}_{u,q,1}) = p(u|q)\boldsymbol{\delta}_q \\ E(\boldsymbol{x}_{r,q,1}) = p(r|q)\boldsymbol{\delta}_q \end{cases};$$

$$cov(\boldsymbol{x}_{u,p,1}, \boldsymbol{x}_{r,q,1}) = cov(\boldsymbol{x}_{u,q,1}, \boldsymbol{x}_{r,q,1}) = -p(u|q)p(r|q)diag(\boldsymbol{\delta}_q);$$

$$
\begin{aligned}
tr[\boldsymbol{A}_{i,j}cov(\boldsymbol{x}_{u,p,1},\boldsymbol{x}_{r,q,1})] &= tr[\boldsymbol{A}_{i,j}cov(\boldsymbol{x}_{u,q,1},\boldsymbol{x}_{r,q,1})] \\
&= tr[-p(u|q)p(r|q)diag(\boldsymbol{\delta}_q)] \\
&= -p(u|q)p(r|q)tr[\boldsymbol{A}_{i,j}diag(\boldsymbol{\delta}_q)] \\
&= -p(u|q)p(r|q)tr[\boldsymbol{A}_{i,j}diag(\boldsymbol{\delta}_q)];
\end{aligned}
$$

$$
\begin{aligned}
&E(\boldsymbol{x}'_{u,p,1}\boldsymbol{A}_{i,j}\boldsymbol{x}_{r,q,1}) \\
&= tr[\boldsymbol{A}_{i,j}cov(\boldsymbol{x}_{u,p,1},\boldsymbol{x}_{r,q,1})] + E(\boldsymbol{x}'_{u,p,1})\boldsymbol{A}_{i,j}E(\boldsymbol{x}_{r,q,1}) \\
&= tr[\boldsymbol{A}_{i,j}cov(\boldsymbol{x}_{u,q,1},\boldsymbol{x}_{r,q,1})] + E(\boldsymbol{x}'_{u,q,1})\boldsymbol{A}_{i,j}E(\boldsymbol{x}_{r,q,1}) \\
&= -p(u|q)p(r|q)tr[\boldsymbol{A}_{i,j}diag(\boldsymbol{\delta}_q)] + p(r|q)p(u|q)\boldsymbol{\delta}'_q\boldsymbol{A}_{i,j}\boldsymbol{\delta}_q.
\end{aligned}
$$

We combine **Case 1** and **Case 2** to obtain the followings.

$$
\begin{aligned}
&E(\boldsymbol{X}'_{p,1}\boldsymbol{A}_{i,j}\boldsymbol{X}_{q,1}) \\
&= tr[\boldsymbol{A}_{i,j}diag(\boldsymbol{\delta}_q)][diag(\vec{P}(q)) - \vec{P}(q)\vec{P}'(q)] + \vec{P}(q)\vec{P}'(q)\boldsymbol{\delta}'_q\boldsymbol{A}_{i,j}\boldsymbol{\delta}_q,
\end{aligned}
$$

where $\vec{P}(q)$ is an $R \times 1$ vector of $p(r|q)$, $r = 1, \ldots, R$ for $q = 1, \ldots, Q$, i.e., if q=1, R=2, we have $\vec{P}(1) = [p(1 \mid 1), p(2 \mid 1)]'$.

2. **For $p \neq q$:**

$$
\begin{cases}
E(\boldsymbol{x}_{u,p,1}) = p(u|p)\boldsymbol{\delta}_p \\
E(\boldsymbol{x}_{r,q,1}) = p(r|q)\boldsymbol{\delta}_q
\end{cases};
$$

$$
cov(\boldsymbol{x}_{u,p,1},\boldsymbol{x}_{r,q,1}) = 0;
$$

$$
tr[\boldsymbol{A}cov(\boldsymbol{x}_{u,p,1},\boldsymbol{x}_{r,q,1})] = 0;
$$

$$
\begin{aligned}
&E(\boldsymbol{x}'_{u,p,1}\boldsymbol{A}_{i,j}\boldsymbol{x}_{r,q,1}) \\
&= tr[\boldsymbol{A}_{i,j}cov(\boldsymbol{x}_{u,p,1},\boldsymbol{x}_{r,q,1})] + E(\boldsymbol{x}'_{u,p,1})\boldsymbol{A}_{i,j}E(\boldsymbol{x}_{r,q,1}) \\
&= 0 + p(u|p)\boldsymbol{\delta}'_p\boldsymbol{A}_{i,j}p(r|q)\boldsymbol{\delta}_q \\
&= p(u|p)p(r|q)\boldsymbol{\delta}'_p\boldsymbol{A}_{i,j}\boldsymbol{\delta}_q.
\end{aligned}
$$

Thus,

$$E(\boldsymbol{X}'_{p,1}\boldsymbol{A}_{i,j}\boldsymbol{X}_{q,1}) = \vec{P}(q)\vec{P}'(p)\boldsymbol{\delta}'_p\boldsymbol{A}_{i,j}\boldsymbol{\delta}_q.$$

We thus have the following formula, which can be easily built in a computer program to calculate the elements in $E\{\boldsymbol{M}_d\}$. In particular, the $R \times R$ matrix $E(\boldsymbol{X}'_{p,1}\boldsymbol{A}_{i,j}\boldsymbol{X}_{q,1}) =$

$$\begin{cases} tr[\boldsymbol{A}_{i,j}diag(\boldsymbol{\delta}_q)][diag(\vec{P}(q)) - \vec{P}(q)\vec{P}'(q)] + \vec{P}(q)\vec{P}(q)'\boldsymbol{\delta}'_q\boldsymbol{A}_{i,j}\boldsymbol{\delta}_q, & p = q; \\ \vec{P}(q)\vec{P}'(p)\boldsymbol{\delta}'_p\boldsymbol{A}_{i,j}\boldsymbol{\delta}_q, & p \neq q. \end{cases}$$

We make use of this analytical result to conduct some case studies in the next Chapter to demonstrate the usefulness of our proposed approach.

Chapter 4

CASE STUDIES

In the following case studies, we consider cases with one stimulus type ($Q = 1$), and cases with two stimulus types ($Q = 2$). The length N of a design $\boldsymbol{d} = (d_1, ..., d_N)'$ is 255 for $Q = 1$, and is 242 for $Q = 2$. For each stimulus in $\boldsymbol{d}$, we assume that the subject selects an 'answer' from two possible answers; i.e. $R = 2$. In other words, corresponding to $\boldsymbol{d}$, we have a vector $\boldsymbol{r} = (r_1, ..., r_N)'$ that consists of the answers of the subject; here, $r_n = 0$ if $d_n = 0$ and $r_n = 1$ or 2 when $d_n > 0$. Both the inter-stimulus interval τ_{ISI} and the time to repetition τ_{TR} are set to 2 seconds. Thus, ΔT is equal to 2. The drift of time series, $\boldsymbol{S\gamma}$ is assumed to be a second-order Legendre polynomial. We also assume that the noise follows an stationary AR1 process with a correlation coefficient of 0.3. The duration of HRF is 32 seconds. Consequently, the number of HRF heights is $K = \lfloor 1 + (32/2) \rfloor = 17$. Such a model assumption is not uncommon in the fMRI design literature.

For the conditional probability $P(r \mid q)$ that the subject select the r^{th} answer when there is a q^{th}-type stimulus, we consider the following situations. In all these situations, we set $P(0 \mid 0) = 1$, although this assumption is not essential.

1. For cases with one stimulus type ($Q = 1$), we consider two situations, including (i) $p(1 \mid 1) = p(2 \mid 1)$, i.e., for each stimulus, the subject has equal probability to select any of the two answers; and (ii) $p(1 \mid 1) \neq p(2 \mid 1)$, i.e., the probabilities for selecting two answers are different. In particular,

Equal Probability		Unequal Probability	
$p(1 \mid 1)$	$p(2 \mid 1)$	$p(1 \mid 1)$	$p(2 \mid 1)$
0.5	0.5	0.2	0.8

2. For $Q = 2$, we have the following three situations:

Equal Probability		Unequal Probability (I)		Unequal Probability (II)	
$p(1 \mid 1)$	$p(2 \mid 1)$	$p(1 \mid 1)$	$p(2 \mid 1)$	$p(1 \mid 1)$	$p(2 \mid 1)$
0.5	0.5	0.5	0.5	0.7	0.3
$p(1 \mid 2)$	$p(2 \mid 2)$	$p(1 \mid 2)$	$p(2 \mid 2)$	$p(1 \mid 2)$	$p(2 \mid 2)$
0.5	0.5	0.2	0.8	0.2	0.8

We note that the previously specified $p(r \mid q)$'s are needed for calculating $E(\boldsymbol{M}_d)$ for each $\boldsymbol{d}$ as described in Chapter 3. These conditional probabilities are also used to generate m realizations of $\boldsymbol{M}_d$ for each given $\boldsymbol{d}$. These realizations are then used calculate the approximation $\phi_2(\boldsymbol{d}; m)$ of $\phi_2(\boldsymbol{d})$. In what follows, we will first adapt the genetic algorithm of Kao (2009) to obtain a design, $\boldsymbol{d}_{GA}$, that maximizes $\phi_1(\boldsymbol{d})$. With this optimality criterion, we compare our obtained designs with some traditional designs that are popular in practice (for different purposes). These traditional designs include random designs, m-sequences, and blocked designs. Each element of a random design, $\boldsymbol{d}_{rand}$, is generated from a discrete uniform distribution over $\{0, 1, ..., Q\}$. The m-sequences, $\boldsymbol{d}_{mseq}$, or maximum-length shift register sequences, are introduced into fMRI by Buračas and Boynton (2002). These designs are known to be perform well for estimating the HRF, and can be easily generated by the MATLAB program of Liu and Frank (2004). For blocked design $\boldsymbol{d}_{block}$, we consider designs having a 16-s-off-16-s-on pattern. For example, when $Q = 1$, the first 16 seconds is the off-period, and no stimulus is shown to the subject. In the next 16 seconds, stimuli of the same type is shown to the subject every τ_{ISI} seconds. This is repeated for several cycles until the end of the experiment. In particular, a $\boldsymbol{d}_{block}$ may

look like $\{0000000011111111100000000\cdots 0\}$ when $Q = 1$. For $Q = 2$, $\boldsymbol{d}_{block}$ may be $\{00000000111111112222222200000000\cdots 0\}$. The blocked designs are known to useful for another study objective of fMRI, namely the detection of activated brain voxels. They do not perform well when the focus is on the estimation of the HRF, and may give rise to confounding psychological effects such as subject habituation or anticipation. For all these traditional designs, we compare their ϕ_1-values to that of $\boldsymbol{d}_{GA}$.

In addition, we use the genetic algorithm to obtain a design $\boldsymbol{d}_{r100}$ that maximizes $\phi_2(\boldsymbol{d}; 100)$. The resulting designs is compared with $\boldsymbol{d}_{GA}$ in terms of ϕ_1. To demonstrate that ϕ_1 provides a good surrogate for ϕ_2, we also compare the ϕ_2-values of $\boldsymbol{d}_{r100}$, and $\boldsymbol{d}_{GA}$ as well as the CPU times needed for generating these two types of designs. For this latter comparison, $\phi_2(\boldsymbol{d})$ is approximated by $\phi_2(\boldsymbol{d}; m = 1000)$ even though $\boldsymbol{d}_{r100}$ is obtained with $\phi_2(\boldsymbol{d}; m = 100)$. We note that $\phi_2(\boldsymbol{d}; 1000)$ is expected to have a higher precision than $\phi_2(\boldsymbol{d}; 100)$ for approximating $\phi_2(\boldsymbol{d})$. However, the calculation of $\phi_2(\boldsymbol{d}; 1000)$ is computationally very expensive, and is thus difficult, if not infeasible, to be considered for obtaining $\boldsymbol{d}_{r100}$. For ϕ_1 and ϕ_2, we will consider both A- and D-optimality criteria.

4.1 Design Comparisons in Terms of ϕ_1

We evaluate the designs described above with both A- and D-optimality criteria. The results are presented in **Table** 4.1 and 4.2 for cases with $Q = 1$. **Table** 4.3 and 4.4 provide results for cases with $Q = 2$. We also compute the ratio $\phi_1(\boldsymbol{d})/\phi_1(\boldsymbol{d}_{GA})$ for the traditional fMRI designs and $\boldsymbol{d}_{r100}$ in **Figure** 4.1 and 4.2 for $Q = 1$ and $Q = 2$ respectively. For these results, we generate 100 random designs, and the mean and standard deviation of the ϕ_1-value over these 100 random designs are calculated.

The good performance of $\boldsymbol{d}_{GA}$ is consistently demonstrated in **Tables** 4.1 to 4.4

Table 4.1: The values of ϕ_1 for the different designs that we consider with $Q = 1$ when the A-optimality criterion is considered.

Case	$\boldsymbol{d}_{GA}$	$\boldsymbol{d}_{rand}$ (mean±std)	$\boldsymbol{d}_{Block}$	$\boldsymbol{d}_{mseq}$	$\boldsymbol{d}_{r100}$
Equal Probability	40.88	36.45 ± 0.0032	0.64	39.12	40.86
Unqual Probability	30.90	26.72 ± 0.0027	0.64	28.08	30.42

Table 4.2: The values of ϕ_1 for the different designs that we consider with $Q = 1$ when the D-optimality criterion is considered.

Case	$\boldsymbol{d}_{GA}$	$\boldsymbol{d}_{rand}$ (mean±std)	$\boldsymbol{d}_{Block}$	$\boldsymbol{d}_{mseq}$	$\boldsymbol{d}_{r100}$
Equal Probability	48.88	43.38 ± 0.0034	12.08	45.02	48.80
Unqual Probability	39.11	34.44 ± 0.0027	9.66	35.59	38.81

Table 4.3: The values of ϕ_1 for the different designs that we consider with $Q = 2$ when the A-optimality criterion is considered.

Case	$\boldsymbol{d}_{GA}$	$\boldsymbol{d}_{rand}$ (mean±std)	$\boldsymbol{d}_{Block}$	$\boldsymbol{d}_{mseq}$	$\boldsymbol{d}_{r100}$
Equal Probability	24.90	21.92 ± 0.0016	0	23.51	24.52
Unequal Probability (I)	21.15	18.56 ± 0.0014	0	19.15	20.72
Unequal Probability (II)	19.94	17.66 ± 0.0012	0	17.90	19.61

Table 4.4: The values of ϕ_1 for the different designs that we consider with $Q = 2$ when the D-optimality criterion is considered.

Case	$\boldsymbol{d}_{GA}$	$\boldsymbol{d}_{rand}$ (mean±std)	$\boldsymbol{d}_{Block}$	$\boldsymbol{d}_{mseq}$	$\boldsymbol{d}_{r100}$
Equal Probability	31.97	28.67 ± 0.0018	0	29.29	31.90
Unequal Probability (I)	28.60	25.73 ± 0.0016	0	26.07	28.36
Unequal Probability (II)	27.38	24.50 ± 0.0017	0	25.84	27.02

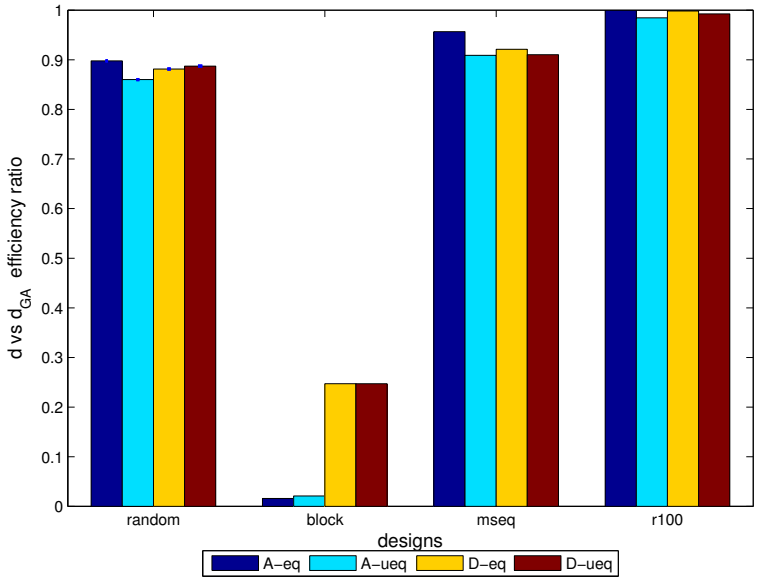

Figure 4.1: Relative Design Efficiencies for $Q = 1$: This plot provides the relative efficiency $\phi_1(\boldsymbol{d})/\phi_1(\boldsymbol{d}_{GA})$ of different designs $\boldsymbol{d}$ with $Q = 1$ for four different cases corresponding to equal $p(r \mid q)$ with A-optimality/D-optimality criterion and unequal $p(r \mid q)$ with A-optimality/D-optimality criterion

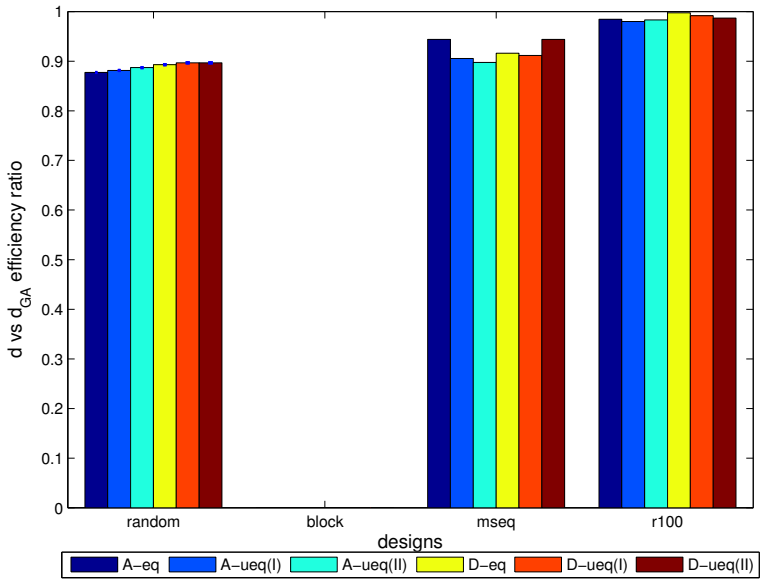

Figure 4.2: Relative Design Efficiencies for $Q = 2$: This plot provides the relative efficiency $\phi_1(\boldsymbol{d})/\phi_1(\boldsymbol{d}_{GA})$ of different designs $\boldsymbol{d}$ with $Q = 2$ for five different cases corresponding to equal $p(r \mid q)$ with A-optimality/D-optimality criterion and the two unequal $p(r \mid q)$ settings with A-optimality/D-optimality criterion

in all the cases that we studied. From **Figures** 4.1 and 4.2, it also is clear that no design that we consider has higher ϕ_1-value than $\boldsymbol{d}_{GA}$ since the ratio $\phi_1(\boldsymbol{d})/\phi_1(\boldsymbol{d}_{GA})$ is less than 1 for any other $\boldsymbol{d}$. It is noteworthy that the blocked designs perform poorly in terms of the ϕ_1 criterion. These designs are not recommended when the study objective lies in the estimation of the HRF.

4.2 Design Comparisons in Terms of $\phi_2(\boldsymbol{d}; 1000)$

In this section, we compare $\boldsymbol{d}_{GA}$, $\boldsymbol{d}_{r100}$ in terms of $\phi_2(\boldsymbol{d}; 1000)$ and the CPU time required for obtaining them. Specifically, for $\boldsymbol{d}_{GA}$ and $\boldsymbol{d}_{r100}$, we generate $m = 1000$ corresponding vectors $\boldsymbol{r}$ of the subject's answers for the calculation of $\phi_2(\boldsymbol{d}_{GA}; 1000)$, and $\phi_2(\boldsymbol{d}_{r100}; 1000)$.

All the results in **Tables** 4.5 and 4.6 for $Q = 1$, and those in **Tables** 4.7 and 4.8 for $Q = 2$ suggest that $\boldsymbol{d}_{GA}$ and $\boldsymbol{d}_{r100}$ have similar performance with respect to $\phi_2(\boldsymbol{d}; 1000)$. This observation holds for both A- and D-optimality criteria, and is even clear as presented in **Figures** 4.3 and 4.4, where the bars correspond to $\phi_2(\boldsymbol{d}_{r100}; 1000)/\phi_2(\boldsymbol{d}_{GA}; 1000)$. While $\boldsymbol{d}_{GA}$ is obtained by considering ϕ_1-criterion, it outperforms $\boldsymbol{d}_{r100}$ in some cases, when $\phi_2(\boldsymbol{d}; 1000)$ is considered for design evaluations.

Table 4.5: The values of $\phi_2(\boldsymbol{d}_{GA}; 1000)$ and $\phi_2(\boldsymbol{d}_{r100}; 1000)$ with $Q = 1$ under the A-optimality criterion.

Case	$\boldsymbol{d}_{GA}$	$(\boldsymbol{d}_{r100}; 1000)$
Equal Probability	37.26	37.20
Unequal Probability	27.85	27.28

Table 4.6: The values of $\phi_2(\boldsymbol{d}_{GA}; 1000)$ and $\phi_2(\boldsymbol{d}_{r100}; 1000)$ with $Q = 1$ under the D-optimality criterion.

Case	$\boldsymbol{d}_{GA}$	$(\boldsymbol{d}_{r100}; 1000)$
Equal Probability	46.41	46.36
Unequal Probability	37.13	36.75

Table 4.7: The values of $\phi_2(\boldsymbol{d}_{GA}; 1000)$ and $\phi_2(\boldsymbol{d}_{r100}; 1000)$ with $Q = 2$ under the A-optimality criterion.

Case	$\boldsymbol{d}_{GA}$	$(\boldsymbol{d}_{r100}; 1000)$
Equal Probability	19.71	19.49
Unequal Probability (I)	16.46	16.15
Unequal Probability (II)	15.39	15.19

Table 4.8: The values of $\phi_2(\boldsymbol{d}_{GA}; 1000)$ and $\phi_2(\boldsymbol{d}_{r100}; 1000)$ with $Q = 2$ under the D-optimality criterion.

Case	$\boldsymbol{d}_{GA}$	$(\boldsymbol{d}_{r100}; 1000)$
Equal Probability	28.40	26.99
Unequal Probability (I)	25.39	25.12
Unequal Probability (II)	24.29	23.95

The current results show that designs optimizing ϕ_1 can also perform well with respect to ϕ_2. One major advantage for considering the former criterion for obtaining designs is further evident in **Tables** 4.9 and 4.10 that present the CPU times needed for obtaining $\boldsymbol{d}_{GA}$ under ϕ_1 and $\boldsymbol{d}_{r100}$ under $\phi_2(\cdot; 100)$. As also can be seen from **Figures** 4.5 and 4.6, it takes much less CPU time to obtain a $\boldsymbol{d}_{GA}$ than $\boldsymbol{d}_{r100}$. The use of the former design is thus recommended.

Table 4.9: CPU times for obtaining $\boldsymbol{d}_{GA}$ and $\boldsymbol{d}_{r100}$ for $Q = 1$.

	A-optimality		D-optimality	
Case	$\boldsymbol{d}_{GA}$	$\boldsymbol{d}_{r100}$	$\boldsymbol{d}_{GA}$	$\boldsymbol{d}_{r100}$
Equal Probability	0.88	24.82	0.60	37.97
Unequal Probability	1.17	17.47	0.60	19.53

Table 4.10: CPU times for obtaining $\boldsymbol{d}_{GA}$ and $\boldsymbol{d}_{r100}$ for $Q = 2$.

	A-optimality		D-optimality	
Case	$\boldsymbol{d}_{GA}$	$\boldsymbol{d}_{r100}$	$\boldsymbol{d}_{GA}$	$\boldsymbol{d}_{r100}$
Equal Probability	2.00	15.37	1.54	32.59
Unequal Probability (I)	2.97	44.61	1.53	32.69
Unequal Probability (II)	2.50	52.17	2.46	16.97

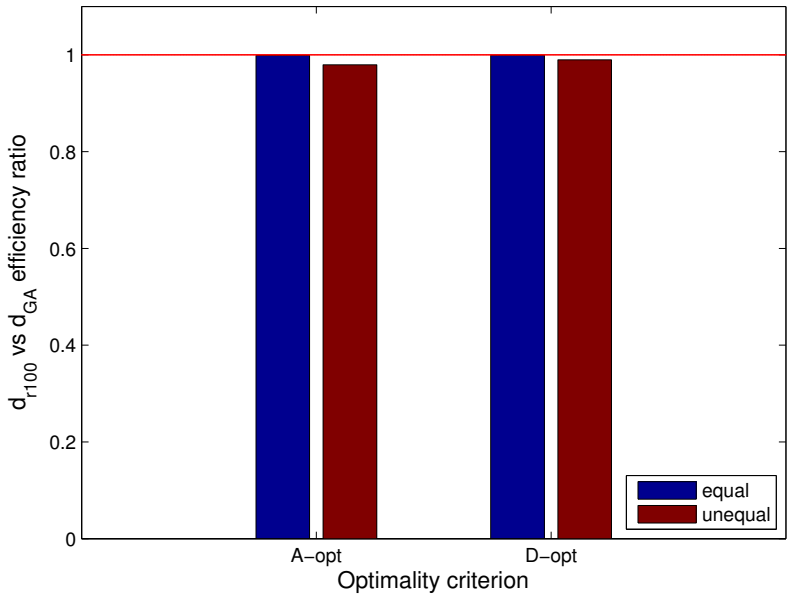

Figure 4.3: Relative efficiency of $\boldsymbol{d}_{r100}$ to $\boldsymbol{d}_{GA}$ for $Q = 1$: This plot proces $\phi_2(\boldsymbol{d}_{r100}; 1000)/\phi_2(\boldsymbol{d}_{GA}; 1000)$ for $Q = 1$ with equal/unequal $p(r \mid q)$ under A-/D-optimality.

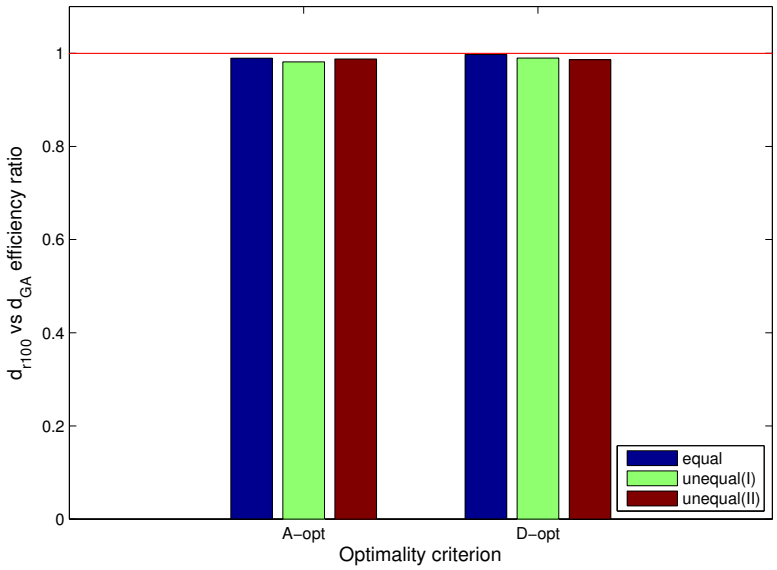

Figure 4.4: Relative efficiency of $\boldsymbol{d}_{r100}$ to $\boldsymbol{d}_{GA}$ for $Q = 2$: This plot proces $\phi_2(\boldsymbol{d}_{r100}; 1000)/\phi_2(\boldsymbol{d}_{GA}; 1000)$ for $Q = 2$ with equal/unequal $p(r \mid q)$ under A-/D-optimality.

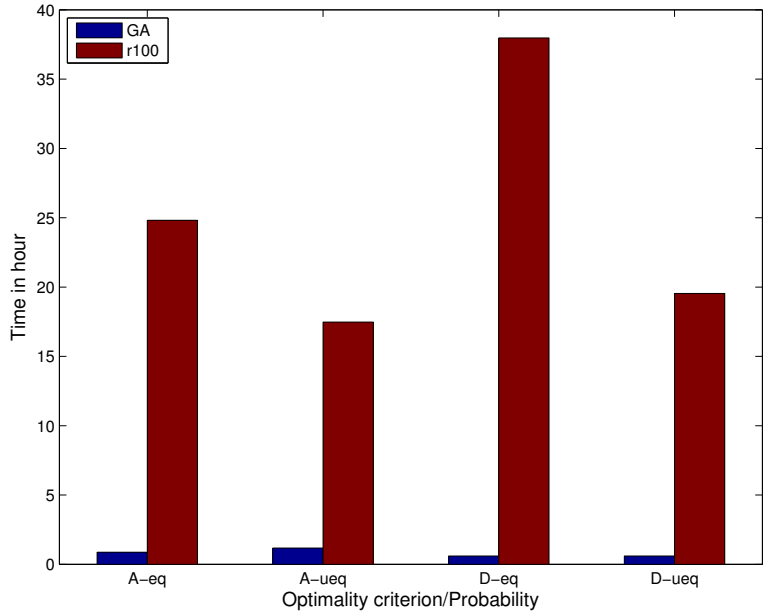

Figure 4.5: CPU times for obtaining $\boldsymbol{d}_{GA}$ and $\boldsymbol{d}_{r100}$ for $Q = 1$: This plot presents CPU times needed for obtaining $\boldsymbol{d}_{GA}$ and $\boldsymbol{d}_{r100}$ for the different cases that we studied.

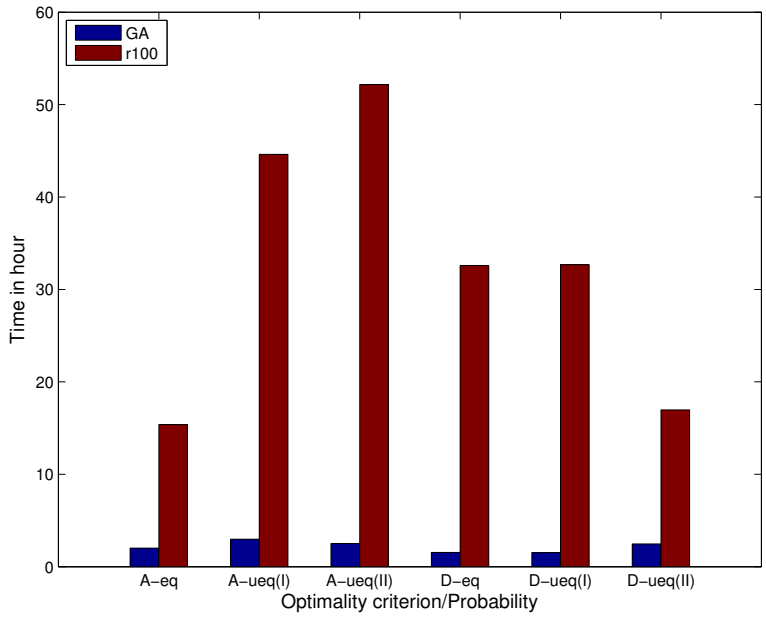

Figure 4.6: CPU times for obtaining $\boldsymbol{d}_{GA}$ and $\boldsymbol{d}_{r100}$ for $Q = 1$: This plot presents CPU times needed for obtaining $\boldsymbol{d}_{GA}$ and $\boldsymbol{d}_{r100}$ for the different cases that we studied.

Chapter 5

CONCLUSION AND DISCUSSION

We propose an efficient approach to obtain a robust designs for fMRI experiments when the design matrix depends not only on the selected designs, but also on the subject's probabilistic behavior during the experiment. The main idea is by considering the optimality criterion $\phi_1(\boldsymbol{d}) = \phi(E\{\boldsymbol{M}_d\})$. A computer algorithm such as the genetic algorithm technique can then be considered to find a design $\boldsymbol{d}$ that optimizes ϕ_1. Through case studies, we show that our obtained designs outperform some traditional fMRI designs. We also show that ϕ_1 provides a very good surrogate for $\phi_2(\boldsymbol{d}) = E\{\phi(\boldsymbol{M}_d)\}$, which is also not uncommon in practice. The value of ϕ_2 is normally unavailable and needs to be approximated. One possible way is to conduct a Monte Carlo simulation to generate m realizations of $\boldsymbol{M}_d$ for each $\boldsymbol{d}$, and calculate a summary statistic such as mean/median of the resulting m realizations of $\phi(\boldsymbol{M}_d)$ as an approximation of $\phi_2(\boldsymbol{d})$. Such an approach has recently been considered by Cordes *et al.* (2012) for tackling the same design issue. We show that, with a much less CPU time than this latter approach, our method can obtain designs that perform very well in terms of the ϕ_2-value. We thus recommend the proposed method for obtaining high-quality fMRI designs even when ϕ_2 is considered.

REFERENCES

Atkinson, A. C., A. N. Donev and R. D. Tobias, *Optimum experimental designs, with SAS*, vol. 34 (Oxford University Press Oxford, 2007).

Buračas, G. T. and G. M. Boynton, "Efficient design of event-related fmri experiments using m-sequences", Neuroimage **16**, 3, 801–813 (2002).

Cordes, D., G. Herzmann, R. Nandy and T. Curran, "Optimization of contrast detection power with probabilistic behavioral information", NeuroImage **60**, 3, 1788 – 1799 (2012).

Dale, A. M., "Optimal experimental design for event-related fmri", Human brain mapping **8**, 2-3, 109–114 (1999).

et al., J. J. M., "Functional magnetic resonance imaging: The current science of clinical application", Neuropsychology Review **17** (2007).

Kao, M.-H., "Multi-objective optimal experimental designs for er-fmri using matlab", Journal of Statistical Software **30**, 11, 1–13 (2009).

Kao, M.-H., D. Majumdar, A. Mandal, J. Stufken *et al.*, "Maximin and maximin-efficient event-related fmri designs under a nonlinear model", The Annals of Applied Statistics **7**, 4, 1940–1959 (2013).

Kao, M.-H., A. Mandal, N. Lazar and J. Stufken, "Multi-objective optimal experimental designs for event-related fMRI studies", NeuroImage **44**, 3, 849 – 856 (2009).

Lindquist, M. A., "The statistical analysis of fMRI data", Statist. Sci. **23**, 4, 439–464 (2008).

Liu, T. T. and L. R. Frank, "Efficiency, power, and entropy in event-related fMRI with multiple trial types: Part I: theory", NeuroImage **21**, 1, 387 – 400 (2004).

Maus, B., G. J. van Breukelen, R. Goebel and M. P. Berger, "Robustness of optimal design of fMRI experiments with application of a genetic algorithm", NeuroImage **49**, 3, 2433 – 2443 (2010).

Rencher, A. C. and G. B. Schaalje, *Linear models in statistics* (John Wiley & Sons, 2008).

Wager, T. D. and T. E. Nichols, "Optimization of experimental design in fMRI: a general framework using a genetic algorithm", NeuroImage **18**, 2, 293 – 309 (2003).

APPENDIX A

GENETIC ALGORITHM

We adapt genetic algorithm reported on Kao (2009) 's paper, here, we provide some details that described in the paper. This algorithm is an efficient and effective approach for finding optimal designs for ER-fMRI. The following steps present the outline of the genetic algorithm:

Step 1 Initial designs: Generate $2G$ initial designs including block designs of various block sizes, random designs, m-sequence designs and mixed designs of their combinations. Evaluate the fitness of each initial design by objective function. The objective function can set as single design criterion or weighted sum of standardized criteria for multi-objective studies.

Step 2 Crossover and mutation: Based on the probability proportional to fitness, select G pairs of different designs with replacement; these G paris are then used to generate G pair of designs via crossover and mutation, which are the offspring designs. For crossover, randomly select a cut-point and exchange between the paired designs. For mutation, randomly select a portion of elements of the resulting design, replace these elements by integers randomly generated from discrete uniform distribution over $1, 2, \ldots, Q$.

Step 3 Immigration: Generate another I designs from random designs, blocked designs and mixed designs and add to the population.

Step 4 Fitness: Evaluate the fitness of each designs in the population.

Step 5 Natural selection: With the value of fitness, Keep the best $2G$ designs to form the parents of the next generation.

Step 6 Stop: Repeat Step 2 through Step 5 until meet a stopping rule (e.g. no significant improvement is made).