INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
OPTIMAL DESIGNS FOR COMPUTER
EXPERIMENTS FOR DETERMINING OPTIMAL
OPERATING CONDITIONS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

David Wayne Donley, B.S., M.S.

* * * * *

The Ohio State University

1996

Dissertation Committee:

T. J. Santner
W. I. Notz
A. M. Dean

Approved by

Co-Adviser
Department of Statistics

Co-Adviser
Department of Statistics
To my mom who saw that her boys received college degrees ... and much more.
ACKNOWLEDGMENTS

I would like to thank Profs. Thomas J. Santner and William I. Notz for the time they devoted to the development of my research and the production of this thesis. Thanks are also extended to Prof. Angela Dean who served on my defense committee and provided some much needed encouragement during the course of my research. To my wife, Denise, I present you a promise fulfilled. Your support and encouragement made this possible, and I am eternally grateful. To Maddie and "Frito", thank you for giving me perspective on what is important.
VITA

March 20, 1967 .......................... Born - Ellsworth, Kansas

May 1989 ................................ B.S. Mathematics, University of Kansas, Lawrence, Kansas

June 1991 .............................. M.S. Statistics, The Ohio State University, Columbus, Ohio

Publications

Research Publications


Fields of Study

Major Field: Statistics

Studies in:

Computer Experimentation    T. J. Santner
Optimal Experimental Design  W. I. Notz
TABLE OF CONTENTS

DEDICATION ................................................................. ii
ACKNOWLEDGMENTS ................................................ iii
VITA ........................................................................ iv
LIST OF TABLES ........................................................... viii
LIST OF FIGURES .......................................................... xi
LIST OF NOTATION ....................................................... xiii

CHAPTER PAGE
I Background and Problem Formulation ................. 1

1.1 Introduction ......................................................... 1
   1.1.1 Modeling Computer Generated Data ................. 2
   1.1.2 Design Issues in Computer Experiments .......... 7

1.2 A Method for Quality Improvement by Computer Experimentation 12
   1.2.1 The Loss Model Approach to Parameter Design .... 13
   1.2.2 The Response Model Approach to Parameter Design . 15
   1.2.3 The Model for the Quality Characteristic .......... 16
   1.2.4 Design Criteria for the Parameter Design Problem ... 19

1.3 Preliminary Theorems and Lemmas ....................... 23
   1.3.1 Experimental Designs ................................... 24
   1.3.2 Elliptical Distributions and Stochastic Domination . 30
   1.3.3 Completely Symmetric Matrices ..................... 34
II  Determining the Optimal Univariate Operating Conditions ............... 37
  2.1  Introduction ................................................................. 37
  2.2  Defining the Design Region $\mathcal{X}$ ................................. 39
      2.2.1  Defining $\mathcal{X}_d$ .............................................. 40
      2.2.2  Defining $\mathcal{X}_e$ .............................................. 42
  2.3  Reparameterizing the Model ........................................... 42
      2.3.1  Reparameterizing the Parameter Design Problem .......... 44
  2.4  The Asymptotic Variance of the MLE .................................. 45
  2.5  Iav and Lav Optimal Designs ........................................... 46
  2.6  Mav Optimal Designs ...................................................... 53
  2.7  Local Approximate Stochastic Domination Optimal Designs .... 55
  2.8  A Method for Constructing Optimal Designs ....................... 57
  2.9  Examples of Optimal Designs .......................................... 60

III  Determining the Optimal Multivariate Operating Conditions ............ 71
  3.1  Introduction ................................................................. 71
  3.2  Defining the Design Region $\mathcal{X}$ .................................. 73
      3.2.1  Defining $\mathcal{X}_&$ ................................................ 73
      3.2.2  Defining $\mathcal{X}_e$ ................................................ 77
  3.3  Reparameterizing the Model ........................................... 77
      3.3.1  Reparameterizing the Parameter Design Problem .......... 79
  3.4  The Asymptotic Variance of the MLE .................................. 81
  3.5  IavA Optimal Designs ...................................................... 84
  3.6  IavD Optimal Designs ...................................................... 97
  3.7  Constructing Efficient Designs for the IavA and IavD Criteria .... 104
  3.8  Examples of IavA and IavD Optimal Designs ....................... 114

IV  A Comparison of Parameter Design Methodologies ............................ 131
  4.1  Introduction ................................................................. 131
  4.2  Defining the Methodology Components .................................. 132
      4.2.1  Models ............................................................. 133
      4.2.2  Design Criteria .................................................... 134
      4.2.3  Estimating the Optimal Operating Levels .................... 138
      4.2.4  Candidate Methodologies ....................................... 141
  4.3  The $(k_d, k_e) = (2, 4)$ Problem ..................................... 142
4.4 The \((k_d, k_e) = (4, 5)\) Problem .................................................. 146

V Stochastic Domination Designs for Linear Estimators .......... 162
   5.1 Introduction ................................................................. 162
   5.2 SD-Optimal Designs ..................................................... 163

VI Future Work ................................................................. 167

APPENDICES

A Calculations for the IavA-optimal design .......................... 171
   A.1 Proof of Lemma 3.5.2 .............................................. 171
   A.2 Calculating \(M(\xi^*)\) in Lemma 3.5.3 ....................... 176
   A.3 Calculating \(M^{-}(\xi^*)\) in Lemma 3.5.4 ................... 179

B Calculations for the IavD-optimal design ......................... 181
   B.1 Proof of Lemma 3.6.4 .............................................. 181
   B.2 Proof of Lemma 3.6.5 .............................................. 183
   B.3 The IavD criterion with respect to the prior \(\pi^*\) ......... 184

BIBLIOGRAPHY .............................................................. 187
List of Tables

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>An efficient sixteen point design for the lav criterion for estimating $x_d^*$ when $k_d = 1$ and $k_e = 2$.</td>
</tr>
<tr>
<td>2</td>
<td>A sixteen point Lav optimal design for estimating $x_d^<em>$ at $x_d^</em> = 5$ when $k_d = 1$ and $k_e = 2$.</td>
</tr>
<tr>
<td>3</td>
<td>An efficient sixteen point design for the Mav criterion for estimating $x_d^*$ when $k_d = 1$ and $k_e = 2$.</td>
</tr>
<tr>
<td>4</td>
<td>All designs $\mathcal{D}(i, j, k)$ supported on three isozero sets that are integrated asymptotic variance D-optimal designs for estimating $u^<em>$ (resp. $x_d^</em>$) when $k_d = 4$ and $s_d = 1$.</td>
</tr>
<tr>
<td>5</td>
<td>Reproduction of the 10 point sample design in Example 3.8.1.</td>
</tr>
<tr>
<td>6</td>
<td>Values of $r_1, r_2, t_1$ and $t_2$ in Conditions ii.) through v.) of Theorem 3.5.1 for lavA-optimal designs with respect to a prior $\pi^*$ for which the hypotheses in Assumptions 3.2.1 are true.</td>
</tr>
<tr>
<td>7</td>
<td>A 25 point lavD-optimal (resp. lavA-optimal with $\theta = 1$) design on $[-1, 1]^6$ for $(k_d, k_e) = (2, 4)$.</td>
</tr>
<tr>
<td>8</td>
<td>A 64 point lavD-optimal (resp. lavA-optimal with $\theta = 1$) design on $[-1, 1]^9$ for $(k_d, k_e) = (4, 5)$.</td>
</tr>
<tr>
<td>9</td>
<td>The five competing methodologies in the two simulated parameter design computer experiments.</td>
</tr>
</tbody>
</table>
10 Proportion of all simulation runs greater than or equal to the given loss efficiency levels for the \((k_d, k_e) = (2,4)\) parameter design problem. 144

11 Proportion of the simulation runs, with \(x_d^* \in [-0.8,0.8]^2\), that achieve the given loss efficiency levels for the \((k_d, k_e) = (2,4)\) parameter design problem. .......................... 145

12 Mean and median coordinate-wise ranks of the five competing methodologies for the simulation runs in the \((k_d, k_e) = (2,4)\) parameter design problem. .......................... 146

13 Proportion of the simulation runs greater than or equal to the given loss efficiency levels for the \((k_d, k_e) = (4,5)\) parameter design problem. 148

14 Proportion of the simulation runs, with \(x_d^* \in [-0.8,0.8]^4\), that achieve the given loss efficiency levels for the \((k_d, k_e) = (4,5)\) parameter design problem. .......................... 149

15 Mean and median coordinate-wise ranks of the five competing methodologies for the simulation runs in the \((k_d, k_e) = (4,5)\) parameter design problem. .......................... 150

16 The 25 point AMSE\(_Q\) optimal design on \([-1,1]^6\) that is used in the simulation experiment for the \((k_d, k_e) = (2,4)\) parameter design problem.153

17 The 64 point AMSE\(_Q\) optimal design on \([-1,1]^9\) that is used in the simulation experiment for the \((k_d, k_e) = (4,5)\) parameter design problem.154

18 The 25 point AMSE\(_\mu\) optimal design on \([-1,1]^6\) that is used in the simulation experiment for the \((k_d, k_e) = (2,4)\) parameter design problem.156

19 The 64 point AMSE\(_\mu\) optimal design on \([-1,1]^9\) that is used in the simulation experiment for the \((k_d, k_e) = (4,5)\) parameter design problem.157

20 The 25 point strength two orthogonal array based Latin hypercube design on \([-1,1]^6\) that is used in the simulation experiment for the \((k_d, k_e) = (2,4)\) parameter design problem. .......................... 159
21 The 64 point OA$_2$-based Latin hypercube design on $[-1,1]^9$ that is used in the simulation experiment for the $(k_d,k_c) = (4,5)$ parameter design problem. .......................................................... 160

22 The control variable portion of a 41 point central composite design for the $(k_d,k_c) = (4,5)$ parameter design problem with $s_d = 1$ and $\theta = 1$. 168

23 The control variable portion of a 163 point central composite design for the $(k_d,k_c) = (7,5)$ parameter design problem with $s_d = 1.5$ and $\theta = 0.5$. .......................................................... 169
FIGURE

1. An efficient sixteen point design for the Iav criterion for estimating $x_d^*$ when $k_d = 1$ and $k_e = 2$. ................. 63
2. A sixteen point Lav optimal design for estimating $x_d^*$ at $x_d^* = 5$ when $k_d = 1$ and $k_e = 2$. ......................... 67
3. An efficient sixteen point design for the Mav criterion for estimating $x_d^*$ when $k_d = 1$ and $k_e = 2$. ......................... 68
4. A double-fold of a bivariate slice of the design region used to obtain the isozero set weights for a design of the form $D(i, j, k)$. .............. 107
5. A double-fold of a bivariate slice of the design region used to obtain the isozero set weights for designs of the form $D(0, j_1^{(0)}, j_1^{(1)}, j_2^{(0)}, j_2^{(1)})$ and $D(j_1^{(0)}, j_1^{(1)}, j_2^{(0)}, j_2^{(1)}, k_d)$. 111
6. Perspective plots of the loss function $L(\cdot | x_d^*)$ for the $(k_d, k_e) = (2, 4)$ parameter design problem for (a) $x_d^* = (0, 0)^t$ and (b) $x_d^* = (-0.75, 0.50)^t$. 143
7. Perspective plots of the true integrated loss function $L(\cdot | x_d^*)$ for the $(k_d, k_e) = (4, 5)$ parameter design problem as a function of each pair in $x_d$ with the remaining variables in $x_d$ fixed at their levels in $x_d^*$. The true optimal operating conditions are $x_d^* = 0_4$. ....................... 151
Perspective plots of the true integrated loss function $L(\cdot|x_d^*)$ for the $(k_d, k_o) = (4, 5)$ parameter design problem as a function of each pair in $x_d$ with the remaining variables in $x_d$ fixed at their levels in $x_d^*$. The true optimal operating conditions are $x_d^* = (-0.5, 0.75, 0.5, -0.75)^t$.  

152
LIST OF NOTATION

\( \mathbf{x} \)  
A real-valued vector (also \( \mathbf{w}, \mathbf{y}, \mathbf{z} \)). (page 5)

\( \mathbf{x}_d \)  
The vector of controllable variables. (page 13)

\( \mathbf{x}_d^* \)  
The vector of optimal operating conditions for the controllable variables \( \mathbf{x}_d \). (page 16)

\( \mathbf{x}_e \)  
The vector of environmental variables. (page 13)

\( \mathbf{1}_d \)  
The vector of ones in \( \mathbb{R}^d \). (page 5)

\( \mathbf{0}_d \)  
The vector of zeroes in \( \mathbb{R}^d \). (page 31)

\( \mathbf{A} \)  
A real-valued matrix (also \( \mathbf{B}, \mathbf{C}, \) etc.). (page 5)

\( \mathbf{I}_d \)  
The \( d \)-dimensional identity matrix. (page 34)

\( \mathbf{R} \)  
The matrix whose columns form an eigensystem for set of all completely symmetric matrices. (page 35)

\( \mathbf{J}_d \)  
The \( d \)-dimensional matrix of ones (i.e., \( \mathbf{1}_d \mathbf{1}_d^T \)). (page 34)

\( \mathcal{X} \)  
The design space. (page 24)
\( \mathcal{X}_d \) The design space for the controllable variables. (page 39)

\( \mathcal{X}_e \) The design space for the environmental variables. (page 39)

\( \pi_d^* \) The assumed conditional (marginal) distribution for the parameter \( x_d^* \). (page 21)

\( \mathcal{M}_d^* \) The assumed support space for the parameter \( x_d^* \). (page 21)

\( \pi_e \) The known distribution for the environmental variables. (page 14)

\( \Xi \) A class of design measures on a Borel measurable space \( \mathcal{X} \). (page 24)

\( \Xi_1 \) A class of design measures on a Borel measurable space \([-1, 1]^k\). (page 45)

\( \xi \) A design measure either in \( \Xi \) or \( \Xi_1 \). (page 19)

\( \xi^* \) An optimal design measure either in \( \Xi \) or \( \Xi_1 \). (page 21)

\( \| \cdot \| \) Euclidean distance, i.e., \( (x^tx)^{1/2} \). (page 22)

\( \| \cdot \|_D \) Generalized Euclidean distance, i.e., \( (x^tDx)^{1/2} \) (page 32).

\( \mathcal{H}^j \) Points in the lattice \( \{-1, 0, 1\}^k \) that have \( j \) coordinates equal to zero. (page 104)

\( \mathcal{H}^{(0)} \) Points in the lattice \( \{-1, 0, 1\}^k \) that have \( j \) coordinates equal to zero and the first coordinate is 0. (page 110)

\( \mathcal{H}^{(1)} \) Points in the lattice \( \{-1, 0, 1\}^k \) that have \( j \) coordinates equal to zero and the first coordinate is \( \pm 1 \). (page 110)
Design supported on the three isozero sets $\mathcal{H}^i$, $\mathcal{H}^j$ and $\mathcal{H}^k$ with $0 \leq i < j < k \leq k_d$. (page 105)

Design supported on the isozero (subisozero) sets $\mathcal{H}^0$, $\mathcal{H}^{j_1(0)}$, $\mathcal{H}^{j_1(1)}$, $\mathcal{H}^{j_2(0)}$ and $\mathcal{H}^{j_2(1)}$, with $0 < j_1 < j_2 < k_d$. (page 111)

Design supported on the isozero (subisozero) sets $\mathcal{H}^{j_1(0)}$, $\mathcal{H}^{j_1(1)}$, $\mathcal{H}^{j_2(0)}$, $\mathcal{H}^{j_2(1)}$ and $\mathcal{H}^{k_d}$, with $0 < j_1 < j_2 < k_d$. (page 111)
CHAPTER I

Background and Problem Formulation

1.1 Introduction

The quality of a product or process typically depends on some factors that can be controlled by the manufacturer and other factors that cannot. The goal of a quality improvement experiment is to determine the levels of the controllable factors that make the product or process insensitive to manufacturing and environmental variations. When the experiment can be conducted physically, well-known methodologies, such as response surface methodology (RSM) (Myers 1991) and parameter design (Taguchi 1986), can be used to determine these optimal operating conditions. However, for products or processes that are studied using computer experimentation, the experimental strategies for determining optimal operating conditions are less well established. In this dissertation we develop a statistical methodology for achieving quality improvement in a product or process that explicitly considers the special circumstances inherent to computer experimentation.

The proposed methodology is based on the response model approach to parameter design advanced by Welch, Yu, Kang, and Sacks (1990). This approach adopts
principles from both RSM and parameter design to conduct quality improvement experiments. We present brief overviews of parameter design and the response model approach in Section 1.2.1 and Section 1.2.2, respectively. Section 1.2 also contains the description of our methodology including the model for the response (Section 1.2.3) and design criteria (Section 1.2.4). In Chapter II we determine the experimental designs that are used in our methodology in the case of a single controllable factor. The experimental designs for the multivariate case are derived in Chapter III. In Chapter IV we perform two simulation experiments comparing our proposed methodology against several ad hoc methods from the literature that are likewise based on the response model approach to parameter design. Chapter V presents some ancillary results for one of the univariate criteria defined in Section 1.2.4, and Chapter VI contains some concluding remarks and states directions for future research.

We conclude this section with a review of the literature pertinent to the problem of quality improvement by means of computer experimentation. Our discussion highlights the nuances to consider when modeling (Section 1.1.1) computer generated data and selecting the input configurations for the computer code (Section 1.1.2). Throughout the dissertation we recall established results and statistical definitions. The relevant theorems, lemmas, and definitions are given in Section 1.3.

1.1.1 Modeling Computer Generated Data

To conduct a computer experiment for quality improvement using the response model approach, an experimenter must specify a statistical model that relates the output $y$ to
the inputs $x$. We assume $y(x)$ is a real-valued scalar function of a $d$-dimensional input vector $x$, and although $y(x)$ is actually deterministic, for the purpose of statistical modeling we treat it as a realization of a stochastic process, $Y(x)$, that includes a regression model. We assume, following Sacks, Welch, Mitchell and Wynn (1989), that

$$Y(x) = f(x)^T \beta + Z(x) \quad (1.1.1)$$

where $f(x)$ is the vector of known regression functions, $\beta$ is the vector of unknown regression coefficients and $Z(\cdot)$ is a mean zero Gaussian stochastic process.

The idea of the model in (1.1.1) is that the regression model $f(x)^T \beta$ captures the large-scale behavior of the output, and the stochastic process term $Z(\cdot)$ describes the small-scale deviations. To capture the correlations between the systematic errors for pairs of input sites, the covariance function for the stochastic process $Z(\cdot)$ needs to be specified. An intuitive choice for the covariance function is

$$\text{Cov}(Z(x), Z(w)) = \sigma^2 R(x, w) = \sigma^2 \prod_{i=1}^{d} \exp\{-\theta_i |x_i - w_i|^\rho_i\} \quad (1.1.2)$$

where $\sigma^2 > 0$ is the constant process variance and $R(x, w)$ is the correlation function with parameters $\theta_i > 0$ and $\rho_i \in (0, 2]$. The correlation function in (1.1.2) relates the belief that the residuals for nearby pairs of inputs are more positively correlated than the residuals for pairs of inputs that are farther apart.

The correlation function defines characteristics of the stochastic process $Z(\cdot)$. For example, a process with correlation function (1.1.2) is continuous for all values of $\theta_i$ and $\rho_i$, and the process is mean square differentiable in the $i^{th}$ coordinate direction if and only if $\rho_i = 2$ (Sacks et al., 1989). In this case, the process is infinitely
mean square differentiable. Consequently, the correlation function in (1.1.2) does not allow processes that are differentiable yet not infinitely differentiable. Alternative correlation functions that allow more degrees of differentiability have been studied by Stein (1989), Handcock (1991) and Currin, Mitchell, Morris, and Ylvisaker (1988). The increased flexibility afforded by these alternative correlation functions comes at expense of model simplicity, since there are more parameters that must be estimated.

How best to specify the regression model $\mathbf{f}(\mathbf{x})'\beta$ in (1.1.1) is rather uncertain, since there is a complexity "trade-off" that occurs between the regression model $\mathbf{f}(\mathbf{x})'\beta$ and the stochastic process error term $Z(\cdot)$. Consider the covariance function in (1.1.2). If an experimenter specifies a regression model with many terms, then the resulting residuals after fitting the model would likely be weakly correlated with one another for all pairs of input sites. The stochastic error term $Z(\cdot)$ is already assumed to be a mean zero Gaussian process with $R(\mathbf{x},\mathbf{x}) = 1$, so the small correlations mean we could assume independence for draws from the process $Z(\cdot)$. In contrast, if the experimenter specifies the constant mean regression model, i.e., $\mathbf{f}(\mathbf{x})'\beta = \beta_0$, then the Gaussian stochastic error term $Z(\cdot)$ must capture all of the behavior of output. Thus, the estimated correlation function in (1.1.2) would be quite complicated.

In practice, the constant mean stochastic process model is used frequently, because the simplification does not appear to affect the quality of predictions. Examples in the literature suggest that modeling the stochastic error term properly is more important than the selection of the regression model in (1.1.1) for the purpose of prediction. In Chapter IV we propose some ad hoc competitors to our methodology that use
the constant mean stochastic process model. The following example presents the procedures that we use in Chapter IV for forming a predictor of the output and for estimating the model parameters.

**Example 1.1.1 (Constant Mean Stochastic Process Model)** Assume the output $y(x)$ of a computer code is a realization of the Gaussian stochastic process $Y(x)$ in (1.1.1) with constant mean $\beta_0$ and covariance function given by (1.1.2). For a set of input sites $\{v_1, v_2, \ldots, v_n\}$, let $Y = (Y(v_1), Y(v_2), \ldots, Y(v_n))^t$ be the vector of unobserved draws of the process at these input sites. The vector $Y$ is a multivariate normal random variable with $E[Y] = \beta_0 \mathbf{1}_n$ and $\text{Cov}[Y, Y] = \sigma^2 \mathbf{R}$ where $\mathbf{R}_{i,j} = R(v_i, v_j)$.

We use realizations $y$ of the vector $Y$ to predict the output $y(x)$ at an untried input $x$. We use the best linear unbiased predictor (BLUP) of $Y(x)$,

$$\hat{y}(x) = \lambda^t(x)y,$$

where the vector $\lambda(x)$ minimizes

$$MSE(\hat{y}(x)) = E\left[\left(\lambda^t(x)Y - Y(x)\right)^2\right] = \sigma^2 \left(\lambda^t(x)\mathbf{R}\lambda(x) - 2\lambda^t(x)r(x) + 1\right)$$

subject to the “unbiasedness” constraint

$$E\left[\lambda^t(x)Y\right] = E[Y(x)].$$

The vector $r(x)$ in (1.1.4) is the vector of correlations between the random variable $Y(x)$, at the untried input $x$, and the vector of inputs $Y$, i.e., $r(x) = \text{Corr}[Y, Y(x)]$. 

To determine \( \lambda(x) \), suppose the vectors \( \theta \) and \( \rho \) which contain the parameters of the correlation function in (1.1.2) are fixed and known. Using the method of Lagrange multipliers, we obtain that the minimizing choice for the vector \( \lambda(x) \) is

\[
\lambda(x) = R^{-1/2}(1_n R^{-1/2})^{-1} + \left( I_n - R^{-1}1_n (1_n' R^{-1} 1_n)^{-1} 1_n' \right) R^{-1} r(x) \tag{1.1.6}
\]

Thus, the BLUP of \( y(x) \) is

\[
y(x) = \hat{\beta}_0 + r'(x) R^{-1} (Y - 1_n \hat{\beta}_0) \tag{1.1.7}
\]

where \( \hat{\beta}_0 = (1_n' R^{-1} 1_n)^{-1} 1_n' R^{-1} Y \) is the generalized least squares estimator of \( \beta_0 \).

In practice, the correlation function parameters \( \theta \) and \( \rho \) are not known but must be estimated from the data in order to compute the BLUP in (1.1.7). We use the method of maximum likelihood to estimate the unknown model parameters \( (\beta_0, \sigma^2, \theta, \rho) \). The random vector \( Y \) is multivariate normal. Hence, for given values of the correlation parameters \( (\theta, \rho) \), the maximum likelihood estimate of \( \beta_0 \) is the generalized least-square estimate

\[
\hat{\beta}_0 = (1_n' R^{-1} 1_n)^{-1} 1_n' R^{-1} y \tag{1.1.8}
\]

and the MLE of \( \sigma^2 \) is

\[
\hat{\sigma}^2 = \frac{1}{n} (y - 1_n \hat{\beta}_0)' R^{-1} (y - 1_n \hat{\beta}_0). \tag{1.1.9}
\]

Replacing the parameters \( \beta_0 \) and \( \sigma^2 \) in the likelihood with their estimates in (1.1.8) and (1.1.9), the maximum likelihood estimates for \( (\theta, \rho) \) are the values that minimize the function

\[
n \log \hat{\sigma}^2 + \log \det R. \tag{1.1.10}
\]
We obtain the values of \((\theta, \rho)\) that minimize (1.1.10) using a numerical optimization routine. The software that we use is a public domain implementation of the simulated annealing algorithm. The algorithm is given in Corana et al. (1987), and the implementation of the algorithm is described in Goffe, Ferrier and Rogers (1994).

From (1.1.6), we see the BLUP (1.1.7) is the sum of two independent linear combinations of the multivariate normal vector \(Y\). Therefore, the fitting of the BLUP can be obtained in two stages. First, the generalized least-squares predictor, \(\hat{\theta}_0\), of the output is obtained. Then, the second term in (1.1.7) is a smoothing interpolation of the remaining residuals. Thus, the choice of the BLUP (1.1.7) to predict \(y(x)\) has the desirable consequence that the resulting predicted surface will interpolate the observed data.

1.1.2 Design Issues in Computer Experiments

Complex computer codes that describe complicated physical processes require extensive computing time even on the most powerful machines. Each observation is generated at great expense. Therefore, careful consideration of the set of input configurations at which the code is run during the experiment is crucial.

The standard experimental design methodologies (e.g., Fisher, 1935; and Box, Hunter and Hunter, 1978), that were proposed for the analysis of data collected from physical experiments and for determining sources of variation are probably not appropriate for computer experiments because of the deterministic nature of the data. The uncertainty in a computer experiment arises solely from systematic departures of
the regression model from the model embodied in the computer code, so the statistical design principles of blocking, randomization, and replication that reduce random error are irrelevant. Therefore, experimental design methodologies that address the effects of systematic errors (bias) must be considered.

Experimental design strategies that consider both the error that arises due to random variation and bias error caused by model inadequacy have been investigated by several authors (e.g., Box and Draper, 1959; Karson, Manson and Hader, 1969; Kiefer, 1973; Welch, 1983; and Sacks and Ylvisaker, 1984). Each of these articles propose a design criterion that is a function of the mean squared error for prediction averaged over the set of possible design sites. These mean squared error criteria can be divided into two terms representing the variance and the bias error, respectively. While it is generally impossible to simultaneously minimize both the variance and the bias term by choice of design, in each article the authors found that the all-bias optimal design adequately approximates the true optimal design provided the variance term does not dominate the bias term. Further, the design points for these all-bias optimal designs are uniformly distributed over the design space.

A concomitant problem with the approaches of Box and Draper (1959) and Karson, Manson, and Hader (1969) is the requirement that the experimenter must not only specify the true polynomial model but must also parametrically specify the difference of the fitted model from the true expected value of the response surface. Welch (1983) and Sacks and Ylvisaker (1984) address this problem by modeling the departures of the approximating polynomial model from the expected value of the
response nonparametrically. This protects the experimenter against a large class of possible departures.

The procedure of Welch (1983) has been used to construct designs for computer experiments using the stochastic process models of Section 1.1.1 with the stochastic error term modeled nonparametrically. The optimal designs are determined using the Average Mean Squared Error of the Response Estimators (AMSE) design criterion in the ACED experimental design software package (Welch, 1985). Optimal experimental designs for computer experiments have also been constructed using the integrated mean squared error (e.g., Sacks, Schiller, and Welch, 1989; and Sacks, Welch, Mitchell and Wynn, 1989), and the maximum mean squared error (Sacks and Schiller, 1988). Bayesian approaches to design of computer experiments (e.g., Currin, Mitchell, Morris and Ylvisaker, 1988; and Mitchell, Morris and Ylvisaker, 1993) use the determinant of the Gaussian posterior variance-covariance matrix to determine optimal designs. This approach has an information theoretic interpretation, since minimizing this determinant is equivalent to maximizing the information as measured by Shannon's entropy (Shewry and Wynn, 1987). Johnson, Moore and Ylvisaker (1990) proposed design criteria for computer experiments that optimize the intersite distances between the input points.

All design strategies mentioned above were developed to provide accurate predictors of the response (output). However, prediction of the output is not the primary objective in this thesis. Instead, it is determining the location of the optimal levels of the controllable factors. The above approaches also involve the minimization of a
design criterion that is nonlinear, so construction of optimal designs requires numerical optimization procedures. These procedures can be inefficient if the number of design points or the dimension of the inputs is large. Therefore, we consider simpler and more appropriate design alternatives.

McKay, Conover, and Beckman (1979) suggested a simple and inexpensive alternative — Latin hypercube sampling. Latin hypercube sampling was originally proposed as an alternative sampling scheme for Monte Carlo integration. However, due to their ease of construction and their uniformity of distribution over the design region, which is a key feature of the "all-bias" designs, Latin hypercube designs have been used for Kriging prediction of computer generated data (Welch et al., 1992), Bayesian prediction of deterministic data (Mitchell, Morris and Ylvisaker, 1993), and advocated for adaptive modeling procedures, like Friedman's, (1991) multivariate adaptive regression splines, (Owen, 1992a).

An $N$ point uniform Latin hypercube sample is constructed as follows. Assume that there are $K$ inputs each ranging over $[0, 1]$. For each coordinate, $k \in \{1, \ldots, K\}$, generate a random permutation, $p_k(1), \ldots, p_k(N)$, of the integers $1, \ldots, N$ and a random sample, $U_k^{(1)}, \ldots, U_k^{(N)}$, of size $N$ from the Uniform$(0,1)$ distribution in such a way that the $K$ permutations and $NK$ uniform variates are independent. Then, the $k^{th}$ coordinate of the $j^{th}$ input configuration in the Latin hypercube sample is

$$X_k^{(j)} = \frac{p_k(j) - U_k^{(j)}}{N},$$

(1.1.11)

where $k \in \{1, \ldots, K\}$ and $j \in \{1, \ldots, N\}$. The Latin hypercube sample given by (1.1.11) samples once from each of the equiprobable regions $((i - 1)/N, i/N)$,
i \in \{1, \ldots, N\}, for each marginal. This stratifies the design across each marginal thereby providing some very desirable variance reduction properties for the estimator \( \hat{Y} \).

In particular, McKay et al. (1979) showed that the variance of \( \hat{Y} \) is smaller under Latin hypercube sampling (LHS) than it is under simple random sampling (SRS) whenever the behavior of \( y(x) \) is monotone in each of the inputs. Stein (1987) proved the complementary and stronger asymptotic result that the variance of \( \hat{Y} \) under LHS is smaller than the variance of \( \hat{Y} \) under SRS variance provided \( y(x) \) is a square integrable function. The amount of the reduction in variance due to LHS depends on the extent that \( y(x) \) is additive — the more additive the greater the reduction.

Essentially, the stratifying of the design across each marginal filters out the additive component of the output. Owen (1992b) and Tang (1993) independently studied designs stratified across higher order margins for achieving greater variance reduction. Using strength two (or three) orthogonal arrays, they constructed Latin hypercube designs, stratified across all bivariate (or trivariate) margins, that eliminated the portion of the Monte Carlo variance attributable to the lower (up to second or third) order effects of \( y(x) \) thereby reducing the variance of \( \hat{Y} \). Of course, this approach relies on the existence of appropriately dimensioned orthogonal arrays of the proper order. Other authors (e.g., Iman and Conover, 1982; and Stein, 1987) proposed enhancements to Latin hypercube sampling that incorporate dependencies between the inputs into the design. However, in general, these approaches do not work when \( N \) is small.
Even with the existence of an appropriately dimensioned orthogonal array, the applicability of these designs as well as other Latin hypercube designs for the parameter design problem is still open to question. Currently, these designs along with the AMSE-optimal designs (Welch, 1983) are two of the most widely used alternatives for the parameter design problem in computer experimentation, yet neither was originally intended for this purpose. However, they are not unreasonable choices, since they have been shown to have "good" properties in the presence of bias. In this dissertation we present a design strategy for the parametric design problem in computer experimentation that specifically considers the estimation of the optimal operating characteristics of the simulated product or process.

1.2 A Method for Quality Improvement by Computer Experimentation

In this section we describe our methodology for achieving quality improvement by means of computer experimentation. Our methodology is based on the response model approach to parameter design. It has three components: a model for the computer generated output (Section 1.2.3), a design criterion (Section 1.2.4) and a method for estimating the optimal operating conditions (Section 1.2.3). We begin this section with background information on parameter design (Section 1.2.1) and the response model approach (Section 1.2.2).
1.2.1 The Loss Model Approach to Parameter Design

Over the past two decades the quality improvement of products has become increasingly more important in companies worldwide, and one of the most widely implemented methodologies for bettering quality has been Taguchi's Robust Design method (Taguchi, 1986; and Phadke, 1989). The Robust Design method uses planned experiments during research and development to produce high quality products that can be manufactured quickly and with low cost. An important part of this methodology is designing products that are less sensitive to manufacturing and environmental variations. This aspect of the Robust Design method is called parameter design.

A defining feature of the Robust Design method is the delineation of the inputs into factors that can be controlled by the manufacturer, denoted by \( x_d \), and factors that are determined by environmental conditions, denoted by \( x_e \). The goal of parameter design is to choose the levels of the factors in \( x_d \), so that the resulting product performs as expected regardless of the levels of the factors in \( x_e \). To determine the levels of the controllable factors \( x_d \) that provide robust function of the product, there must be a measure of the product’s performance, called the quality characteristic, for any combination of inputs \((x_d, x_e)\). For example, this measure might quantify product performance by the size of the output or by the squared deviance of the output from some target value \( \Delta_0 \). We denote this measure by \( y(x_d, x_e) \) and assume that it is a real-valued scalar quantity.

In the Taguchi approach, a.k.a. the loss model approach, to parameter design, the experimenter observes realizations of the quality characteristic \( y(x_d, x_e) \) for each
point determined by crossing a pair of orthogonal arrays for $x_d$ and $x_c$ (Phadke 1989). This means for each level of $x_d$ in the "inner array", "replicate" observations of $y(x_d, \cdot)$ are generated using each level of $x_c$ in the "outer array". These "replicate" observations are then pooled into a loss statistic that is a function of $x_d$ only, e.g.,

$$L(x_d) = \int y(x_d, x_c) d\pi_c(x_c),$$  \hspace{1cm} (1.2.1)

which is just the average of the quality characteristic over the distribution of $x_c$. The loss statistic $L(x_d)$ is, in turn, modeled using a main effects, possibly a second-order, regression model in the $x_d$. The model is fit by least squares, and a predictor $\hat{L}(\cdot)$ of $L(\cdot)$ is obtained. The optimal operating levels of $x_d$ are then estimated by minimizing the predictor $\hat{L}(x_d)$ over the levels of $x_d$.

Critics of the Taguchi approach to parameter design believe the modeling of the loss statistic $L(\cdot)$ is inferior to modeling the response $Y(\cdot)$. They conclude that an opportunity to gain understanding about the underlying mechanics of the product has been forsaken for the purpose of simplicity. Advocates, like Shin Taguchi, respond by saying, "The goal of parameter design is not to characterize the system but to achieve robust function," (Nair, 1992). We agree with this statement, but we seek a compromise by modeling the response $Y(\cdot)$ directly, yet choosing the design points to best estimate the minimizing levels of $x_d$ for the loss function $L(\cdot)$. This results in designs that are considerably smaller than the product designs in the Taguchi approach.
1.2.2 The Response Model Approach to Parameter Design

The response model approach to parameter design in computer experimentation regards the quality characteristic \( y(x_d, x_c) \) as a realization of the stochastic process \( Y(x_d, x_c) \). A model for the process \( Y(x_d, x_c) \) is proposed. This model is fitted, and a predictor \( \hat{y}(\cdot, \cdot) \) of the output is obtained. The predictor \( \hat{L}(\cdot) \) is then computed by integrating the predictor \( \hat{y}(\cdot, \cdot) \) over the distribution of the environmental factors, i.e.,

\[
\hat{L}(x_d) = \int \hat{y}(x_d, x_c) d\pi_c(x_c) \tag{1.2.2}
\]

where \( \pi_c(\cdot) \) is the known joint probability distribution for the environmental variables \( x_c \). Estimates of the optimal operating conditions are then determined by minimizing \( \hat{L}(x_d) \) in (1.2.2) over the possible values of \( x_d \).

By modeling the stochastic process \( Y(x_d, x_c) \), the experimental design is not a product array. This provides substantial savings from the number of experimental units required by Taguchi (1986) approach. For example, in a parameter design computer experiment for determining the optimal widths of six transistors in a VLSI integrated circuit, Welch et al (1990) found that a 60 point experimental design based on the response model approach performed better in terms of predicting the integrated loss function and determining the optimal operating conditions than a 200 (40 \( \times \) 5) point design based on the loss model approach. As an added bonus, there is a predictor of the response which allows a more thorough investigation of the causal relationships and the underlying mechanics of the product under study.
1.2.3 The Model for the Quality Characteristic

We model the relationship between the random quality characteristic $Y(x_d, x_e)$ and the inputs $(x_d, x_e)$ as

$$Y(x_d, x_e) = \beta_0 + \beta_d^i x_d + x_d^i B_d x_d + \beta_e^i x_e + x_e^i B_I x_e + Z(x_d, x_e)$$  \hspace{1cm} (1.2.3)$$

where $x_d \in X_d \subset \mathbb{R}^{k_d}$, $x_e \in X_e \subset \mathbb{R}^{k_e}$, the error term $Z(\cdot, \cdot)$ is a mean zero Gaussian process with constant variance $\sigma^2 > 0$ and zero correlations for draws at distinct input sites, and $\beta_0$, $\beta_d$, $B_d$, $\beta_e$ and $B_I$ are conformable arrays of unknown regression coefficients. In addition, we consider only applications where a unique configuration of the controllable variables, denoted by $x_d^*$, minimizes the value of the expected quality characteristic averaged over the distribution of the environmental variables.

This implies that the symmetric matrix $B_d$ is positive definite.

Under the model assumptions given in the previous paragraph, the expected integrated loss function is a convex quadratic function of $x_d$, i.e.

$$L(x_d) = \int E[Y(x_d, x_e)]d\pi_e(x_e)$$

$$= \beta_0 + \beta_d^i x_d + x_d^i B_d x_d + \beta_e^i \mu_e + x_e^i B_I \mu_e$$  \hspace{1cm} (1.2.4)$$

where $E^\pi_e(x_e) = \mu_e$. This expected loss function attains its minimum value, uniquely, at the unknown operating levels

$$x_d^* = -\frac{1}{2} B_d^{-1}(\beta_d + B_I \mu_e).$$  \hspace{1cm} (1.2.5)$$
To determine an estimator of $x_d^*$ in (1.2.5), we expand the coefficient arrays in (1.2.3) to rewrite the model as

$$Y(x_d, x_e) = \beta_0 + \sum_{i=1}^{k_d} \beta_d, x_d + \sum_{i=1}^{k_d} \beta_d, x_d^2 + \sum_{1 \leq i < j \leq k_d} \beta_{d,ij} x_d x_{d,ij} + \sum_{i=1}^{k_r} \beta_{r,ic} x_c + \sum_{i=1}^{k_r} \sum_{j=1}^{k_r} \beta_{r,ij} x_d x_{e;ij} + Z(x_d, x_e)$$

(1.2.6)

where

$$\beta_d = (\beta_d)_i \text{ for } i \in \{1, \ldots, k_d\},$$

(1.2.7)

$$\beta_{d,ij} = \frac{2}{1 + \delta_{ij}} (B_d)_{ij} \text{ for } i < j \in \{1, \ldots, k_d\},$$

(1.2.8)

$$\beta_{r,ij} = (\beta_r)_j \text{ for } j \in \{1, \ldots, k_r\}, \text{ and}$$

(1.2.9)

$$\beta_{l,ij} = (B_l)_{ij} \text{ for } i \in \{1, \ldots, k_d\} \text{ and } j \in \{1, \ldots, k_r\}.$$  

(1.2.10)

The term $\delta_{ij}$ in (1.2.8) is the Kroneker delta function; $\delta_{ij} = 1$ if $i = j$ and equals zero otherwise. Using the least squares estimators for the regression coefficients in (1.2.6), our estimator of the optimal conditions conditions is

$$\hat{x}_d^* = -\frac{1}{2} \hat{B}_d^{-1} (\hat{\beta}_d + \hat{B}_l \mu_e)$$

(1.2.11)

where $\hat{\beta}$, $\hat{B}_d$ and $\hat{B}_l$ are the arrays of least-squares estimators for the parameters in $\beta$, $B_d$ and $B_l$.

Let $f(x_d, x_e)$ be the $\left(\frac{k_d+2}{2}\right) + (k_d+1)k_e$-dimensional vector of linearly independent regression functions in (1.2.6), and let $\beta$ denote the associated vector of unknown regression parameters. Under the assumptions stated above for the process $Z(\cdot, \cdot)$, the least-squares estimator $\hat{\beta}$ of $\beta$ is also the maximum likelihood estimator (m.l.e.),
and it has a Gaussian distribution. Thus, by invariance of the m.l.e. under continuous mappings, $\hat{x}_d^*$ is the maximum likelihood estimator of $x_d^*$.

To construct designs for estimating $x_d^*$ with $\hat{x}_d^*$, we appeal to the large sample distribution of $\hat{x}_d^*$. To begin, fix a design $\xi$ in the class of all design measures $\Xi$, and let $\{\xi_n\}_{n>1} \subset \Xi$ be a sequence of exact design measures (see Section 1.3.1) of increasing size $n$ such that $\xi_n$ converges weakly to $\xi$. By the continuity of the regression functions in $f(x_d, x_e)$, the sequence of information matrices for the designs $\{\xi_n\}_{n>1}$ satisfies

$$M(\xi_n) \xrightarrow{n \to \infty} M(\xi). \quad (1.2.12)$$

Next, let $\hat{\beta}_n$ be the m.l.e. of $\beta$ for the design $\xi_n$. The random vector $\hat{\beta}_n$ has a Gaussian distribution with mean $\beta$ and dispersion matrix $\sigma^2 n^{-1} M^{-}(\xi_n)$ where $M^{-}(\xi_n)$ is the Moore-Penrose generalized inverse of $M(\xi_n)$. Using (1.2.12) and the continuity of the characteristic function of a Gaussian random variable, we have that

$$\sqrt{n} \left( \hat{\beta}_n - \beta \right) \xrightarrow{\mathcal{L}} N \left( 0, \sigma^2 M^{-}(\xi) \right) \quad (1.2.13)$$

where $M^{-}(\xi)$ is the Moore-Penrose generalized inverse of the information matrix for $\beta$ for the design $\xi$. Finally, consider $x_d^*$ in (1.2.5) as a function of $\beta$. The gradient of $x_d^*$ with respect to $\beta$, i.e. $\frac{\partial x_d^*}{\partial \beta_i}$ (see Section 3.4) is nonzero for all values $\beta$ for which $B_d$ is positive definite. It follows that

$$\sqrt{n} \left( \hat{x}_d^{*(n)} - x_d^* \right) \xrightarrow{\mathcal{L}} N \left( 0, \sigma^2 \left( \frac{\partial x_d^*}{\partial \beta_i} \right)^T M^{-}(\xi) \left( \frac{\partial x_d^*}{\partial \beta_i} \right) \right) \quad (1.2.14)$$

where $\hat{x}_d^{*(n)}$ is the m.l.e. of $x_d^*$ for the design $\xi_n$. Hence, $\hat{x}_d^*$ in (1.2.11) is consistent for estimating $x_d^*$, and it is asymptotically normal with covariance matrix proportional
to
\[
\left( \frac{\partial \hat{x}_d^*}{\partial \beta^i} \right) M^{-1}(\xi) \left( \frac{\partial \hat{x}_d^*}{\partial \beta^i} \right)^t
\]  
(1.2.15)
for any given \( \xi \in \Xi \).

While our assumptions concerning \( Z(\cdot, \cdot) \) sacrifice some of the modeling flexibility of existing methodologies, this proposed response model approach formulates the problem of interest more precisely and provides an explicit estimator, \( \hat{x}_d^* \), for the unknown operating levels, \( x_d^* \). In Section 1.2.4 we define design criteria, in both the univariate and multivariate cases, that allow \( \hat{x}_d^* \) to estimate \( x_d^* \) as accurately as possible. Thus, our proposed methodology directly addresses the parameter design problem.

1.2.4 Design Criteria for the Parameter Design Problem

The design criteria that we define in this section for use in our methodology are convex functions of the asymptotic variance-covariance matrix of \( \hat{x}_d^* \) given in (1.2.15). The parameter of interest \( x_d^* \) (1.2.5) is a nonlinear function of the parameters in \( \beta \), so this matrix will depend on the values of \( \beta \) and \( x_d^* \). This means real-valued functions of the matrix (1.2.15) and the associated optimal designs will likewise depend on these unknown parameters.

Abdelbasit and Plackett (1983) describe four methods that can be used to determine optimal designs in this situation. One can define initial estimates of \( x_d^* \) and/or \( \beta \), define a prior distribution for the parameters \( x_d^* \) and/or \( \beta \) and then take expectations, use a sequential method in which we continually update our estimate, or
choose the design points in such a way that the value of design function is free of the parameters.

This thesis uses the first two methods. If we regard initial estimates as the expected values of degenerate priors, then our method of removing the design function's dependence on the unknown parameters can be considered as a limited-scale Bayesian method. Whereas a full-scale Bayesian method would use a complete prior for the unknown parameters and then eliminate the unknown parameters by taking expectations, we require only the values for a sufficient number of moments for the unknown parameters $\beta$ and $x_d^*$ to determine an optimal design. We also investigate, in the univariate case, a minimax strategy for eliminating the unknown parameters from the design function.

**Design Criteria for the Univariate Case**

In this section we define the four design criteria for the univariate case. When there is only a single controllable factor, the parameter of interest $x_d^*$ in (1.2.5) and the large sample variance of $\tilde{x}_d^*$ in (1.2.15) are scalar quantities. This means no special real-valued functions of (1.2.15) must be defined in order to facilitate the determination of optimal experimental designs. The first three design criteria that we define in this section are functions of (1.2.15), while the fourth is a new criterion that we propose for making inferences on $x_d^*$.

To define the criteria, let $\mathcal{X}$ be the design space for the computer experiment and $\Xi$ be the class of competing design measures defined on the Borel sets of $\mathcal{X}$. In
addition, let $\pi_d^*$ be a given prior distribution for $x_d^*$, and let $\mathcal{M}_d^*$ be a given closed interval in the real line $\mathbb{R}$ such that $x_d^* \in \mathcal{M}_d^*$. The assumptions that we impose on $\pi_d^*$ and $\mathcal{M}_d^*$ for determining the optimal designs of Chapter II are given in Section 2.2.

**Design Criterion 1.2.1 (Lav)** A design measure $\xi^* \in \Xi$ is called a Local asymptotic variance optimal design for estimating $x_d^*$ at the value $\mu_d^*$ if and only if

$$
\left( \frac{\partial x_d^*}{\partial \beta^t} \right) \mathcal{M}^-(\xi^*) \left( \frac{\partial x_d^*}{\partial \beta^t} \right)^t \bigg|_{x_d^*=\mu_d^*} = \min_{\xi \in \Xi} \left( \frac{\partial x_d^*}{\partial \beta^t} \right) \mathcal{M}^-(\xi) \left( \frac{\partial x_d^*}{\partial \beta^t} \right)^t \bigg|_{x_d^*=\mu_d^*}.
$$

(1.2.16)

**Design Criterion 1.2.2 (Lav)** A design measure $\xi^* \in \Xi$ is called an Integrated asymptotic variance optimal design for estimating $x_d^*$ with respect to the prior $\pi_d^*$ if and only if

$$
E_{x_d^*} \left[ \left( \frac{\partial x_d^*}{\partial \beta^t} \right) \mathcal{M}^-(\xi^*) \left( \frac{\partial x_d^*}{\partial \beta^t} \right)^t \right] = \min_{\xi \in \Xi} E_{x_d^*} \left[ \left( \frac{\partial x_d^*}{\partial \beta^t} \right) \mathcal{M}^-(\xi) \left( \frac{\partial x_d^*}{\partial \beta^t} \right)^t \right].
$$

(1.2.17)

**Design Criterion 1.2.3 (Mav)** A design measure $\xi^* \in \Xi$ is called a Minimax asymptotic variance optimal design for estimating $x_d^*$ with respect to the interval $\mathcal{M}_d^*$ if and only if

$$
\max_{x_d^* \in \mathcal{M}_d^*} \left( \frac{\partial x_d^*}{\partial \beta^t} \right) \mathcal{M}^-(\xi^*) \left( \frac{\partial x_d^*}{\partial \beta^t} \right)^t = \min_{\xi \in \Xi} \max_{x_d^* \in \mathcal{M}_d^*} \left( \frac{\partial x_d^*}{\partial \beta^t} \right) \mathcal{M}^-(\xi) \left( \frac{\partial x_d^*}{\partial \beta^t} \right)^t.
$$

(1.2.18)

The reason we rely on the asymptotic distribution of the ratio estimator $\hat{x}_d^*$ for the first three design criteria is that the mean and variance of $\hat{x}_d^*$ do not exist (Fieller, 1932). This means most of the well-known design criteria can not be applied directly to the estimator $\hat{x}_d^*$. For our final design criterion in the univariate case, we consider the stochastic closeness of the fixed estimator $\hat{x}_d^*$ to the unknown parameter
Design Criterion 1.2.4 (Lsd) A design measure $\xi^* \in \Xi$ is called a Local stochastic domination optimal design for estimating $x^*_d$ at the value $\mu^*_d$ if and only if for all $\epsilon > 0$

$$P_{\xi^*}(|\hat{x}^*_d - \mu^*_d| > \epsilon) = \min_{\xi \in \Xi} P_{\xi}(|\hat{x}^*_d - \mu^*_d| > \epsilon)$$

(1.2.19)

where $|\cdot|$ is the usual Euclidean distance.

Sinha (1970) was the first to consider the stochastic domination design criterion. He used it for problems of allocating experimental units to treatment groups. In Chapter V we consider the stochastic domination criterion for making inferences on estimable functions of the regression parameters $\beta$, and we determine a sufficient condition for a stochastic domination optimal design to exist.

Design Criteria for the Multivariate Case

In this section we define the two design criteria for the multivariate case. Both criteria are real valued functions of the asymptotic dispersion matrix of $\hat{x}^*_d$ given in (1.2.15). Once again, to define the criteria, we let $\mathcal{X}$ be the design space for the computer experiment and $\Xi$ be the class of competing design measures defined on the Borel sets of $\mathcal{X}$. We also let $\pi^*$ be a given prior distribution for $x^*_d$. The two criteria that we define here require different assumptions for the prior $\pi^*$. The prior assumptions that make for each criterion are given in 3.2.1.
Design Criterion 1.2.5 (IavA) A design measure $\xi^* \in \Xi$ is called an Integrated asymptotic variance $A$-optimal design for estimating $x_d^*$ with respect to the prior $\pi^*$ if and only if it minimizes

$$E^* \left[ \text{tr} \left( \frac{\partial x_d^*}{\partial \beta^i} M^- (\xi^*) \left( \frac{\partial x_d^*}{\partial \beta^i} \right)^t \right) \right]. \quad (1.2.20)$$

Design Criterion 1.2.6 (IavD) A design measure $\xi^* \in \Xi$ is called an Integrated asymptotic variance $D$-optimal design for estimating $x_d^*$ with respect to the prior $\pi^*$ if and only if it minimizes

$$\left| E^* \left[ \left( \frac{\partial x_d^*}{\partial \beta^i} \right)^t M^- (\xi^*) \left( \frac{\partial x_d^*}{\partial \beta^i} \right) \right] \right|. \quad (1.2.21)$$

1.3 Preliminary Theorems and Lemmas

In this section we present the definitions and known results that we use to prove the results in the later chapters of this thesis. In Section 1.3.1 we provide relevant background information on experimental designs. This includes the equivalence theorem of Kiefer and Wolfowitz (1960) which we use repeatedly in the sequel. Section 1.3.2 contains pertinent properties and definitions for the family of elliptically distributed random variables and the statistical property of stochastic domination. Finally, Section 1.3.3 provides results for the class of completely symmetric matrices.
1.3.1 Experimental Designs

Let \( \mathcal{X} \) be a compact region in \( k \)-dimensional Euclidean space, and let \( \mathcal{B} \) be a Borel \( \sigma \)-field defined on \( \mathcal{X} \). In addition, define \( \Xi \) to be the class of all design measures that are defined on the sets of \( \mathcal{B} \). By definition, each \( \xi \in \Xi \) satisfies \( \int_{\mathcal{X}} \xi(dx) = 1 \).

The design measure \( \xi \in \Xi \) specifies the amount of the design mass that is to be placed on each Borel measurable set of \( \mathcal{X} \). When the only Borel measurable sets of \( \mathcal{X} \) that have positive measure with respect to \( \xi \) are the \( n \) points \( \{x_i\}_{i=1}^n \subset \mathcal{X} \), the design \( \xi \) can be represented as

\[
\xi = \left\{ \begin{array}{cccc}
x_1 & x_2 & \cdots & x_n \\
\xi(x_1) & \xi(x_2) & \cdots & \xi(x_n)
\end{array} \right\}
\]  

where \( \xi(x_i) \) is the design weight on the point \( x_i \). However, even in this simple form, the design \( \xi \) in (1.3.1) is not obtainable for a fixed sample size \( N \) unless \( N\xi(x_i) \) is an integer for each \( i \in \{1,2,\ldots,n\} \).

**Definition 1.3.1** A design that specifies the integer number of observations to be taken at each point \( x \in \mathcal{X} \), for a given sample size, is called exact. Otherwise, it is called continuous or approximate.

Every experimental design used in practice must be exact, so the experimental design problem is to determine an exact design that is, in some sense, optimal. However, the problem of determining an optimal design can be greatly simplified if we disregard the constraint of being exact and consider all possible design measures in \( \Xi \). The optimal design that is obtained may not be exact, but near-optimal exact designs can often be achieved by integer approximation of an optimal continuous design.
Assume the model for the relationship between the real-valued random observable \( Y \) and the vector of regressor variables \( \mathbf{x} \) is given by

\[
Y = \mathbf{f}(\mathbf{x})^T \beta + \epsilon
\]  

(1.3.2)

where \( \mathbf{f}(\mathbf{x})^T = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_p(\mathbf{x})) \) is a \( p \)-dimensional vector of known continuous, linearly independent regression functions of \( \mathbf{x} \in \mathcal{X}, \beta \in \mathbb{R}^p \) is the unknown vector of regression parameters, and \( \epsilon \) has \( \mathbb{E}[\epsilon] = 0 \) and \( \text{Var}[\epsilon] = \sigma^2 \). Suppose an \( N \) point experimental design \( \xi_N \) specifies that independent observations of \( Y \) are to be collected at the sites \( \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\} \), with at least \( p \) distinct sites. Then, the least squares estimator of \( \beta \) is \( \hat{\beta} = (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T \mathbf{Y} \) where \( \mathbf{F} \) is the \( N \times p \) matrix with \( i^{th} \) row \( \mathbf{f}(\mathbf{x}_i)^T \) and \( \mathbf{Y} = (Y_1, Y_2, \ldots, Y_N)^T \) is the vector of unobserved \( Y \)'s. The variance-covariance matrix of the unbiased estimator \( \hat{\beta} \) is \( \sigma^2 (\mathbf{F}^T \mathbf{F})^{-1} \). Hence, the mean squared error of the estimator \( \hat{\beta} \) depends on the design only through the \( p \times p \) symmetric nonnegative definite matrix \( \mathbf{F}^T \mathbf{F} \).

**Definition 1.3.2** For the statistical model in (1.3.2) and a design \( \xi \in \Xi \), the information matrix \( \mathbf{M}(\xi) \) for the vector \( \beta \) is the \( p \times p \) symmetric nonnegative definite matrix defined by

\[
\mathbf{M}(\xi) = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x}) \mathbf{f}(\mathbf{x})^T \xi(\mathbf{x}) \, d\mathbf{x} = \mathbb{E}[\mathbf{f}(\mathbf{x}) \mathbf{f}(\mathbf{x})^T].
\]  

(1.3.3)

To determine optimal designs, we minimize functions, \( \Psi(\mathbf{M}(\cdot)) \), of the information matrix over the class of competing designs \( \Xi \). The \( \Psi(\cdot) \) that we defined in Section 1.2.4 are convex and differentiable, i.e., directional derivatives exist. According to the equivalence theorem of Kiefer and Wolfowitz (1960), continuous designs that are
optimal with respect to these $\Psi$ also satisfy a second criterion. Often the second criterion is easier to verify thus simplifying the construction and verification of an optimal design. Before we present the equivalence theorem of Kiefer and Wolfowitz, we define the Fréchet directional derivative.

**Definition 1.3.3** Let $M_1$ be positive definite matrix and $M_2$ be a nonnegative definite matrix. The Fréchet derivative of $\Psi$ at $M_1$ in the direction of $M_2$ is given by

$$F_{\Psi}(M_1, M_2) = \lim_{\epsilon \to 0^+} \frac{\Psi((1-\epsilon)M_1 + \epsilon M_2) - \Psi(M_1)}{\epsilon}.$$  \hspace{1cm} (1.3.4)

In the equivalence theorem the matrix $M_2$ in (1.3.4) will be a matrix of the form $M(\xi_x)$ where $\xi_x$ is a design that puts unit mass on the point $x$. Using (1.3.3), we can express this design as $M(\xi_x) = f(x)f(x)^t$.

**Theorem 1.3.4 (Equivalence Theorem)** Let $\Psi$ be a real-valued convex design function for which the Fréchet derivative exists, then the following are equivalent.

i.) The design $\xi^* \in \Xi$ minimizes $\Psi(M(\xi^*))$.

ii.) For every point $x \in \mathcal{X}$, $F_{\Psi}(M(\xi^*), M(\xi_x)) \geq 0$.

iii.) For every point $x \in \mathcal{X}$ such that $\xi^*(x) > 0$, $F_{\Psi}(M(\xi^*), M(\xi_x)) = 0$.

We use the results of this theorem to determine optimal designs in Chapter II and Chapter III. Therefore, we must know the value of the Fréchet derivative for each design criterion. We conclude this section with two examples of the Fréchet derivative that will encompass all the criteria that we consider.
Example 1.3.1 (IavA Fréchet derivative) The design function (1.2.20) for the Integrated asymptotic variance A-optimal design criterion is equivalent to

$$\Psi(M(\xi)) = \text{tr} \left[ M^{-1}(\xi) E^\pi \left[ \left( \frac{\partial x_d^*}{\partial \beta^i} \right)^t \left( \frac{\partial x_d^*}{\partial \beta^i} \right) \right] \right] \quad (1.3.5)$$

for all $\xi \in \Xi$. The Fréchet derivative of $\Psi$ at the positive definite matrix $M_1$ in the direction of the nonnegative definite matrix $M_2$ is

$$F_\Psi(M_1, M_2) = \lim_{\epsilon \to 0^+} \epsilon^{-1} \text{tr} \left[ (M_1 - \epsilon (M_1 - M_2))^{-1} E^\pi \left[ \left( \frac{\partial x_d^*}{\partial \beta^i} \right)^t \left( \frac{\partial x_d^*}{\partial \beta^i} \right) \right] \right] -$$

$$\epsilon^{-1} \text{tr} \left[ M_1^{-1} E^\pi \left[ \left( \frac{\partial x_d^*}{\partial \beta^i} \right)^t \left( \frac{\partial x_d^*}{\partial \beta^i} \right) \right] \right]$$

$$= \lim_{\epsilon \to 0^+} \epsilon^{-1} \text{tr} \left[ (I - \epsilon (I - M_1^{-1} M_2))^{-1} M_1^{-1} E^\pi \left[ \left( \frac{\partial x_d^*}{\partial \beta^i} \right)^t \left( \frac{\partial x_d^*}{\partial \beta^i} \right) \right] \right] -$$

$$\epsilon^{-1} \text{tr} \left[ M_1^{-1} E^\pi \left[ \left( \frac{\partial x_d^*}{\partial \beta^i} \right)^t \left( \frac{\partial x_d^*}{\partial \beta^i} \right) \right] \right] \quad (1.3.6)$$

Replacing the matrix $(I - \epsilon (I - M_1^{-1} M_2))^{-1}$ in (1.3.6) with $I + \epsilon (I - M_1^{-1} M_2) + o(\epsilon)$ where $o(\epsilon)/\epsilon \xrightarrow{\epsilon \to 0} 0$, its Taylor series expansion, we get

$$F_\Psi(M_1, M_2) = \lim_{\epsilon \to 0^+} \frac{\text{tr} \left[ (I - M_1^{-1} M_2) M_1^{-1} E^\pi \left[ \left( \frac{\partial x_d^*}{\partial \beta^i} \right)^t \left( \frac{\partial x_d^*}{\partial \beta^i} \right) \right] + o(\epsilon) \right]}{\epsilon}$$

$$= \text{tr} \left[ (M_1^{-1} - M_1^{-1} M_2 M_1^{-1}) E^\pi \left[ \left( \frac{\partial x_d^*}{\partial \beta^i} \right)^t \left( \frac{\partial x_d^*}{\partial \beta^i} \right) \right] \right]. \quad (1.3.7)$$
For given $x \in \mathcal{X}$, the Fréchet derivative of $\Psi$ at $M(\xi^*)$ for the IavA optimal design $\xi^*$ in the direction $f(x)f(x)^t$ equals
\[
\text{tr} \left[ \left( M(\xi^*)^{-1} - M(\xi^*)^{-1}f(x)f(x)^tM(\xi^*)^{-1} \right) E_{\xi^*}^\beta \left[ \left( \frac{\partial x_i^*}{\partial \beta^t} \right)^t \left( \frac{\partial x_i^*}{\partial \beta^t} \right) \right] \right]. \tag{1.3.8}
\]
Hence, Condition ii.) of Theorem 1.3.4 can be written as
\[
f(x)^tM(\xi^*)^{-1}E_{\xi^*}^\beta \left[ \left( \frac{\partial x_i^*}{\partial \beta^t} \right)^t \left( \frac{\partial x_i^*}{\partial \beta^t} \right) \right] M(\xi^*)^{-1}f(x) \leq \text{tr} \left[ M(\xi^*)^{-1}E_{\xi^*}^\beta \left[ \left( \frac{\partial x_i^*}{\partial \beta^t} \right)^t \left( \frac{\partial x_i^*}{\partial \beta^t} \right) \right] \right], \tag{1.3.9}
\]
for all $x \in \mathcal{X}$. Condition iii.) of Theorem 1.3.4 says that equality in (1.3.9) is obtained at every support point of the optimal design $\xi^*$. □

Example 1.3.2 (IavD Fréchet derivative) Since $\log(\cdot)$ is a monotone increasing function, minimizing the design function (1.2.21) for the Integrated asymptotic variance D-optimal design criterion is equivalent to minimizing
\[
\Psi(M(\xi)) = \log \left| E_{X^*}^\beta \left[ \left( \frac{\partial x_i^*}{\partial \beta^t} \right) M^{-1}(\xi) \left( \frac{\partial x_i^*}{\partial \beta^t} \right)^t \right] \right| \tag{1.3.10}
\]
over $\Xi$. Let $G$ be the gradient matrix $(\partial x_i^*/\partial \beta^t)$, and let $H_1 = GM_1^{-1}G^t$ for given positive definite matrix $M_1$. The Fréchet derivative of $\Psi$ at $M_1$ in the direction of the nonnegative definite matrix $M_2$, $F_\Psi(M_1, M_2)$, is
\[
= \lim_{\epsilon \to 0^+} \frac{\log \left| E_{X^*}^\beta \left[ G (\epsilon M_1 + \epsilon M_2)^{-1} G^t \right] \right| - \log |H_1|}{\epsilon} = \frac{\log \left| E_{X^*}^\beta \left[ G (I - \epsilon (I - M_1^{-1}M_2))^{-1} M_1^{-1}G^t \right] \right|}{\epsilon} \tag{1.3.11}
\]
Replacing the matrix \((I - \epsilon(I - M_1^{-1}M_2))^{-1}\) in (1.3.11) with \(I + \epsilon(I - M_1^{-1}M_2) + o(\epsilon)\), its Taylor series expansion, we get \(F_\Psi(M_1, M_2)\) is
\[
= \lim_{\epsilon \to 0^+} \frac{\log \left| G \left[ G \left( I + \epsilon (I - M_1^{-1}M_2) \right) M_1^{-1}G' \right] H_1^{-1} + o(\epsilon) \right|}{\epsilon}
\]
\[
= \lim_{\epsilon \to 0^+} \frac{\log \left| I + \epsilon \left( I - E^{\ast\ast} [G M_1^{-1}M_2M_1^{-1}G'] H_1^{-1} + o(\epsilon) \right) \right|}{\epsilon}
\]
\[
= \lim_{\epsilon \to 0^+} \frac{\log \left\{ 1 + \epsilon \text{tr} \left[ I - E^{\ast\ast} [G M_1^{-1}M_2M_1^{-1}G'] H_1^{-1} \right] + o(\epsilon) \right\}}{\epsilon}. \quad (1.3.12)
\]
Using the Taylor Series expansion \(\log(1 + x) \approx x - x^2\), we see
\[
F_\Psi(M_1, M_2) = k - \text{tr} \left[ E^{\ast\ast} \left[ G M_1^{-1}M_2M_1^{-1}G' \left( E^{\ast\ast} [G M_1^{-1}G'] \right)^{-1} \right] \right]. \quad (1.3.13)
\]
For given \(x \in X\), the Fréchet derivative of \(\Psi\) at \(M(\xi^*)\) for the lavD optimal design \(\xi^*\) in the direction \(f(x)f(x)'\) equals
\[
F_\Psi(M(\xi^*), ff') = k - \text{tr} \left[ E^{\ast\ast} \left[ G M(\xi^*)^{-1}ff'M(\xi^*)^{-1}G' \left( E^{\ast\ast} [G M(\xi^*)^{-1}G'] \right)^{-1} \right] \right]. \quad (1.3.14)
\]
Hence, Condition ii.) of Theorem 1.3.4 can be written as
\[
f' M(\xi^*)^{-1} E^{\ast\ast} \left[ \left( \frac{\partial x^*}{\partial \beta} \right)' \left( E^{\ast\ast} \left[ \left( \frac{\partial x^*}{\partial \beta} \right) M(\xi^*)^{-1} \left( \frac{\partial x^*}{\partial \beta} \right)' \right] \right)^{-1} \left( \frac{\partial x^*}{\partial \beta} \right)' \right] M(\xi^*)^{-1} f \leq k
\]
for all \(x \in X\). Condition iii.) of Theorem 1.3.4 says that equality in (1.3.15) is obtained at every support point of the optimal design \(\xi^*\). □
1.3.2 Elliptical Distributions and Stochastic Domination

The results and definitions of this section are used in Chapter V where we examine a stochastic design criterion similar to (1.2.19) for making inferences about $K\beta$ where $\beta$ is the regression parameter vector in (1.3.2) and $K$ is a known $k \times p$ matrix of rank $k$. Since the design criterion that we define in Chapter V is stochastic, the classical linear model in (1.3.2) with only moment assumptions is not sufficient for constructing optimal designs. Therefore, we assume the distribution of the error term $\epsilon$ belongs to a known family of distributions. In this section we define the family of distributions that we assume for $\epsilon$. We also present known results for this family that are relevant to the problem given in Chapter V, but first we define the majorization (Schur) partial ordering of vectors and the associated functional property of Schur-concavity (convexity).

**Definition 1.3.5** Suppose $x, y \in (0, \infty)^n$ have ordered coordinates $0 < x^{(1)} \leq x^{(2)} \leq \cdots \leq x^{(n)}$ and $0 < y^{(1)} \leq y^{(2)} \leq \cdots \leq y^{(n)}$, respectively. The vector $x$ is said to weakly majorize $y$ if

$$\sum_{i=1}^{k} x^{(i)} \geq \sum_{i=1}^{k} y^{(i)} \text{ for } k = 1, 2, \ldots, n. \quad (1.3.16)$$

If, in addition, there is equality in (1.3.16) for $k = n$, then $x$ is said to majorize $y$.

**Definition 1.3.6** Suppose $x$ majorizes $y$. The function $g(\cdot)$ is Schur concave (convex) if $g(x) \geq g(y)$ ($g(x) \leq g(y)$).
**Definition 1.3.7** Let \( \eta \) be a vector in \( \mathbb{R}^d \) and \( \Sigma \) be a \( d \times d \) positive definite matrix, then a random vector \( Y \in \mathbb{R}^d \) is elliptically distributed if and only if the characteristic function of \( Y \) is of the form

\[
\Psi_Y(t) = \psi(t' \Sigma) \exp\{it' \eta\}.
\]  

(1.3.17)

**Definition 1.3.8** A random vector \( X \in \mathbb{R}^d \) is spherically distributed if and only if \( PX \) has the same distribution as \( Y \) for every real \( d \times d \) orthogonal matrix \( P \).

The family of elliptical distributions contains the family of normal distributions as well as heavier-tailed distributions such as the multivariate t and multivariate Cauchy. Hence, for some, but not all, members of the family of distributions that we assume, \( Y \) has finite first and second moments. When the density function \( f_Y(\cdot) \) of an elliptically distributed random variable \( Y \) exists, it is proportional to

\[
f_Y(y) \propto |\Sigma|^{-1/2} \phi\left( (y - \eta)' \Sigma^{-1} (y - \eta) \right).
\]  

(1.3.18)

(Kelker, 1970). An elliptical distribution is spherical if and only if \( \eta = 0_d \) and \( \Sigma = \sigma^2 I_d \) for some \( \sigma^2 > 0 \). When \( Y \) is spherically distributed, the density function in (1.3.18) is Schur concave (Eaton and Perlman, 1991). Hence, the following probability inequality of Tong (1982), which requires a Schur concave density function, is valid for all spherically distributed random vectors for which the density function exists.

**Theorem 1.3.9** If \( Y \in \mathbb{R}^d \) is a random variable with a Schur-concave density function \( f_Y \) and \( g(\cdot) \) is a positive, increasing and convex function, then the probability

\[
P \left\{ \sum_{i=1}^{d} \frac{Y_i^2}{\lambda_i} \leq 1 \right\}
\]  

(1.3.19)
is a Schur-concave function of \((g(\lambda_1), g(\lambda_2), \ldots, g(\lambda_d))\).

The remainder of this section is devoted to material that is needed for defining the design criterion of Chapter V. The actual definition is reserved for Section 5.1, but this section defines all of the components used in the definition. We begin by defining the concept of a stochastically smaller random variable. We also define the multivariate extension of the Euclidean distance in (1.2.19) that we use in our design criterion. Finally, we state the result that shows the least squares estimator \(K\hat{\beta}\) is the best linear unbiased estimator of \(K\beta\) in this set-up.

**Definition 1.3.10** For two real-valued positive random variables \(X\) and \(Y\), \(X\) is said to be stochastically smaller than \(Y\), denoted \(X <_d Y\), if and only if for every \(\epsilon > 0\)

\[
P\{X < \epsilon\} \geq P\{Y < \epsilon\}
\]

with strict inequality for at least one \(\epsilon > 0\).

**Definition 1.3.11** Let \(\delta(Y)\) be an estimator of a parameter \(\theta \in \mathbb{R}^d\). For a given \(d \times d\) real positive definite matrix \(D\), the generalized Euclidean error of \(\delta(Y)\) for estimating \(\theta\), with respect to \(D\), is given by

\[
|\delta(Y) - \theta|_D = \left( (\delta(Y) - \theta)^T D (\delta(Y) - \theta) \right)^{1/2}.
\]

For \(D = I_d\), the generalized Euclidean error reduces to the usual Euclidean error. Together the previous two definitions lead to a desirable property for an estimator \(\delta(\cdot)\) for \(\theta\), as given in the next definition.
Definition 1.3.12 The estimator $\delta_1(Y)$ for the parameter $\theta$ stochastically dominates the estimator $\delta_2(Y)$ under the generalized Euclidean error with respect to $D$ if for every value of $\theta \in \mathbb{R}^d$ and every $\epsilon > 0$

$$P\{|\delta_1(Y) - \theta|_D \leq \epsilon\} \geq P\{|\delta_2(Y) - \theta|_D \leq \epsilon\}$$

(1.3.22)

with $|\delta_1(Y) - \theta|_D <_d |\delta_2(Y) - \theta|_D$ for some $\theta \in \mathbb{R}^d$.

Hwang (1995) showed that if $\delta_1$ stochastically dominates $\delta_2$, then for a large class of loss functions the risk of $\delta_1$ is less than or equal to the risk of $\delta_2$ with strict inequality for some $\theta \in \mathbb{R}^d$. This means if there exists an estimator $\delta_1(\cdot)$ that stochastically dominates every other estimator within some class $\Delta$, then the estimator $\delta_1(\cdot)$ simultaneously minimizes a large class of loss functions over $\Delta$. The following result of Hwang is an analog of the Gauss-Markov Theorem. It shows that the least squares estimator stochastically dominates every other linear unbiased estimator.

Theorem 1.3.13 Consider a random sample $Y_1, Y_2, \ldots, Y_N$, from the model (1.3.2) where the vector of unobservable errors $\epsilon \in \mathbb{R}^N$ is spherically distributed. In addition, let $D$ be a given $p \times p$ positive definite matrix. The least squares estimator $\hat{\beta} = (F'F)^{-1}F'Y$ stochastically dominates every other linear unbiased estimator of $\beta$ under the generalized Euclidean error with respect to $D$.

The consequence of Theorem 1.3.13 is stronger than the result of the usual Gauss-Markov Theorem. However, the assumption that the errors are spherically distributed
is stronger than the assumptions in the Gauss-Markov Theorem. The result of Theorem 1.3.13 can be extended for making inferences on linear functions of the regression coefficients $\beta$ and for more general structures for the variance-covariance matrix.

**Corollary 1.3.14** Consider a random sample $Y_1, Y_2, \ldots, Y_N$, from the model (1.3.2) where the vector of unobservable errors $\mathbf{e} \in \mathbb{R}^N$ is elliptically distributed with $\mathbf{e} = \mathbf{0}_N$ and $\Sigma = \sigma^2 \Sigma_0$, $\Sigma_0$ known and $\sigma^2$ unknown. In addition, let $\mathbf{D}$ be a given $k \times k$ positive definite matrix. For a known $k \times p$ matrix $\mathbf{K}$ of rank $k$, the generalized least squares estimator $\mathbf{K} \hat{\beta}$ stochastically dominates every other linear unbiased estimator of $\mathbf{K} \beta$ under the generalized Euclidean error with respect to $\mathbf{D}$.

### 1.3.3 Completely Symmetric Matrices

We use properties of positive definite completely symmetric matrices to prove results in Chapter III. In this section we present the relevant properties of this class of matrices. Proofs of the lemmas and corollaries that we present in this section can be found in almost any text on matrix results (e.g., Graybill, 1969). We begin this section with the definition of a completely symmetric matrix. Let $\mathbf{I}_n$ be the $n \times n$ identity matrix and $\mathbf{J}_n = \mathbf{1}_n \mathbf{1}_n^T$ be the $n \times n$ matrix of ones.

**Definition 1.3.15** An $n \times n$ real symmetric matrix, $\mathbf{M}$, is completely symmetric if and only if it has the form

$$\mathbf{M} = a\mathbf{I}_n + b\mathbf{J}_n$$

(1.3.23)

where $a$ and $b$ are real-valued.
Completely symmetric matrices have equal entries along the diagonal and equal off-diagonal elements. In addition, as the next lemma shows, completely symmetric matrices have the same set of orthonormal eigenvectors. This well-known lemma is key to proving the remaining results of this section.

**Lemma 1.3.16** Every $n \times n$ real completely symmetric matrices is diagonalized by any set of orthonormal eigenvectors that includes the normalized vector of ones, $n^{-1/2}1_n$, and $n - 1$ orthonormal contrasts, $\{v_i\}_{i=1}^{n-1}$.

Note that $1_n^t v_i = 0$ for each $i \in \{1, \ldots, n - 1\}$. By (1.3.23), this implies that a completely symmetric matrix has at most two distinct eigenvalues: $a + nb$ with multiplicity 1 and $a$ with multiplicity $n - 1$. Since a matrix is positive definite if and only if all its eigenvalues are positive, the following corollary is immediate.

**Corollary 1.3.17** A completely symmetric matrix is positive definite if and only if $a > 0$ and $b > -a/n$.

Fix $n - 1$ orthonormal contrasts $\{v_i\}_{i=1}^{n-1}$ and let $R = (n^{-1/2}1_n, v_1, \ldots, v_{n-1})$. Each completely symmetric matrix $A$ can be spectrally decomposed as $A = RA_A R^t$ where $A_A$ is the diagonal matrix with the eigenvalues of $A$ along the diagonal and the eigenvalue of multiplicity one is the first entry on the diagonal. If we multiply the two diagonal matrices $A_A$ and $A_B$ for the completely symmetric matrices $A$ and $B$, the result will be a diagonal matrix of similar form. Since matrix multiplication is a commutative and associative on the set of diagonal matrices, the following two corollaries are immediate.
Corollary 1.3.18 Matrix multiplication for the set of completely symmetric matrices is commutative and associative.

Corollary 1.3.19 The set of $n \times n$ real positive definite completely symmetric matrices is closed under the operation of matrix multiplication.

If, in addition, we assume the completely symmetric matrix $A$ is positive definite, then the matrix $\Lambda_A^{1/2}$ for which $\Lambda_A^{1/2} \Lambda_A^{1/2} = \Lambda_A$ is positive definite with form similar to $\Lambda_A$. This leads to the following corollary.

Corollary 1.3.20 The square root matrix $A^{1/2} = RA_A^{1/2}R^t$ is a positive definite completely symmetric matrix.
CHAPTER II

Determining the Optimal Univariate Operating Conditions

2.1 Introduction

In this chapter we consider the problem of designing computer experiments so as to accurately determine the level \( x^*_d \) of the controllable factor \( x_d \) that minimizes the expected integrated loss function in (1.2.1). The goal is to find, for a fixed sample size \( N \), an optimal collection of input configurations at which the computer code should be run. In this chapter optimal means the set of inputs that allow the most accurate estimation of the parameter \( x^*_d \) where accuracy is defined by the univariate criteria in Section 1.2.4.

We assume realizations of the output \( y \) are generated by a computer code that requires input values for a single controllable factor \( x_d \) and \( k_e \) environmental factors, \( x_e \). In addition, we regard the deterministic output \( y(x_d, x_e) \) as a realization of the stochastic process \( Y(x_d, x_e) \) given in (1.2.3). When there is only a single controllable variable, the model in (1.2.3) can be written

\[
Y(x_d, x_e) = \beta_0 + \beta_d x_d + \beta_{d1} x_d^2 + \beta' e x_e + x_d \beta' f x_e + Z(x_d, x_e)
\]
where $\beta_0, \beta_{d_1}, \beta_{d_11} \in \mathbb{R}$, and $\beta_c, \beta_f \in \mathbb{R}^{k_e}$ are unknown regression coefficients, and the error term $Z(\cdot, \cdot)$ is a mean zero Gaussian process with unknown constant variance $\sigma^2$ and zero correlations between distinct input sites. The coefficient for the quadratic term $x_d^2$ is assumed to be positive, i.e. $\beta_{d_11} > 0$. This implies the integrated expected loss function $L(x_d)$ in (1.2.4) is a strictly convex quadratic function of $x_d$. Hence, $L(\cdot)$ attains its minimum value at the unique level

$$x_d^* = -\frac{\beta_{d_1} + \beta_f^\prime \mu_c}{2\beta_{d_11}}$$

(2.1.2)

of the control factor $x_d$ where $\mu_c = E^\prime [x_c]$ is known.

With the formulation of the problem given in the previous paragraphs, designing experiments for the parameter design problem is similar to the problem of designing experiments to estimate the optimal point of a univariate quadratic regression model. However, designs for estimating the optimal point of the univariate quadratic regression model do not typically consider the environmental factors $x_e$. Optimal designs for estimating the maximal point of a quadratic response have been determined by Murty and Studden (1972), Mandal (1978) and Buonaccorsi and Iyer (1986).

The estimand, $x_d^*$, is a ratio of linear combinations of the unknown regression parameters in (2.1.1). Buonaccorsi (1985) showed that under the assumption of normal errors an unbiased estimator of $x_d^*$ exists if and only if both the variance $\sigma^2$ and the sign of $\beta_{d_11}$ are known. Since $\sigma^2$ is unknown, an unbiased estimator of $x_d^*$ does not exist. We estimate $x_d^*$ using the maximum likelihood estimator $\hat{x}_d^*$ in (1.2.11). In this
univariate case, this estimator has the form

\[ \hat{x}_d^* = -\frac{\hat{\beta}_d + \hat{\beta}_f \mu_c}{2\hat{\beta}_{d11}} \]  

(2.1.3)

where \(\hat{\beta}_d\), \(\hat{\beta}_{d11}\), and \(\hat{\beta}_f\) are the least squares estimators of \(\beta_d\), \(\beta_{d11}\), and \(\beta_f\), respectively.

In this chapter we determine optimal designs for each of the univariate criteria defined in Section 1.2.4. In Section 2.5 we find both a local and an integrated asymptotic variance optimal design, while Section 2.6 contains the determination of a minimax asymptotic variance optimal design. In Section 2.7 we show the equivalence of an approximate version of the local stochastic domination criterion and the local asymptotic variance criterion. A method of construction for an optimal design is given in Section 2.8, and examples for each of the criteria are presented in Section 2.9. We begin this chapter with the definition of the design region \(\mathcal{X}\) (Section 2.2) and a subsequent reparameterization of the model (Section 2.3). Section 2.4 gives the value of the asymptotic variance of \(\hat{x}_d^*\).

### 2.2 Defining the Design Region \(\mathcal{X}\)

In this section we define the common support space \(\mathcal{X}\), for the set of all competing designs. The space \(\mathcal{X}\) is the Cartesian product of the support spaces \(\mathcal{X}_d\) and \(\mathcal{X}_e\) for the control and environmental variables, respectively. We allow generality in the definition of \(\mathcal{X} = \mathcal{X}_d \times \mathcal{X}_e\) by requiring only that \(\mathcal{X}_d \subset \mathbb{R}\) and \(\mathcal{X}_e \subset \mathbb{R}^{k_e}\) are the respective images of nonsingular affine transformations of the symmetric regions.
The region $X$ is the image of a nonsingular affine transformation of the symmetric cube $[-1, 1]^{k+1}$. Thus, it is a closed, bounded and connected region in $\mathbb{R}^{k+1}$.

### 2.2.1 Defining $X_d$

When defining the design space $X_d$, we assume the optimal operating condition $x_d^*$ is in the interior of $X_d$. In addition, since $x_d^*$ is a nonlinear function of the regression coefficients in (2.1.1), we also assume a priori knowledge about the value of $x_d^*$. These a priori assumptions will, in turn, impact the way we define $X_d$. For the univariate design criteria given in Section 1.2.4, we add knowledge of $x_d^*$ into the design problem in two different ways: by means of moment conditions for $x_d^*$ and by specifying the region of support for $x_d^*$.

When a priori knowledge about $x_d^*$ is added to the design problem by means of a prior distribution $\pi_d^*$ for $x_d^*$, we assume the following moment conditions for $x_d^*$.

**Assumptions 2.2.1** For the model in (2.1.1), assume the scalar $x_d^*$ in (2.1.2) is random with distribution $\pi_d^*$ for which the following hypotheses are true.

1. $\mu_d^* = E^{\pi_d^*}[x_d^*] \in \mathbb{R}$ is known.
2. $\sigma_d^2 = E^{\pi_d^*}[(x_d^* - \mu_d^*)^2] \in \mathbb{R}$ is known and nonnegative.

In this case a natural choice for the design space of the control variables would be $X_d = \{s_d\sigma_d u + \mu_d^* : u \in [-1, 1]\}$, where $s_d > 0$ is a specified sizing scalar. However, when we assume $\sigma_d = 0$, which is required for the Lav design criterion, the space $X_d$
is only a single point, and the coefficients $\beta_{d1}$ and $\beta_{d11}$ in (2.1.1) can not be estimated. Therefore, we consider the Lav design criterion as a limiting case ($\sigma_d \searrow 0$) of the Lav design criterion.

It is convenient to let $s_d = \gamma/\sigma_d$ where $\gamma > 0$ is a specified constant. Then, if $\sigma_d > 0$ is fixed and known (Lav criterion), $s_d$ is a known quantity, and the design space for $x_d$ can be given by

$$\mathcal{X}_d = \{ \gamma u + \mu_d^*: u \in [-1,1] \}. \quad (2.2.1)$$

Similarly, as $\sigma_d \searrow 0$ (Lav criterion), $s_d \nearrow \infty$ at such a rate that $s_d\sigma_d = \gamma$, so that the design space for $x_d$ for the Lav criterion can also be given by (2.2.1).

When a priori knowledge of the parameter $x_d^*$ is added to the design problem by defining an interval of support $\mathcal{M}_d^*$, the assumptions that we make about $x_d^*$ are as follows.

**Assumptions 2.2.2** For the model in (2.1.1), assume $\mathcal{M}_d^*$, a closed finite interval of the real line, is the support space for the prior $\pi_d^*$ of $x_d^*$.

In this case, a natural choice for the design space of the controllable variables is $\mathcal{X}_d = \mathcal{M}_d^*$. However, we provide slightly more flexibility in the definition of $\mathcal{X}_d$. Let $m_d^*$ be the midpoint of the region $\mathcal{M}_d^*$, and let $\gamma > 0$ be a specified scalar. The design region for the control variable $x_d$ for the Mav design criterion is given by

$$\mathcal{X}_d = \{ \gamma u + m_d^*: u \in [-1,1] \}. \quad (2.2.2)$$
The arbitrariness of the constant \( \gamma \) in (2.2.2) means that either \( \mathcal{X}_d \subseteq \mathcal{M}_d^\ast \) or \( \mathcal{M}_d^\ast \subseteq \mathcal{X}_d \) can be true. In either case, the design space \( \mathcal{X}_d \) is centered about the midpoint \( m_d^\ast \) of the interval \( \mathcal{M}_d^\ast \).

For both methods of adding prior knowledge of \( x_d^\ast \) into the design problem, the resulting design space \( \mathcal{X}_d \) is an interval that is centered about a hypothesized value (\( \mu_d^\ast \) or \( m_d^\ast \)) for \( x_d^\ast \). This means two things. First, the design space \( \mathcal{X}_d \) is the image of \([-1, 1]\) for some nonsingular mapping. Second, we can use a single reparameterization of the model that will be valid for both methods of adding prior knowledge.

### 2.2.2 Defining \( \mathcal{X}_e \)

To complete the definition of the support space \( \mathcal{X} \), we define the design space \( \mathcal{X}_e \) for the random environmental variables \( x_e \). Recall the random vector \( x_e \) has known probability distribution \( \pi_e \). In addition, we assume that the first two moments of \( x_e \) with respect to \( \pi_e \) are known. Let \( \mu_e = E^{\pi_e}[x_e] \) and let \( \Sigma_e = E^{\pi_e}[(x_e - \mu_e)(x_e - \mu_e)^t] \), a positive definite matrix. Then, the design space \( \mathcal{X}_e \) is given by

\[
\mathcal{X}_e = \{ s_e \Sigma_e^{1/2} v + \mu_e : v \in [-1, 1]^k \} 
\]

where \( \Sigma_e^{1/2} \Sigma_e^{1/2} = \Sigma_e \) and \( s_e \) is a user-supplied scaling constant.

### 2.3 Reparameterizing the Model

In Section 2.2 we defined the design space \( \mathcal{X} \) with sufficient generality to ensure that the proposed set-up is applicable for a variety of situations. It would be arduous to
consider the parameter design problem in each of these situations separately. Therefore, in this section we define a reparameterization of the model in (2.1.1) that will allow us to determine optimal designs exclusively on the canonical region $[-1,1]^{kr+1}$. Since we imposed certain restrictions on the space $X$, any optimal design obtained on $[-1,1]^{kr+1}$ can be mapped directly to an optimal design on $X$.

We defined the support space for the set of competing designs by $X = X_d \times X_c$ where $X_c = \{s_c\Sigma_e^{1/2}v + \mu_c : v \in [-1,1]^{kr}\}$ and $X_d = \{\gamma u + \mu_d^* : u \in [-1,1]\}$ with the midpoint $m_d^*$ of $M_d^*$ replacing $\mu_d^*$ in $X_d$ for the Mav design criterion. If we replace $x_d$ and $x_c$ in (2.1.1) with $\gamma u + \mu_d^*$ (or $\gamma u + m_d^*$) and $s_c\Sigma_e^{1/2}v + \mu_c$, respectively, the model for the computer output expresses $Y$ as a function of the transformed inputs $(u,v) \in [-1,1]^{kr+1}$. That is,

$$Y(u,v) = \alpha_0 + \alpha_{d_1} u + \alpha_{d_11} u^2 + \alpha_{c}^t v + u\alpha_{c}^t v + Z(u,v) \quad (2.3.1)$$

where

$$\alpha_0 = \beta_0 + \beta_{d_1}\mu_d^* + \beta_{d_11}\mu_d^{*2} + \mu_d^*\beta_{d_1}^t\mu_c \in \mathbb{R} \quad (2.3.2)$$

$$\alpha_{d_1} = \gamma(\beta_{d_1} + 2\beta_{d_11}\mu_d^* + \beta_{d_1}^t\mu_c) \in \mathbb{R} \quad (2.3.3)$$

$$\alpha_{d_11} = \gamma^2\beta_{d_11} \in \mathbb{R} \text{ is positive} \quad (2.3.4)$$

$$\alpha_{c} = s_c\Sigma_e^{1/2}(\beta_c + \mu_d^*\beta_{d_1}) \in \mathbb{R}^{kr} \quad (2.3.5)$$

$$\alpha_{d_1} = s_c\Sigma_e^{1/2}\beta_{d_1} \in \mathbb{R}^{kr} \quad (2.3.6)$$

The error term $Z(\cdot,\cdot)$ in (2.3.1) is still a mean zero Gaussian process with constant variance $\sigma^2 > 0$ and zero correlations between pairs of distinct input sites.
2.3.1 Reparameterizing the Parameter Design Problem

The vector $v \in [-1,1]^k$ equals $s_e^{-1}\Sigma_e^{-1/2}(x_e - \mu_e)$; therefore, it is random with probability law dictated by the known distribution $\pi_e$ of $x_e$. The expected value of $v$ is

$$E^\pi[v] = s_e^{-1}\Sigma_e^{-1/2}E^\pi[x_e - \mu_e] = 0_k,$$

and its variance equals

$$E^\pi[vv'] = s_e^{-2}\Sigma_e^{-1/2}E^\pi[(x_e - \mu_e)(x_e - \mu_e)']\Sigma_e^{-1/2} = s_e^{-2}I_k.$$

Using (2.3.7), (2.3.8) and the model in (2.3.1), the expected integrated loss function in (1.2.4) can be expressed as a function of the standardized control variable $u$, i.e.,

$$L(u) = \alpha_0 + \alpha_{d1}u + \alpha_{d1}u^2.$$  \hspace{1cm} (2.3.9)

Since $\alpha_{d1} = \gamma^2\beta_{d1}$ where $\beta_{d1} > 0$, the function $L(u)$ is minimized, uniquely, at the point

$$u^* = -\frac{\alpha_{d1}}{2\alpha_{d1}} = \gamma^{-1}(x_d^* - \mu_d^*).$$ \hspace{1cm} (2.3.10)

The second equality in (2.3.10) shows that the minimizing point $u^* \in [-1,1]$ is the image of $x_d^* \in \mathcal{X}_d$ for the inverse of the mapping that defines $\mathcal{X}_d$. Hence, $u^*$ is random with probability law governed by the prior $\pi_d^*$. The expected value of $u^*$ is

$$E^\pi_d[u^*] = \gamma^{-1}E^\pi_d[x_d^* - \mu_d^*] = 0,$$

and the variance is

$$E^\pi_d[u^{*2}] = \gamma^{-2}E^\pi_d[(x_d^* - \mu_d^*)^2] = \frac{\sigma_d^2}{\gamma^2} = s_d^{-2}.$$ \hspace{1cm} (2.3.12)
see Section 2.2.1. By invariance, the maximum likelihood estimator of \( u^* \) is

\[
\hat{u}^* = -\frac{\hat{\alpha}_{d_1}}{2\hat{\alpha}_{d_{11}}} = \gamma^{-1}(\hat{x}_d^* - \mu_d^*). \tag{2.3.13}
\]

### 2.4 The Asymptotic Variance of the MLE

In this section we find the value of the asymptotic variance for the estimator \( \hat{u}^* \). The value of the asymptotic variance of \( \hat{u}^* \) depends on the value of the regression parameters in (2.3.1). Hence, either the distributional hypotheses in Assumptions 2.2.1 or Assumptions 2.2.2 must be applied to the standardized variable \( u^* \) to obtain optimal designs.

Define \( \Xi_1 \) be the set of all design measures with finite support that are defined on the Borel sets of \([-1, 1]^{k_r+1} \), and let \( \xi \in \Xi_1 \). Under the distributional assumptions that we imposed on the stochastic error term \( Z(\cdot, \cdot) \) in (2.3.1), the maximum likelihood estimator \( \hat{u}^* \) is asymptotically Gaussian distributed (Buonaccorsi, 1985). The asymptotic variance of \( \hat{u}^* \) is proportional to the scalar

\[
\left( \frac{\partial u^*}{\partial \alpha^t} \right) M(\xi)^- \left( \frac{\partial u^*}{\partial \alpha^t} \right)^t \tag{2.4.1}
\]

where \( M(\xi)^- \) is the Moore-Penrose generalized inverse of the information matrix \( M(\xi) \) and \( \partial u^*/\partial \alpha^t \) is the gradient of \( u^* \) with respect \( \alpha^t = (\alpha_0, \alpha_{d_1}, \alpha_{d_{11}}, \alpha_{e}, \alpha_{\ell}) \). Calculation gives

\[
\left( \frac{\partial u^*}{\partial \alpha^t} \right) = \left( 0, -1, -u^*, \frac{0_{k_e}}{\alpha_{d_{11}}}, \frac{0_{k_e}}{\alpha_{d_{11}}} \right). \tag{2.4.2}
\]
Hence, the asymptotic variance depends on the values of \( \alpha_{d_{11}} \) and \( u^* \). The parameter \( \alpha_{d_{11}} \) can be factored out of the asymptotic variance, so knowledge of its value is not necessary for determining optimal designs. This explains why the a priori knowledge in Section 2.2.1 concerned only the parameter \( x_d^* \) (resp. \( u^* \)).

### 2.5 Iav and Lav Optimal Designs

In this section we determine sufficient conditions for design measures \( \xi^* \), defined on \([-1, 1]^{k+1} \), to minimize the Iav and Lav design criteria of Section 1.2.4. We assume the parameter of interest \( u^* \) is a random variable with probability law dictated by a distribution function \( \pi_d^* \) for which the hypotheses in Assumptions 2.2.1 are satisfied.

The goal of this section is to find conditions for design measures \( \xi \in \Xi_1 \) to minimize

\[
E^{\pi_d^*} \left[ \left( \frac{\partial u^*}{\partial \alpha^t} \right) M^{-1}(\xi) \left( \frac{\partial u^*}{\partial \alpha^t} \right)^t \right]
\]

over the set \( \Xi_1 \). When \( \pi_d^* \) is a degenerate distribution function, (2.5.1) is equivalent to the left hand side of (1.2.16) thus justifying the joint treatment of the Iav and Lav criteria in this section.

It is equivalent to minimize the quantity

\[
\text{tr} \left\{ M^{-1}(\xi) E^{\pi_d^*} \left[ \left( \frac{\partial u^*}{\partial \alpha^t} \right)^t \left( \frac{\partial u^*}{\partial \alpha^t} \right) \right] \right\}
\]

over the set of design measures in \( \Xi_1 \). The quantity in (2.5.2) is the univariate version of the convex and differentiable design function in Example 1.3.1. Therefore, we can use the results of Theorem 1.3.4 to determine an optimal design.
Theorem 2.5.1 (Iav Optimal Design) Assume the model in (2.3.1) and let $\pi^*_d$ be a probability distribution for $u^*$ for which the hypotheses in Assumptions 2.2.1, with $\sigma_d > 0$, hold. A design $\xi^* \in \Xi_1$ is an integrated asymptotic variance optimal design with respect to the prior $\pi^*_d$ if

i.) $E^\xi u = E^\xi u^3 = 0,$

ii.) $E^\xi u^2 = E^\xi u^4 = 1 + 4s_d^{-2} - \sqrt{4s_d^{-2}(1 + 4s_d^{-2})},$ and

iii.) for each $j \in \{1, 2, \ldots, k_r\}$, $E^\xi v_j = E^\xi v_j u = E^\xi v_j u^2 = 0.$

Proof of Theorem 2.5.1: For the model in (2.3.1), let

$$f^t = (1, u, u^2, v^t, uv^t)$$ (2.5.3)

where $u$ is in $[-1, 1]$ and $v \in [-1, 1]^{k_r}$. According to Condition ii.) of Theorem 1.3.4, a design measure $\xi \in \Xi_1$ is Iav optimal with respect to the prior $\pi^*_d$ if and only if

$$f^t M(\xi)^{-E^\pi_d} \left[ \left( \frac{\partial u^*}{\partial \alpha^t} \right)^t \left( \frac{\partial u^*}{\partial \alpha^t} \right) \right] M(\xi)^{-t} f \leq \text{tr} \left\{ M^{-}(\xi) E^\pi_d \left[ \left( \frac{\partial u^*}{\partial \alpha^t} \right)^t \left( \frac{\partial u^*}{\partial \alpha^t} \right) \right] \right\}$$ (2.5.4)

for every $(u, v) \in [-1, 1]^{k_r+1}$. We only need to verify (2.5.4) for the candidate design $\xi^*$ to prove the theorem.

Using the three hypotheses of Theorem 2.5.1, the information matrix for the parameter vector $\alpha$ for the design measure $\xi^*$ is equal to

$$M(\xi^*) = \int (1, u, u^2, v^t, uv^t)(1, u, u^2, v^t, uv^t)^t d\xi^*(u, v)$$
\[
\begin{pmatrix}
1 & 0 & r & 0_{2k_e} \\
0 & r & 0 & 0_{2k_e} \\
r & 0 & r & 0_{2k_e} \\
0_{2k_e} & 0_{2k_e} & 0_{2k_e} & V_u^t V_u
\end{pmatrix}
\]

where

\[r = E^t u^2 = 1 + 4s_d^{-2} - \sqrt{4s_d^{-2}(1 + 4s_d^{-2})}\] (2.5.6)

and

\[V_u^t V_u = \int \left( \begin{pmatrix} v & u \end{pmatrix}^t \begin{pmatrix} v & u \end{pmatrix} \right) d\xi^* \left( \begin{pmatrix} u \\ v \end{pmatrix} \right),\] (2.5.7)

For fixed finite \(s_d^{-2} > 0, \ r \in (0, 1).\) This means the upper \(3 \times 3\) submatrix is nonsingular, and the coefficients \(\alpha_0, \alpha_{d_1},\) and \(\alpha_{d_{11}}\) are all estimable. The Moore-Penrose generalized inverse of the information matrix in (2.5.5) is

\[
M(\xi^*)^{-1} = \begin{pmatrix}
\frac{1}{1-r} & 0 & -\frac{1}{1-r} & 0_{2k_e} \\
0 & \frac{1}{r} & 0 & 0_{2k_e} \\
-\frac{1}{1-r} & 0 & \frac{1}{r-r^2} & 0_{2k_e} \\
0_{2k_e} & 0_{2k_e} & 0_{2k_e} & (V_u^t V_u)^{-1}
\end{pmatrix}
\] (2.5.8)

where \((V_u^t V_u)^{-1}\) is the Moore-Penrose generalized inverse for the matrix \(V_u^t V_u\) in (2.5.7).

The matrix \(E_u^\pi \left[ (\partial u^* / \partial \alpha^t) (\partial u^* / \partial \alpha^t) \right],\) on the right hand side of (2.5.4), equals

\[
E_u^\pi \left[ \begin{pmatrix} \partial u^* \\ \partial \alpha^t \end{pmatrix} \begin{pmatrix} \partial u^* \\ \partial \alpha^t \end{pmatrix} \right] = \alpha_{d_{11}}^{-2} \begin{pmatrix}
0 & 0 & 0 & 0_{2k_e} \\
0 & \frac{1}{4} & 0 & 0_{2k_e} \\
0 & 0 & s_d^{-2} & 0_{2k_e} \\
0_{2k_e} & 0_{2k_e} & 0_{2k_e} & 0_{2k_e \times 2k_e}
\end{pmatrix}.
\] (2.5.9)
where $s_d^2 = E^2[u^2]$. Now using (2.5.3), (2.5.9) and (2.5.8), we can show that the value of the right hand side of (2.5.4) equals

$$\alpha_{d1}^{-2} \left\{ \frac{1}{4r} + \frac{s_d^{-2}}{r - r^2} \right\}, \quad (2.5.10)$$

and the left hand side of (2.5.4) equals

$$\alpha_{d1}^{-2} \left\{ \frac{4s_d^{-2}u^4 - \left( (1 - r)^2 - 8rs_d^{-2} \right) u^2 + 4s_d^{-2}r^2}{4r^2(1 - r)^2} \right\}. \quad (2.5.11)$$

The expression in (2.5.11) is an even, quartic function of $u$ with positive leading coefficient. This implies that this expression has a critical value at $u = 0$ and can therefore only achieve its maximum over the interval $[-1, 1]$ at the points $u \in \{-1, 0, 1\}$. If we show that (2.5.11) evaluated at $u \in \{-1, 0, 1\}$ is less than or equal to (2.5.10), then (2.5.4) holds for the design $\xi^*$, and the theorem is proved.

When $u = 0$, (2.5.11) equals

$$\alpha_{d1}^{-2} \left( \frac{4s_d^{-2}r^2}{4r^2(1 - r)^2} \right), \quad (2.5.12)$$

and when $u = \pm 1$, (2.5.11) equals

$$\alpha_{d1}^{-2} \left( \frac{(4s_d^{-2} + 1)(1 - r)^2}{4r^2(1 - r)^2} \right). \quad (2.5.13)$$

If we factor a common denominator $4r^2(1 - r)^2$ out of (2.5.12), (2.5.13) and (2.5.10), then equality of three expressions is equivalent to

$$4s_d^{-2}r^2 = (4s_d^{-2} + 1)(1 - r)^2 = (4s_d^{-2} + 1 - r)r(1 - r). \quad (2.5.14)$$

Replacing $r$ in (2.5.14) with its value in (2.5.6) and applying a little arithmetic, we can verify the equalities in (2.5.14). This means (2.5.14) is true for the design $\xi^*$, so the theorem is proved. \(\Box\)
Hypothesis iii.) of Theorem 2.5.1 requires that the design columns for the standardized environmental variables \( v \) be orthogonal to the designs columns corresponding to the intercept, linear and quadratic terms for the standardized control variable \( u \). This is the only requirement that the columns of standardized environmental variables must satisfy in an optimal design. When \( N > k_e + 2 \), we can obtain an optimal exact design of size \( N \) in two stages. First, determine a collection of \( N \) levels of \( u \) that satisfy Hypotheses i.) and ii.) of Theorem 2.5.1. Then, select \( k_e \) vectors from the orthogonal complement of the space generated by \( 1_N, u, \) and \( u^2 \) where \((u^2)_i = u_i^2\). This two stage procedure provides great flexibility in the selection of \( k_e \) columns for the standardized environmental variables. A method for selecting the levels of \( v \) is given in Section 2.8. Applications of this method are provided in Section 2.9 when we construct optimal designs for several examples.

The common value of the second and fourth moments of \( u \) with respect to the optimal design measure \( \xi^* \) is a function of only the size \( s_d \) of the design region \( \mathcal{X}_d \). As \( s_d \) decreases, the size of the region over which the model in (2.3.1) is valid gets smaller, and the second and fourth moments decrease; hence, the optimal design must take more observations in the center of the design region. As \( s_d \) increases, the size of the region \( \mathcal{X} \) increases, and the second and fourth moments increase, so we must take more observations on the edges of the design region. This implies that confidence in the validity of the model (2.3.1) is directly related to the confidence in the initial estimate \( \mu_d^* \) of \( x_d^* \). This point is further emphasized by the following corollary where we place utmost confidence, i.e., \( \text{Var}[u^*] = 0 \), in an a priori value for \( u^* \) (resp. \( x_d^* \)).
Corollary 2.5.2 (Lav Optimal Design) Consider the model in (2.3.1). A design $\xi^* \in \Xi_1$ is a local asymptotic variance optimal design at the point $u^* = 0$ if

i.) $E^{\xi^*} u = E^{\xi^*} u^3 = 0,$

ii.) $E^{\xi^*} u^2 = E^{\xi^*} u^4 = 1,$ and

iii.) for each $j \in \{1, 2, \ldots, k_e\}$, $E^{\xi^*} v_j = E^{\xi^*} v_j u = E^{\xi^*} v_j u^2 = 0.$

An Lav optimal design evenly divides the marginal mass for the standardized control variables $u$ between the levels ±1. As a result, the design matrix is singular, and the denominator of $u^*$, i.e., $\alpha_{d_{u^*}}$, is not estimable. Bounacorssi (1985) showed that estimability of both the numerator and denominator is necessary for $\hat{u}^*$ to be invariant to the choice of generalized inverse for $M(\xi^*)$. Therefore, there is little practical interest in Lav optimal designs.

The statement $u^* = 0$ can be considered as $E^{\xi^*}[u^*] = 0$ for a prior $\pi_d^*$ that is degenerate at $u^* = 0$. This means $\text{Var}^{\xi^*}[u^*] = 0$ as well. Using the degenerate prior $\pi_d^*$, we can prove Corollary 2.5.2 in the same way that we proved Theorem 2.5.1.

Proof of Corollary 2.5.2: For the model in (2.3.1), let $f$ be the regression vector in (2.5.3) where $u \in [-1,1]$ and $v \in [-1,1]^{k_e}$. In addition, let $\pi_d^*$ be a degenerate prior for $u^*$ at the point $u^* = 0$. According to Condition ii.) of Theorem 1.3.4, a design measure $\xi \in \Xi_1$ is Lav optimal at $u^*$ if and only if

$$f^t M(\xi^*)^{-1} E^{\xi^*} \left[ \left( \frac{\partial u^*}{\partial \alpha^i} \right)^t \left( \frac{\partial u^*}{\partial \alpha^i} \right) \right] M(\xi^*)^{-1} f \leq \text{tr} \left\{ M^{-1}(\xi^*) E^{\xi^*} \left[ \left( \frac{\partial u^*}{\partial \alpha^i} \right)^t \left( \frac{\partial u^*}{\partial \alpha^i} \right) \right] \right\}$$

for every $(u, v) \in [-1,1]^{k_e+1}$. 

(2.5.15)
Using the hypotheses of Corollary 2.5.2, the information matrix for the parameter vector $\alpha$ for the design measure $\xi^*$ is

$$
\mathbf{M}(\xi^*) = \begin{pmatrix}
1 & 0 & 1 & 0_{2k_r} \\
0 & 1 & 0 & 0_{2k_r} \\
1 & 0 & 1 & 0_{2k_r} \\
0_{2k_r} & 0_{2k_r} & 0_{2k_r} & \mathbf{V}_u^t \mathbf{V}_u
\end{pmatrix},
$$

(2.5.16)

where $\mathbf{V}_u^t \mathbf{V}_u$ is the matrix defined in (2.5.7). The Moore-Penrose generalized inverse for the information matrix $\mathbf{M}(\xi^*)$ is given by

$$
\mathbf{M}(\xi^*)^{-} = \begin{pmatrix}
\frac{1}{4} & 0 & \frac{1}{4} & 0_{2k_r} \\
0 & 1 & 0 & 0_{2k_r} \\
\frac{1}{4} & 0 & \frac{1}{4} & 0_{2k_r} \\
0_{2k_r} & 0_{2k_r} & 0_{2k_r} & (\mathbf{V}_u^t \mathbf{V}_u)^{-}
\end{pmatrix}
$$

(2.5.17)

where $(\mathbf{V}_u^t \mathbf{V}_u)^{-}$ is the Moore-Penrose generalized inverse for the matrix $\mathbf{V}_u^t \mathbf{V}_u$.

The matrix $\mathbf{E}^{\pi_d^*} \left[ (\partial u^*/\partial \alpha^t)^t (\partial u^*/\partial \alpha^t) \right]$, on the right hand side of (2.5.15), for the degenerate prior $\pi_d^*$ is equal to

$$
\mathbf{E}^{\pi_d^*} \left[ \left( \frac{\partial u^*}{\partial \alpha^t} \right)^t \left( \frac{\partial u^*}{\partial \alpha^t} \right) \right] = \alpha_{d11}^{-2} \begin{pmatrix}
0 & 0 & 0 & 0_{2k_r} \\
0 & \frac{1}{4} & 0 & 0_{2k_r} \\
0 & 0 & 0 & 0_{2k_r} \\
0_{2k_r} & 0_{2k_r} & 0_{2k_r} & 0_{2k_r \times 2k_r}
\end{pmatrix}.
$$

(2.5.18)

Using the matrices in (2.5.18) and (2.5.17), the inequality statement in (2.5.15) is

$$
\frac{u^2}{4\alpha_{d11}^2} \leq \frac{1}{4\alpha_{d11}^2}.
$$

(2.5.19)
Clearly, the left hand side is less than or equal to the right hand side for every $u \in [-1, 1]$ as well as for every $v \in [-1, 1]^*$. Therefore, the corollary is proved. □

### 2.6 MAV Optimal Designs

In this section we determine sufficient conditions for a design measure $\xi \in \Xi_1$ to minimize, over the set $\Xi_1$, the maximum of the asymptotic variance of the estimator $\hat{u}^*$ for values of $u^*$ in a given support space. In Assumptions 2.2.2 we defined the support interval $M_d^*$ for the parameter $x_d^*$ to be of the form

$$M_d^* = [m_d^* - \tau, m_d^* + \tau]$$

where $m_d^*$ is a hypothesized value for $x_d^*$, and $\tau > 0$ is a known scalar. The reparametrization given in Section 2.3, where $u^* = \gamma^{-1}(x_d^* - m_d^*)$, transforms the region $M_d^*$ for $x_d^*$ into the interval

$$[-\gamma^{-1}\tau, \gamma^{-1}\tau].$$

for $u^*$. Hence, the goal of this section is to find a design measure $\xi^* \in \Xi_1$ such that

$$\max_{u^* \in [-\gamma^{-1}\tau, \gamma^{-1}\tau]} \left( \frac{\partial u^*}{\partial \alpha^t} \right) M^-(\xi^*) \left( \frac{\partial u^*}{\partial \alpha^t} \right)^t = \min_{\xi \in \Xi_1} \max_{u^* \in [-\gamma^{-1}\tau, \gamma^{-1}\tau]} \left( \frac{\partial u^*}{\partial \alpha^t} \right) M^-(\xi) \left( \frac{\partial u^*}{\partial \alpha^t} \right)^t.$$

(2.6.3)

**Theorem 2.6.1 (MAV Optimal Design)** Assume the model in (2.3.1), and let $\pi_{1/2}$ be a probability mass function (p.m.f.) for which $\pi_{1/2}(-\gamma/\tau) = \pi_{1/2}(\tau/\gamma) = 1/2$. The integrated asymptotic variance optimal design with respect to the p.m.f. $\pi_{1/2}$, considered as a prior for $u^*$, is a minimax asymptotic variance optimal design with respect to the region $[-\gamma/\tau, \tau/\gamma]$. 


Proof of Theorem 2.6.1: The elements of the gradient vector \((\partial u^*/\alpha^t)\) in (2.4.2) are at most linear functions of the parameter \(u^*\), so the asymptotic variance of \(\hat{u}^*\) in (2.4.1) is a convex function of \(u^*\) for every \(\xi \in \Xi_1\). This implies the maximum of the asymptotic variance of \(\hat{u}^*\) over the interval \([-\tau/\gamma, \tau/\gamma]\) occurs at either \(u^* = -\tau/\gamma\) or \(u^* = \tau/\gamma\). That is, for every \(\xi \in \Xi_1\)

\[
\max_{u^* \in [-\tau/\gamma, \tau/\gamma]} \left( \frac{\partial u^*}{\partial \alpha^t} \right) M^{-}(\xi) \left( \frac{\partial u^*}{\partial \alpha^t} \right)^t = \max_{u^* \in [-\tau/\gamma, \tau/\gamma]} \left( \frac{\partial u^*}{\partial \alpha^t} \right) M^{-}(\xi) \left( \frac{\partial u^*}{\partial \alpha^t} \right)^t. \tag{2.6.4}
\]

To prove the theorem we need only show that an Iav optimal design with respect to the p.m.f. \(\pi_{1/2}\), regarded as a prior for \(u^*\), minimizes the maximum of the asymptotic variance evaluated at the points \(u^* = \pm \tau/\gamma\).

Let \(\xi_{1/2}\) be an Iav optimal design with respect to \(\pi_{1/2}\). Then, \(\xi_{1/2}\) is a Bayes strategy for the “prior” \(\pi_{1/2}\). Moreover, since \(\pi_{1/2}\) gives positive probability to each point in the finite set \([-\tau/\gamma, \tau/\gamma]\), \(\xi_{1/2}\) is an admissible strategy (Berger, 1985 — page 253).

Under \(\pi_{1/2}\), \(E^{\pi_{1/2}}[u^*] = 0\), and \(E^{\pi_{1/2}}[u^{*2}] = \tau^2/\gamma^2\), so the information matrix \(M(\xi_{1/2})\) for the vector \(\alpha\), and the Moore-Penrose generalized inverse \(M(\xi_{1/2})^{-}\) are respectively given by (2.5.5) and (2.5.8) where

\[
r = 1 + \frac{4\tau^2}{\gamma^2} + \sqrt{\frac{4\tau^2}{\gamma^2} \left( 1 + \frac{4\tau^2}{\gamma^2} \right)}.
\tag{2.6.5}
\]

The value of the asymptotic variance of \(\hat{u}^*\) on the set \(u^* \in \{-\tau/\gamma, \tau/\gamma\}\) is constant, i.e.,

\[
\left( \frac{\partial u^*}{\partial \alpha^t} \right) M^{-}(\xi) \left( \frac{\partial u^*}{\partial \alpha^t} \right)^t \bigg|_{u^* = \pm \tau/\gamma} = \frac{1}{4r} + \frac{(\tau/\gamma)^2}{r(1-r)}. \tag{2.6.6}
\]
This means that $\xi_{\frac{1}{2}}$ is an equalizer strategy for the set $\{-\tau/\gamma, \tau/\gamma\}$. Since every admissible equalizer strategy is minimax, the design $\xi_{\frac{1}{2}}$ is minimax and therefore minimizes the right hand side of (2.6.4). □

2.7 Local Approximate Stochastic Domination Optimal Designs

In this section we approximate the probability (1.2.19) for Design Criterion 1.2.4 to give a local approximate stochastic domination optimal design. The approximation occurs because the determination of an optimal design with respect to (1.2.19) is not an analytically tractable problem. It will be shown that the approximate version of the Lsd criterion is equivalent to the Lav criterion. Hence, sufficient conditions for a design to be a local approximate stochastic domination optimal design are given by the hypotheses in Corollary 2.5.2.

Following the reparameterization of Section 2.4, the probability (1.2.19) becomes

$$P_\xi \{|\hat{u}^* - u^*| > \epsilon/\gamma\}$$

(2.7.1)

where $|\cdot|$ is the usual Euclidean distance. The design problem is determining a design $\xi \in \Xi_1$ that minimizes (2.7.1) at $u^* = 0$ over $\Xi_1$ for all $\epsilon > 0$. Essentially, the determination of a Lsd optimal design requires finding the design $\xi \in \Xi_1$ for which the tails of the distribution of $\hat{u}^*$ are the "lightest".

The distribution of the ratio estimator $\hat{u}^*$ was first investigated by Fieller (1932). Marsaglia (1965) and Hinckley (1969) later obtained expressions for $\hat{u}^*$'s density
and cumulative distribution functions. For a given design $\xi \in \Xi_1$, the cumulative distribution function for $\hat{u}^*$ is

$$F_{\hat{u}^*}(t) = P_\xi \{ \hat{\alpha}_{d_i} + 2\hat{\alpha}_{d_{11}} t \leq 0, \hat{\alpha}_{d_{11}} \leq 0 \} + P_\xi \{ \hat{\alpha}_{d_i} + 2\hat{\alpha}_{d_{11}} t \geq 0, \hat{\alpha}_{d_{11}} \geq 0 \}. \quad (2.7.2)$$

The probabilities on the right hand side of (2.7.2) depend on the joint distribution of the random variables $\hat{\alpha}_{d_{11}}$ and $\hat{\alpha}_{d_i} + 2\hat{\alpha}_{d_{11}} t$. This makes the derivation of an optimal design using (2.7.1) directly difficult.

Under appropriate conditions on a sequence of design measures $\{\xi_N\}_{N=1}^\infty$ with increasing sample sizes,

$$\frac{\hat{u}^* - u^*}{\sqrt{\text{Var}^\xi_N [\hat{\alpha}_{d_{11}}]}} \xrightarrow{N \to \infty} N(0, 1) \quad (2.7.3)$$

provided $\text{Var}^\xi_N [\hat{\alpha}_{d_{11}}] \xrightarrow{N \to \infty} 0$ (Buonaccorsi 1985). This means that for large sample sizes the probability in (2.7.1) is approximately minimized for every value of $\epsilon > 0$ if we find a design $\xi \in \Xi_1$ that minimizes $\text{Var}^\xi [\hat{\alpha}_{d_i}]$ over $\Xi_1$. The design that minimizes $\text{Var}^\xi [\hat{\alpha}_{d_i}]$ over $\Xi_1$ is called a Local approximate stochastic domination design for estimating $u^*$ at the point $u^* = 0$.

The quantity $\text{Var}^\xi [\hat{\alpha}_{d_i}]$ is proportional to the right hand side of (1.2.16) evaluated at the point $u^* = 0$. Thus, this approximate version of the Lsd design criterion for the value $u^* = 0$ is identical to the Lav design criteria given $u^* = 0$. This means that a design that satisfies the hypotheses in Corollary 2.5.2 is also a local approximate stochastic domination design for estimating $u^*$ at the point $u^* = 0$. 
The Lsd criterion is difficult to solve for estimating \( u^* \). However, we were able to obtain some new results for the stochastic domination design criterion when estimating linear functions of the regression parameters. Since these results are not pertinent to the problem at hand, we present them in Chapter V.

2.8 A Method for Constructing Optimal Designs

In this section we present a method for constructing exact designs that approximate the optimal designs in Theorem 2.5.1, Corollary 2.5.2, and Theorem 2.6.1. For each design criterion, the optimal design \( \xi^* \) places equal marginal mass on the points \( u = \pm 1 \) with the remaining mass at \( u = 0 \). For example, the lav optimal design has

\[
\xi^*_u(-1) = \xi^*_u(1) = 0.5 + 2s_d^{-2} - \sqrt{s_d^{-2}(1 + 4s_d^{-2})}
\]

and

\[
\xi^*_u(0) = 2\sqrt{s_d^{-2}(1 + 4s_d^{-2})} - 4s_d^{-2}
\]

where \( \xi_u^* \) is the marginal of the optimal measure \( \xi^* \) for the standardized control variable \( u \). For a given sample size \( N \), we allocate, as best possible, \( N\xi_u^*(u) \) observations to the standardized control variable level \( u \) for \( u \in \{-1, 0, 1\} \). By best we mean the allocation with an equal number of observations on the values \( u = \pm 1 \) and for which (2.5.1) is minimized over all such allocations with respect to the appropriate prior.

For each of the design criteria, the only constraint used for the selection of the levels of a given standardized environmental variable is that the resulting vector be
orthogonal to the three vectors associated with the coefficients $\alpha_0, \alpha_{d_1}$, and $\alpha_{d_1}$. The space spanned by these three vectors is at most three-dimensional; hence, if $k_r \leq N - 3$, we can select the environmental variables so that the $k_r$ columns for the standardized environmental variables are mutually orthogonal. Any set of $k_r$ orthogonal columns in the $N - 3$ dimensional space could be chosen. We choose the orthogonal columns so that the distribution of design points across each margin $[-1, 1]$ closely resembles the known marginal distributions.

Fix an integer $q > 0$, and partition the margin $[-1, 1]$ for each standardized environmental variable into $q$ equiprobable regions. Under the assumption of independence of the variables in $\mathbf{v}$, this also partitions the space $[-1, 1]^{k_r}$ into $q^{k_r}$ equiprobable regions. The goal here is to choose $N$ levels for each standardized environmental variable, so that each of the $q$ regions is equally represented. In addition, we want the $N$ vectors of standardized environmental variables to be equally divided among the $q^{k_r}$ regions of $[-1, 1]^{k_r}$. To achieve these two objectives we consider maximizing the entropy functions

$$E_m(v_j) = -\sum_{i=1}^{q} p_i^{(m)} \log p_i^{(m)} \text{ for } j = 1, 2, \ldots, k_r,$$

(2.8.3)

where $p_i^{(m)}$, $i \in \{1, 2, \ldots, q\}$, is the proportion of the $N$ levels of $v_j$ in the $i^{th}$ stratum of $[-1, 1]$ and

$$E_j(v_1, \ldots, v_{k_r}) = -\sum_{i=1}^{q^{k_r}} p_i^{(j)} \log p_i^{(j)}$$

(2.8.4)

where $p_i^{(j)}$ is the proportion of the $N$ observations in the $i^{th}$ stratum of $[-1, 1]^{k_r}$. In (2.8.3) and (2.8.4), $p_i \log p_i = 0$ when $p_i = 0$. 

The functions (2.8.3) and (2.8.4) are maximized when the respective proportions are all equal. The maximum of each marginal entropy is
\[ E_{m}^{\text{max}} = -q \cdot \left[ \frac{N}{q} \right] \log \left( \frac{\left[ \frac{N}{q} \right]}{N} \right) - (N - q \cdot \left[ \frac{N}{q} \right]) \cdot \left[ \frac{N}{q} \right] + 1 \cdot \log \left( \frac{\left[ \frac{N}{q} \right] + 1}{N} \right) \] (2.8.5)
where \([\cdot]\) is the greatest integer function, and the maximum of the joint entropy is
\[ E_{j}^{\text{max}} = \log N. \] (2.8.6)

Thus, we want to determine a collection of \(k_e\) orthogonal columns, in the complement of the space generated by \(1_N, u,\) and \(u^2\), for which the \(k_e + 1\) quantities \(E_m(v_j)/E_{m}^{\text{max}},\) for \(j = 1, 2, \ldots, k_e\), and \(E_j(v_1, \ldots, v_{k_e})/E_{j}^{\text{max}}\) are all near 1.

To determine a collection of \(k_e\) orthogonal columns for the standardized environmental variables, we use the following algorithm.

**Step 1.** Set \(j = 1\).

**Step 2.** If \(j \leq k_e\), randomly generate \(w \in \mathbb{R}^N\); else stop.

**Step 3.** Project \(w\) onto the orthogonal complement of the space spanned by \(1_N, u, u^2\) and \(\{v_k\}_{k<j}\).

**Step 4.** Calculate \(E_m(w)/E_{m}^{\text{max}}\) and \(E_j(\{v_k\}_{k<j}, w)/E_{j}^{\text{max}}\).

**Step 5.** If both values are close enough to 1, then set \(v_j = w\) and \(j = j + 1\).

**Step 6.** Go to Step 2.

For the Iav and Mav criteria, an exact design \(\xi_N\) that is constructed using the algorithm given above has an information matrix \(M(\xi_N)\) of the form given by (2.5.5).
In most cases the optimal value of $r$ will not be achieved by the exact design. Therefore, to determine the best exact design, we use (2.5.10) to calculate the efficiency of an exact design, i.e.,

$$
\frac{\text{tr} \left\{ \mathbf{M}^{-1} \mathbb{E}_{\pi^r} \left[ \left( \frac{\partial \mathbf{u}^r}{\partial \mathbf{x}} \right) \left( \frac{\partial \mathbf{u}^r}{\partial \mathbf{x}} \right)^T \right] \right\}}{\text{tr} \left\{ \mathbf{M}^{-1} \mathbb{E}_{\pi^N} \left[ \left( \frac{\partial \mathbf{u}^N}{\partial \mathbf{x}} \right) \left( \frac{\partial \mathbf{u}^N}{\partial \mathbf{x}} \right)^T \right] \right\}}
$$

(2.8.7)

using the appropriate prior. The efficiency in (2.8.7) is less than or equal to 1 for every exact design, so we choose the exact design produced by the algorithm given above for which (2.8.7) is nearest to 1. This optimal exact design minimizes (2.5.1) over the set of all exact designs produced by the algorithm. Note that such issues do not arise for the Lav criterion since an exact design $\xi^r$ that is constructed per our method for the Lav criterion satisfies the hypotheses in Corollary 2.5.2.

### 2.9 Examples of Optimal Designs

Consider a quality improvement computer experiment where the computer code depends on a single controllable variable $x_d$ and two random environmental variables $x_e$. We assume the model in (2.1.1), and we suppose $x_e$ is distributed according to $\pi_e$ for which

$$
\mathbf{\mu}_e = \mathbb{E}_{\pi_e}[x_e] = \begin{pmatrix} 3 \\ -1 \end{pmatrix},
$$

(2.9.1)

and

$$
\mathbf{\Sigma}_e = \mathbb{E}_{\pi_e}[(x_e - \mathbf{\mu}_e)(x_e - \mathbf{\mu}_e)^T] = \begin{pmatrix} 2.5 & 1.5 \\ 1.5 & 2.5 \end{pmatrix}.
$$

(2.9.2)
We let \( s_c = 2 \) be the sizing scalar for the design region for environmental variables; hence, the space \( \mathcal{X}_c \) is given by

\[
\mathcal{X}_c = \left\{ 2 \begin{pmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{pmatrix} v + \begin{pmatrix} 3 \\ -1 \end{pmatrix} \quad \text{for all } v \in [-1,1]^2 \right\} \tag{2.9.3}
\]

see Figure 1. In addition, we define the design space \( \mathcal{X}_d \) for the controllable variable to be the closed interval \([0,10]\). For this experiment we shall construct optimal designs of size \( N = 16 \) for estimating the optimal operating condition \( x_d^* \) for the Iav, Lav, and Mav design criteria.

Defining the support space \( \mathcal{X}_d \) before adding prior information about \( x_d^* \) into the design problem is not consistent with our approach described in Section 2.2.1. However, we state its definition here to emphasize that we are designing experiments on the same design space using different criteria. In each of the examples, we assert the requisite prior assumptions for \( x_d^* \) and then “define” the design space to be \( \mathcal{X}_d = [0,10] \).

For each criterion, an optimal design on \([-1,1]^{k_r+1}\) is determined and constructed using code that is available from the author. In the first example we provide the program session that determined and constructed the optimal design. Since the program sessions for the final two examples are nearly identical to the first, we present only the results for them.

**Example 2.9.1 (Iav Optimal Design)** Suppose that we have specified the values of the first two moments of \( x_d^* \) with respect to \( \pi_d^* \) by

\[
\mu_d^* = \mathbb{E}^d[x_d^*] = 5, \tag{2.9.4}
\]
and

\[ \sigma_d^2 = E^*[(x_d^* - 5)^2] = \frac{1}{2}. \]  \hspace{1cm} (2.9.5)

In addition, assume we specified the sizing scalar \( s_d = 5\sqrt{2} \) for the design region \( X_d \).

Then, \( X_d = [-\gamma + \mu_d^*, \gamma + \mu_d^*] = [0, 10] \) where \( \gamma = s_d\sigma_d \).

We exercise the reparameterization in Section 2.3 so that we can determine an optimal design on \([-1, 1]^3\). By Theorem 2.5.1, an integrated asymptotic variance optimal design for estimating \( u^* \) (resp. \( x_d^* \)) must split \( 1 + 4\sigma_d^{-2} - \sqrt{4\sigma_d^{-2}(1 + 4\sigma_d^{-2})} = 12.58 \) observations between the control levels \( u = \pm 1 \) (resp. \( x_d \in \{0, 10\} \)) with the remaining 3.42 observations taken at \( u = 0 \) (resp. at \( x_d = 5 \)). In addition, the design columns for the environmental variables \( v \) must be orthogonal to the space generated by the design columns \( u, u^2 \) and the vector \( 1_N \).

A sixteen point optimal design does not exist, since the weights on the design points are not integers, but the following session of the author's computer program determines an efficient 16 point experimental design. The design is listed in Table 1.

Enter the number of controllable variates -- Kd
1
Enter the number of environmental variates -- Ke
2
Enter the scaling factor of Xd -- Sd
7.0710678
Enter the number of design points
16

The criterion efficiency is .99311

Divide the noise margins into how many regions
8
The multivariate entropy efficiency is 1.00000
The univariate entropy efficiencies are .984 .984
The correlations between columns are .000

Are these efficiencies okay? (y or n) n

The multivariate entropy efficiency is 1.00000
The univariate entropy efficiencies are 1.000 1.000
The product correlations between columns are .000

Are these efficiencies okay? (y or n) y

Figure 1: An efficient sixteen point design for the Iav criterion for estimating the optimal operating condition \( x_d^* \) for the control variable \( x_d \) in the presence of two environmental variables. The rhomboidal region in each plot designates the space \( X_e \).
Table 1: An efficient sixteen point design for the Iav criterion for estimating $u^*$ (resp. $x_d^*$) given the model in (2.3.1) with $k_d = 1$ and $k_r = 2$. The asymptotic variance of $\hat{u}^*$ is integrated with respect to a prior $\pi_d^*$ for which the hypotheses in Assumptions 2.2.1 are satisfied with $\mu_d^* = 5$ and $\sigma_d^2 = 0.5$. The efficiency of this design with respect to the optimal continuous design is 0.993.

<table>
<thead>
<tr>
<th>$u$</th>
<th>$v_1$</th>
<th>$v_2$</th>
<th>$x_d$</th>
<th>$x_{e_1}$</th>
<th>$x_{e_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>-0.753</td>
<td>-0.998</td>
<td>10.0</td>
<td>-0.257</td>
<td>-4.746</td>
</tr>
<tr>
<td>1.0</td>
<td>0.427</td>
<td>-0.886</td>
<td>10.0</td>
<td>3.396</td>
<td>-3.230</td>
</tr>
<tr>
<td>1.0</td>
<td>0.098</td>
<td>0.927</td>
<td>10.0</td>
<td>4.219</td>
<td>1.877</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.602</td>
<td>0.472</td>
<td>10.0</td>
<td>1.667</td>
<td>-0.185</td>
</tr>
<tr>
<td>1.0</td>
<td>0.998</td>
<td>-0.215</td>
<td>10.0</td>
<td>5.778</td>
<td>-0.648</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.168</td>
<td>0.700</td>
<td>10.0</td>
<td>3.197</td>
<td>0.931</td>
</tr>
<tr>
<td>0.0</td>
<td>-0.936</td>
<td>-0.282</td>
<td>5.0</td>
<td>-0.091</td>
<td>-2.781</td>
</tr>
<tr>
<td>0.0</td>
<td>-0.115</td>
<td>-0.274</td>
<td>5.0</td>
<td>2.383</td>
<td>-1.935</td>
</tr>
<tr>
<td>0.0</td>
<td>0.274</td>
<td>0.417</td>
<td>5.0</td>
<td>4.237</td>
<td>0.523</td>
</tr>
<tr>
<td>0.0</td>
<td>0.777</td>
<td>0.139</td>
<td>5.0</td>
<td>5.471</td>
<td>0.194</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.226</td>
<td>-0.644</td>
<td>0.0</td>
<td>3.033</td>
<td>-2.706</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.504</td>
<td>0.766</td>
<td>0.0</td>
<td>5.279</td>
<td>1.802</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.653</td>
<td>-0.621</td>
<td>0.0</td>
<td>4.338</td>
<td>-2.210</td>
</tr>
<tr>
<td>-1.0</td>
<td>-0.593</td>
<td>-0.150</td>
<td>0.0</td>
<td>1.071</td>
<td>-2.043</td>
</tr>
<tr>
<td>-1.0</td>
<td>-0.332</td>
<td>0.072</td>
<td>0.0</td>
<td>2.076</td>
<td>-1.116</td>
</tr>
<tr>
<td>-1.0</td>
<td>-0.458</td>
<td>0.577</td>
<td>0.0</td>
<td>2.203</td>
<td>0.274</td>
</tr>
</tbody>
</table>

The first columns for the standardized environmental variables that were given by the program have univariate entropy efficiencies $\mathcal{E}_m(v_1) = \mathcal{E}_m(v_2) = 0.984$. Although this was the best pair of columns obtained after generating 5000 candidates for each column, we felt that better choices might exist. Therefore, we rejected these and requested the program to generate 5000 additional candidates for each column. The
total computing time for this complete session was approximately 30 seconds on an
HP workstation.

The first three columns of Table 1 give the design points for an efficient design on
$[-1, 1]^3$ for the Iav criterion. The last three columns of Table 1 give the corresponding
design on $\mathcal{X} = \mathcal{X}_d \times \mathcal{X}_e$. From the table we see that an efficient 16 point design for the
Iav criterion with respect to the prior $\pi^*_d$ takes 6 observations at each of the endpoints
of the control variable design space with the remaining 4 observations taken at the
midpoint. In Figure 1, we draw the scatterplots of $x_{c1}$ vs. $x_{c2}$ for each level of $x_d$ in
the design. These plots show how the environmental variable levels are spread over
the design space $\mathcal{X}_e$. □

Example 2.9.2 (Lav Optimal Design) Suppose that we have specified a degener­
ate prior $\pi^*_d$ for $x^*_d$ at the point

$$\mu^*_d = E^{\pi^*_d}[x^*_d] = 5. \quad (2.9.6)$$

This implies that

$$\sigma^2_d = E^{\pi^*_d}[(x^*_d - 5)^2] = 0. \quad (2.9.7)$$

To define the design space $\mathcal{X}_d$, we treat $\sigma_d$ as an arbitrarily small positive number
and assume $s_d$ is arbitrarily large so that $s_d\sigma_d = 5$. Then, the design region $\mathcal{X}_d =
[-\gamma + \mu^*_d, \gamma + \mu^*_d] = [0, 10]$ where $\gamma = s_d\sigma_d$. In the design program, we set $s_d = 1000$.

We apply the reparameterization of Section 2.3, so that we can determine the
optimal design on $[-1, 1]^3$. By Corollary 2.5.2, we see that a local asymptotic variance
optimal design for estimating $u^*$ (resp. $x^*_d$) puts half, i.e., 8, of its observations on
each of the control variable levels ±1 (resp. \( x_d \in \{0, 10\} \)). The space spanned by the columns for the model terms 1, \( x_d \), and \( x_d^2 \) is only of dimension two, so the columns for the standardized environmental variables are selected from a space of dimension fourteen.

Table 2: A sixteen point Lav optimal design for estimating \( u^* \) (resp. \( x_d^* \)) given the model in (2.3.1) with \( k_d = 1 \) and \( k_e = 2 \). The asymptotic variance of \( \hat{x}_d^* \) is integrated with respect to a degenerate prior \( \pi_d^* \) for which the hypotheses in Assumptions 2.2.1 are satisfied with \( \mu_d^* = 5 \) and \( \sigma_d^2 = 0 \). This design has efficiency 1 with respect to the optimal continuous design.

<table>
<thead>
<tr>
<th>( u )</th>
<th>( v_1 )</th>
<th>( v_2 )</th>
<th>( x_d )</th>
<th>( x_{e1} )</th>
<th>( x_{e2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.313</td>
<td>0.839</td>
<td>10.0</td>
<td>4.778</td>
<td>1.830</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.874</td>
<td>0.023</td>
<td>10.0</td>
<td>0.401</td>
<td>-1.806</td>
</tr>
<tr>
<td>1.0</td>
<td>0.996</td>
<td>-0.950</td>
<td>10.0</td>
<td>5.038</td>
<td>-2.854</td>
</tr>
<tr>
<td>1.0</td>
<td>0.090</td>
<td>-0.997</td>
<td>10.0</td>
<td>2.274</td>
<td>-3.900</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.429</td>
<td>0.706</td>
<td>10.0</td>
<td>2.419</td>
<td>0.689</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.221</td>
<td>0.138</td>
<td>10.0</td>
<td>2.474</td>
<td>-0.807</td>
</tr>
<tr>
<td>1.0</td>
<td>0.170</td>
<td>0.390</td>
<td>10.0</td>
<td>3.899</td>
<td>0.340</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.045</td>
<td>-0.149</td>
<td>10.0</td>
<td>2.717</td>
<td>-1.492</td>
</tr>
<tr>
<td>-1.0</td>
<td>-0.415</td>
<td>-0.640</td>
<td>0.0</td>
<td>1.116</td>
<td>-3.333</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.349</td>
<td>-0.381</td>
<td>0.0</td>
<td>3.668</td>
<td>-1.792</td>
</tr>
<tr>
<td>-1.0</td>
<td>-0.571</td>
<td>0.970</td>
<td>0.0</td>
<td>2.258</td>
<td>1.338</td>
</tr>
<tr>
<td>-1.0</td>
<td>-0.798</td>
<td>-0.568</td>
<td>0.0</td>
<td>0.037</td>
<td>-3.503</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.848</td>
<td>0.469</td>
<td>0.0</td>
<td>6.013</td>
<td>1.256</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.596</td>
<td>-0.010</td>
<td>0.0</td>
<td>4.778</td>
<td>-0.435</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.589</td>
<td>0.611</td>
<td>0.0</td>
<td>5.377</td>
<td>1.421</td>
</tr>
<tr>
<td>-1.0</td>
<td>-0.598</td>
<td>-0.451</td>
<td>0.0</td>
<td>0.754</td>
<td>-2.951</td>
</tr>
</tbody>
</table>

The first three columns of Table 2 give the local asymptotic variance optimal design, at \( u^* = 0 \), on the space \([-1, 1]^3\), while last three columns of Table 2 give the local asymptotic variance optimal design, at \( x_d^* = 5 \), on \( \mathcal{X} = \mathcal{X}_d \times \mathcal{X}_e \). Figure 2
Figure 2: A sixteen point Lav optimal design for estimating the optimal operating condition \( x_d^* \) for the control variable \( x_d \) in the presence of two environmental variables. The rhomboidal region in both plots designates the space \( \mathcal{X}_e \).

contains the two scatterplots of \( x_{e_1} \) vs. \( x_{e_2} \) for the observed levels of \( x_d \) in the design. These plots clearly show that the noise variable levels used in the design are spread evenly over the region \( \mathcal{X}_e \): the multivariate entropy efficiency for the two columns of environmental variables is 1.000, and the both univariate entropy efficiencies are equal to 1.000. In addition, since the exact design matches the optimal design of Corollary 2.5.2, the criterion efficiency is 1.000. □

**Example 2.9.3 (Mav Optimal Design)** Suppose we define the support space \( \mathcal{M}^*_d = [1, 9] \) for the unknown parameter \( x_d^* \). The midpoint of \( \mathcal{M}^*_d \) is \( m_d^* = 5 \), so the design space \( \mathcal{X}_d \) must also be centered at 5. We let \( \mathcal{X}_d = [0, 10] \).
The midranges for $M_d^*$ and $X_d$ are 4 and 5, respectively. This means that $u^*$, when distributed according to the probability mass function $\pi_{\frac{1}{2}}$ in Theorem 2.6.1, has

$$E[\pi_{\frac{1}{2}}[u^*]] = 0,$$

(2.9.8)

and

$$E[\pi_{\frac{1}{2}}[u^{*2}]] = \frac{\tau^2}{\gamma^2} = \left(\frac{4}{5}\right)^2.$$  

(2.9.9)

Since the Iav optimal design with respect to the $\pi_{\frac{1}{2}}$ is also the Mav optimal design and $s_{d}^2 = E[\pi_{\frac{1}{2}}[u^{*2}]]$, $s_d = 1.25$.

![Figure 3: An efficient sixteen point design for the Mav criterion for estimating the optimal operating condition $x_d^*$ for the control variable $x_d$ in the presence of two environmental variables $x_e$. The rhomboidal region in each plot designates the space $X_e$.](image)

We apply the reparameterization in Section 2.3 so that we can determine the optimal design on $[-1, 1]^3$. By Theorem 2.6.1, the 16 point minimax asymptotic variance
Table 3: An efficient sixteen point design for the Mav criterion for estimating $u^*$ (resp. $x^*_d$) given the model in (2.3.1) with $k_d = 1$ and $k_e = 2$. This design minimizes the maximum of the asymptotic variance of $\hat{x}^*_d$ over the support space $M^*_d$ that satisfies the hypotheses in Assumptions 2.2.2. The efficiency of this design with respect to the optimal continuous design is 0.993.

<table>
<thead>
<tr>
<th>$u$</th>
<th>$v_1$</th>
<th>$v_2$</th>
<th>$x_d$</th>
<th>$x_{c_1}$</th>
<th>$x_{c_2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>-0.011</td>
<td>-0.098</td>
<td>10.0</td>
<td>2.870</td>
<td>-1.304</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.349</td>
<td>0.691</td>
<td>10.0</td>
<td>2.643</td>
<td>0.724</td>
</tr>
<tr>
<td>1.0</td>
<td>0.113</td>
<td>-0.918</td>
<td>10.0</td>
<td>2.421</td>
<td>-3.641</td>
</tr>
<tr>
<td>1.0</td>
<td>0.247</td>
<td>0.325</td>
<td>10.0</td>
<td>4.067</td>
<td>0.221</td>
</tr>
<tr>
<td>-1.0</td>
<td>-0.709</td>
<td>-0.634</td>
<td>0.0</td>
<td>0.238</td>
<td>-3.612</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.467</td>
<td>0.141</td>
<td>0.0</td>
<td>4.542</td>
<td>-0.111</td>
</tr>
<tr>
<td>-1.0</td>
<td>-0.746</td>
<td>0.496</td>
<td>0.0</td>
<td>1.259</td>
<td>-0.259</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.988</td>
<td>-0.002</td>
<td>0.0</td>
<td>5.961</td>
<td>-0.018</td>
</tr>
<tr>
<td>0.0</td>
<td>0.671</td>
<td>-0.355</td>
<td>5.0</td>
<td>4.658</td>
<td>-1.394</td>
</tr>
<tr>
<td>0.0</td>
<td>0.314</td>
<td>-0.996</td>
<td>5.0</td>
<td>2.946</td>
<td>-3.673</td>
</tr>
<tr>
<td>0.0</td>
<td>0.765</td>
<td>0.825</td>
<td>5.0</td>
<td>6.120</td>
<td>2.241</td>
</tr>
<tr>
<td>0.0</td>
<td>-0.996</td>
<td>0.818</td>
<td>5.0</td>
<td>0.830</td>
<td>0.459</td>
</tr>
<tr>
<td>0.0</td>
<td>-0.795</td>
<td>-0.464</td>
<td>5.0</td>
<td>0.152</td>
<td>-3.187</td>
</tr>
<tr>
<td>0.0</td>
<td>-0.146</td>
<td>-0.717</td>
<td>5.0</td>
<td>1.844</td>
<td>-3.298</td>
</tr>
<tr>
<td>0.0</td>
<td>0.608</td>
<td>0.733</td>
<td>5.0</td>
<td>5.558</td>
<td>1.807</td>
</tr>
<tr>
<td>0.0</td>
<td>-0.421</td>
<td>0.155</td>
<td>5.0</td>
<td>1.892</td>
<td>-0.956</td>
</tr>
</tbody>
</table>

The near-optimal design for estimating $u^*$ (resp. $x^*_d$) splits $1 + 4s_d^{-2} - \sqrt{4s_d^{-2}(1 + 4s_d^{-2})} = 8.66$ observations between the control levels $u = \pm 1$ (resp. $x_d \in \{0, 10\}$). The remaining 7.34 observations are taken at $u = 0$ (resp. $x_d = 5$).

In the first three columns of Table 3 we give a 99.3% efficient 16 point minimax asymptotic variance design on $[-1, 1]^3$. The last three columns of Table 3 give the corresponding design on $\mathcal{X} = \mathcal{X}_d \times \mathcal{X}_e$. Figure 3 contains the scatterplots of $x_{c_1}$ vs.
$x_{r_2}$ for each level of $x_d$ in the design. These plots show the uniformity in distribution of the noise variable levels used in the design. The multivariate entropy efficiency of the exact design is 1.000, and both univariate entropy efficiencies are equal to 1.00.

□
CHAPTER III

Determining the Optimal Multivariate Operating Conditions

3.1 Introduction

In this chapter we consider the multivariate version of the problem studied in Chapter II. That is, we seek to design computer experiments so as to accurately determine the levels $x_d^*$ of the controllable factors $x_d$ that minimize the expected integrated loss function in (1.2.4). For a given sample size $N$, the goal is to determine an optimal collection of $N$ input configurations at which the computer code should be run. In this chapter optimal means the set of inputs that allow the most accurate estimation of the optimal levels $x_d^*$ with accuracy defined by the multivariate criteria in Section 1.2.4.

We assume realizations of the output $y$ are generated by a computer code that requires input values for $k_d$ controllable factors, $x_d$, and $k_e$ environmental factors, $x_e$. In addition, we assume that the deterministic output $y(x_d, x_e)$ can be consider as a realization of the stochastic process $Y(x_d, x_e)$ given in (1.2.3). For this model the
expected integrated loss function \( L(\cdot) \) in (1.2.4) is minimized, uniquely, at the levels

\[
x_d^* = -\frac{1}{2} B_d^{-1}(\beta_d + B_I \mu_e).
\]  

(3.1.1)

of the controllable variables \( x_d \).

Designing experiments to accurately estimate the levels \( x_d^* \) in (3.1.1) is similar to the problem of designing experiments to estimate the optimal point of a quadratic response surface. However, designs in the latter case do not include the environmental factors \( x_e \). Optimal designs for estimating the optimal point of a quadratic response surface have been considered by Chatterjee and Mandal (1981) and Mandal (1982).

Using the rewrite of the model (1.2.3) in (1.2.6) and the definitions of the coefficient arrays in (1.2.7) through (1.2.10), we see that the unknown estimand \( x_d^* \) is a nonlinear function of the regression parameters in (1.2.6). Buonaccorsi (1985) showed no unbiased estimator for \( x_d^* \) exists when \( k_d = 1 \). Hence, an unbiased estimator cannot exist for the multivariate estimand \( x_d^* \). We estimate \( x_d^* \) using the maximum likelihood estimator

\[
\hat{x}_d^* = -\frac{1}{2} \hat{B}_d^{-1}(\hat{\beta}_d + \hat{B}_I \hat{\mu}_e).
\]

(3.1.2)

where \( \hat{\beta}_d \), \( \hat{B}_d \) and \( \hat{B}_I \) are the arrays of least-squares estimators of the elements of the arrays \( \beta_d \), \( B_d \) and \( B_I \), respectively.

In this chapter we determine optimal designs for the multivariate criteria defined in Section 1.2.4. Both criteria are real-valued functions of the asymptotic variance-covariance matrix for \( \hat{x}_d^* \). In Section 3.4 we determine this asymptotic dispersion matrix. In Section 3.5 we determine an integrated asymptotic variance A-optimal (IavA-optimal) design. Section 3.6 contains the derivation of an integrated asymptotic
variance D-optimal (lavgD-optimal) design. A method of constructing the optimal designs for both criteria is given in Section 3.7, and several examples of optimal designs are provided in Section 3.8. We begin this chapter by defining the design region \( \mathcal{X} \) (Section 3.2) and a subsequent reparameterization of the model (Section 3.3).

### 3.2 Defining the Design Region \( \mathcal{X} \)

We must specify a common compact support space \( \mathcal{X} \) for the class of all competing designs. Similar to our development in Chapter II, we define \( \mathcal{X} \) as the Cartesian product of the support spaces \( \mathcal{X}_d \) and \( \mathcal{X}_e \) for the controllable and environmental variables, respectively. We allow generality in the definition of \( \mathcal{X} = \mathcal{X}_d \times \mathcal{X}_e \) by requiring only that \( \mathcal{X}_d \) and \( \mathcal{X}_e \) are the images of nonsingular affine transformations of the symmetric cubes \([-1, 1]^d \) and \([-1, 1]^e \). This implies \( \mathcal{X} \) is the image of a nonsingular affine transformation of the symmetric cube \([-1, 1]^{d+e} \). Therefore, \( \mathcal{X} \) is a closed, bounded and connected region in \( \mathbb{R}^{k_d+k_r} \).

#### 3.2.1 Defining \( \mathcal{X}_d \)

When defining the design space \( \mathcal{X}_d \), we assume the optimal operating condition \( x_d^* \) of \( x_d \) is in the interior of \( \mathcal{X}_d \). Since \( x_d^* \) is a nonlinear function of the regression coefficients in (1.2.6), we also assume a priori knowledge about the value of \( x_d^* \). For the two multivariate design criteria that we consider (see Section 1.2.4), there are differing amounts of prior knowledge about \( x_d^* \) that we must add to the design problem in order
to determine optimal designs. Since a priori knowledge impacts the way we define $X_d$, we must consider construction of the space $X_d$ separately for the two design criteria.

For both design criteria we add prior knowledge to the design problem by means of a prior distribution for the unknown regression coefficients arrays in (1.2.3). We assume the regression coefficients $\beta_0, \beta_d, B_d, \beta_e$ and $B_I$ in (1.2.3) are random variables. Let $\pi^*_d$ be the conditional distribution of $x^*_d$ given $\beta_0, B_d, \beta_e$ and $B_I$, and define $\pi^*_B$ to be the marginal prior for the parameters $\beta_0, B_d, \beta_e$, and $B_I$.

For the lavg design criterion (Design Criterion 1.2.5) we add the following distributional assumptions to the design problem.

**Assumptions 3.2.1** The parameters $\beta_0, x^*_d, B_d, \beta_e$ and $B_I$ are random variables governed by the probability law $\pi^* = \pi^*_d \pi^*_B$ for which

i.) $\mu^*_d = \mathbb{E}^{x_d}[x^*_d|\beta_0, B_d, \beta_e, B_I] \in \mathbb{R}^{k_d}$ is known,

ii.) the $k_d \times k_d$ real matrix $\Sigma_d = \mathbb{E}^{x_d}[(x^*_d - \mu^*_d)(x^*_d - \mu^*_d)^t|\beta_0, B_d, \beta_e, B_I]$ is positive definite and known,

iii.) the distribution of $x^*_d - \mu^*_d$ under $\pi^*_d$ is exchangeable, and

iv.) the $k_d \times k_d$ real matrix $Q = \Sigma_d^{-1/2}\mathbb{E}^{x_d}[B_d^{-1}\Sigma_d^{-1}B_d^{-1}]\Sigma_d^{-1/2}$ is known and positive definite.

Assumption iii.) implies $P\Sigma_dP^t = \Sigma_d$, $PB_dP^t = B_d$, and $PB_I = B_I$ for every permutation matrix $P$. The equality $PB_I = B_I$ says $B_I$ has equal rows and arbitrary columns, while the first two equalities indicate $\Sigma_d$ and the random matrix $B_d$ are
completely symmetric matrices (see Section 1.3.3). By Corollaries 1.3.20 and 1.3.19, the matrix $Q$ in Assumption iv.) is also completely symmetric.

Assumption iv.) may appear contrived. However, if we consider the results of Corollaries 1.3.18 and 1.3.20, then we see that it is equivalent to

$$E^x[B_d^{-2}] = \Sigma_d^{1/2}Q\Sigma_d^{1/2},$$

(3.2.1)
i.e., knowing the expected value of $B_d^{-2}$ with respect to $\pi_B$. We could have defined $Q$ in Assumption iv.) by (3.2.1), yet we chose the more cumbersome definition because it is in this form that we encounter $Q$ in Section 3.3.

The matrix $Q$ has at most two distinct positive eigenvalues. Let $\gamma > 0$ be the eigenvalue with multiplicity one, $\lambda > 0$ be the eigenvalue with multiplicity $k_d - 1$, and note that $\gamma$ and $\lambda$ may be equal. In addition, let $R$ be a $k_d \times k_d$ real matrix whose columns are the eigenvectors for the set of completely symmetric matrices (Lemma 1.3.16), and let $\Sigma_d^{1/2}$ be the positive definite completely symmetric matrix that satisfies $\Sigma_d = \Sigma_d^{1/2}\Sigma_d^{1/2}$ (Corollary 1.3.20). Then, the positive definite matrix $R\Sigma_d^{-1/2}$ simultaneously diagonalizes the matrices $\Sigma_d$ and $E^x[B_d^{-1}\Sigma_d^{-1}B_d^{-1}]$ as

$$(R\Sigma_d^{-1/2})\Sigma_d(R\Sigma_d^{-1/2})^t = I_{kd}$$

(3.2.2)

and

$$(R\Sigma_d^{-1/2})E^x[B_d^{-1}\Sigma_d^{-1}B_d^{-1}](R\Sigma_d^{-1/2})^t = \Lambda_Q$$

(3.2.3)

where $\Lambda_Q = \text{diag}(\gamma, \lambda, \ldots, \lambda)$. The simultaneous diagonalization in (3.2.2) and (3.2.3) greatly simplifies the design problem in Section 3.5. For the IavA criterion, we define
the space $X_d$ as

$$X_d = \left\{ s_d \Sigma_d^{1/2} R'u + \mu_d^*: u \in [-1,1]^{k_d} \right\}$$  \hspace{1cm} (3.2.4)

where $s_d$ is a given sizing scalar.

The prior assumptions that we have made for the IavA criterion are similar to those made by Chatterjee and Mandal (1981) for the problem of estimating the optimal point of a quadratic response surface. They minimized the “deficiency” criterion

$$\text{tr} \mathbb{E}^\pi \left[ \mathbb{E}^\pi \left[ (\hat{x}_d^* - x_d^*) B_d (\hat{x}_d^* - x_d^*)^t \right] \right].$$  \hspace{1cm} (3.2.5)

using Assumptions i.) and ii.) and by assuming the value of $\mathbb{E}^\pi B_d^{-1}$ is known. Considering the form of Assumption iv.) given by (3.2.1), we see their approach lacked only Assumption iii.) (exchangeability for the components of $x_d^*$).

When $k_d = 2$, they determined an optimal design for arbitrary symmetric and positive definite $\Sigma_d$ and $B_d$. However, when $k_d > 2$, they could not obtain an optimal design unless $\mathbb{E}^\pi B_d^{-1} = \lambda \Sigma_d$. Note that $\mathbb{E}^\pi B_d^{-2} = \lambda \Sigma_d^2$ in (3.2.1) if $\Lambda_Q = \lambda I_{k_d}$. In Section 3.5 it will be shown that using Assumptions 3.2.1 allows us to obtain optimal designs for all values of $\gamma, \lambda > 0$. In particular, the exchangeability assumption makes the design problem tractable.

For the IavD design criterion (Design Criterion 1.2.6) we add the following distributional assumptions to the design problem.

**Assumptions 3.2.2** The parameter $x_d^*$ is independent of the random variables $\beta_0$, $B_d$, $\beta_e$ and $B_I$, and it is distributed according to $\pi_d^*$ for which

1.) $\mu_d^* = \mathbb{E}^\pi [x_d^*] \in \mathbb{R}^{k_d}$ is known, and
ii.) the $k_d \times k_d$ real matrix $\Sigma_d = E^x_d[(x_d^* - \mu_d^*)^t (x_d^* - \mu_d^*)] \text{ is positive definite and known.}$

When determining IavD optimal designs, an appropriate choice for the design space $X_d$ is

$$X_d = \{ s_d \Sigma_d^{1/2} u + \mu_d^*: u \in [-1,1]^{k_d} \}. \quad (3.2.6)$$

From (3.2.4) and (3.2.6), we see the design space $X_d$ is centered about the prior "best guess" $\mu_d^*$ of the unknown optimal operating conditions $x_d^*$, and its orientation is determined, in part, by the variance-covariance matrix $\Sigma_d$.

### 3.2.2 Defining $X_e$

Our knowledge regarding the environmental variables $x_e$ is the same in this chapter as it was in Section 2.2.2. Therefore, let $\mu_e = E^x_e[x_e]$ and let $\Sigma_e = E^x_e[(x_e - \mu_e)(x_e - \mu_e)^t]$, a positive definite matrix. The design space $X_e$ for the environmental factors $x_e$ is

$$X_e = \{ s_e \Sigma_e^{1/2} v + \mu_e: v \in [-1,1]^{k_e} \} \quad (3.2.7)$$

where $\Sigma_e^{1/2}\Sigma_e^{1/2} = \Sigma_e$ and $s_e$ is a user-supplied scaling constant.

### 3.3 Reparameterizing the Model

In this section we define a reparameterization of the statistical model in (1.2.3). Recall we defined the design space $X$ as $X = X_d \times X_e$ where $X_e = \{ s_e \Sigma_e v + \mu_e: v \in [-1,1]^{k_e} \}$
and \( X_d = \{ s_d T u + \mu_d^*: u \in [-1, 1]^k_d \} \) with the matrix \( T \) equal to either \( \Sigma_d^{1/2} R^t \) or \( \Sigma_d^{1/2} \). If \( \xi^* \) is an optimal design measure, defined on the Borel sets of \([ -1, 1 ]^{k_d + k_e} \), then the design measure \( \xi_{\hat{X}}^* \) defined on the Borel sets of \( X \) and given by

\[
\xi_{\hat{X}}^*(s_d T u + \mu_d^*, s_e \Sigma_e^{1/2} v + \mu_e) = \xi^*(u, v) \text{ for all } (u, v) \in [-1, 1]^{k_d + k_e}.
\]

is optimal in the set of all design measures defined on the Borel \( \sigma \)-field of \( X \). This means we can determine optimal designs exclusively on the region \([ -1, 1 ]^{k_d + k_e} \), and any optimal design obtained on \([ -1, 1 ]^{k_d + k_e} \) can be mapped directly to an optimal design on \( X \).

Replacing the vectors \( x_d \) and \( x_e \) in (1.2.3) with the vectors \( s_d T u + \mu_d^* \) and \( s_e \Sigma_e^{1/2} v + \mu_e \), respectively, the statistical model in (1.2.3) becomes

\[
Y(u, v) = \alpha_0 + \alpha_d^t u + u^t A_d u + \alpha_e^t v + u^t A_e v + Z(u, v)
\]

where

\[
\alpha_0 = \beta_0 + \beta_d^t \mu_d^* + \mu_d^t B_d \mu_d^* + \mu_e^t B_e \mu_e \in \mathbb{R}
\]

\[
\alpha_d = s_d T^t (\beta_d + 2 B_d \mu_d^* + B_e \mu_e) \in \mathbb{R}^{k_d}
\]

\[
A_d = s_d^2 T^t B_d T \in \mathbb{R}^{k_d \times k_d} \text{ is positive definite}
\]

\[
\alpha_e = s_e \Sigma_e^{1/2} (\beta_e + B_e^t \mu_d^*) \in \mathbb{R}^{k_e}
\]

\[
A_e = s_d s_e T^t B_e \Sigma_e^{1/2} \in \mathbb{R}^{k_d \times k_e}.
\]

The error term \( Z(\cdot, \cdot) \) in (3.3.2) is a mean zero Gaussian process with constant variance \( \sigma^2 > 0 \) and zero correlations between pairs of distinct input sites.
3.3.1 Reparameterizing the Parameter Design Problem

The vector $v$ of standardized environmental variables equals $s_c^{-1}\Sigma_c^{-1/2}(x_e - \mu_e)$. Therefore, it is governed by the known probability distribution $\pi_c$ for the environmental variables $x_e$. The mean vector for $v$ is

$$E_\pi^c[v] = s_c^{-1}\Sigma_c^{-1/2}E_\pi^c[x_e - \mu_e] = 0_{k_c}, \quad (3.3.8)$$

and its variance-covariance matrix is

$$E_\pi^c[vv'] = s_c^{-2}\Sigma_c^{-1/2}E_\pi^c[(x_e - \mu_e)(x_e - \mu_e)'](x_e - \mu_e)'\Sigma_c^{-1/2} = s_c^{-2}I_{k_c}. \quad (3.3.9)$$

Using (3.3.2) and (3.3.8) the expected integrated loss function in (1.2.4) can be reexpressed as a function of the standardized control variables $u$, i.e.,

$$L(u) = \alpha_o + \alpha_d'u + u'A_du. \quad (3.3.10)$$

The matrix $A_d$ in (3.3.10) is positive definite, so the function $L(u)$ is minimized, uniquely, at the point

$$u^* = -\frac{1}{2}A_d^{-1}\alpha_d = s_d^{-1}T^{-1}(x_d^* - \mu_d^*). \quad (3.3.11)$$

The second equality in (3.3.11) is obtained by substitution of $\alpha_d$ and $A_d$ from (3.3.4) and (3.3.5). This expression shows that the minimizing point $u^* \in [-1, 1]^{k_d}$ is the image of $x_d^* \in X_d$ for the inverse of the nonsingular mapping that was used to define $X_d$. This means the vector $u^*$ is random with probability law governed by the prior for $x_d^*$. 
Let \( \pi_d^* \) be the conditional distribution for \( x_d^* \) given \( \beta_0, B_d, \beta_e \) and \( B_I \), and let \( \pi_B^* \) be the marginal distribution for \( \beta_0, B_d, \beta_e \) and \( B_I \). The expected value of \( u^* \), with respect to \( \pi_d^* \), is

\[
E^\pi_d[u^*|\alpha_0, A_d, \alpha_e, A_I] = \mathbf{s}_d^{-1} \mathbf{T}^{-1} E^\pi_d([x_d^* - \mu_d^*]|\beta_0, B_d, \beta_e, B_I) = 0_k, \tag{3.3.12}
\]

and its variance-covariance matrix is

\[
E^\pi_d[u^*u^*']|\alpha_0, A_d, \alpha_e, A_I] = s_d^{-2} \mathbf{T}^{-1} E^\pi_d((x_d^* - \mu_d^*)(x_d^* - \mu_d^*)'|\beta_0, B_d, \beta_e, B_I) \mathbf{T}^{-1} = s_d^{-2} \Sigma_d \mathbf{T}^{-1} \mathbf{T}^{-1} = s_d^{-2} \mathbf{I}_{k_d}. \tag{3.3.13}
\]

This completes the required moments conditions for the IavD design criterion (see Assumptions 3.2.2). However, we have yet to find the expected value of \( A_d^{-2} \) which is required for the IavA design criterion (see Assumptions 3.2.1). Recall that for the IavA criterion the matrix \( \mathbf{T} \) equals \( \Sigma_d^{1/2} \mathbf{R}' \), hence

\[
A_d^{-2} = s_d^{-4} \mathbf{R} \Sigma_d^{-1/2} B_d^{-1} \Sigma_d^{-1} B_d^{-1} \Sigma_d^{-1/2} \mathbf{R}'. \tag{3.3.14}
\]

Assumption iv) of Assumptions 3.2.1 says the expected value of (3.3.14) with respect to the marginal prior \( \pi_B^* \) is

\[
E^\pi_B[A_d^{-2}] = s_d^{-4} \mathbf{A}_Q. \tag{3.3.15}
\]
3.4 The Asymptotic Variance of the MLE

To determine the asymptotic variance-covariance matrix for the maximum likelihood estimator \( \hat{u}^* \) of \( u^* \), expand the coefficient arrays and rewrite the model in (3.3.2) as

\[
Y(u, v) = \alpha_0 + \sum_{i=1}^{k_d} \alpha_{di}x_{di} + \sum_{i=1}^{k_d} \alpha_{dij}^2 x_{dij}^2 + \sum_{1 \leq i < j \leq k_d} \alpha_{dij} x_{dij} x_{dij} + \sum_{i=1}^{k_r} \alpha_c x_{ci} + \sum_{i=1}^{k_d} \sum_{j=1}^{k_r} \alpha_{lij} x_{dij} x_{cij} + Z(u, v)
\]  

(3.4.1)

where

\[
\alpha_{di} = (\alpha_d)_i \text{ for } i \in \{1, \ldots, k_d\}
\]

(3.4.2)

\[
\alpha_{dij} = \frac{2}{1 + \delta_{ij}} (A_d)_{i,j} \text{ for } i < j \in \{1, \ldots, k_d\}
\]

(3.4.3)

\[
\alpha_c = (\alpha_c)_i \text{ for } i \in \{1, \ldots, k_c\}
\]

(3.4.4)

\[
\alpha_{lij} = (A_{ij})_{i,j} \text{ for } i \in \{1, \ldots, k_d\} \text{ and } j \in \{1, \ldots, k_r\}.
\]

(3.4.5)

The term \( \delta_{ij} \) in (3.4.3) is the Kroneker delta function (see page 17). Then, consider (3.4.1) expressed in the form

\[
Y(u, v) = f^T \alpha + Z(u, v)
\]

(3.4.6)

where \( \alpha \in \mathbb{R}^{(k_d + 2) + (k_d + 1)k_r} \) equals

\[
\alpha^T = (\alpha_0, \alpha_d^T, \alpha_{d11}, \ldots, \alpha_{dk_2k_d}, \alpha_{d12}, \ldots, \alpha_{dk_{d-1}k_d}, \alpha_{c1}, \ldots, \alpha_{k_dk_r})
\]

(3.4.7)

and \( f \) is the associated vector of known regression functions of \( u \) and \( v \).

Because of the distributional assumptions that we have imposed on the stochastic error term \( Z(\cdot, \cdot) \) in (3.4.1), for a given design \( \xi \), the asymptotic variance-covariance
matrix for the consistent estimator $\hat{u}^*$ of $u^*$ is proportional to

$$\left(\frac{\partial u^*}{\partial \alpha^i}\right) M(\xi)^{-1} \left(\frac{\partial u^*}{\partial \alpha^i}\right)^t$$  \hspace{1cm} (3.4.8)

where $M(\xi)^{-1}$ is the Moore-Penrose generalized inverse of the information matrix $M(\xi)$, and $(\partial u^*/\partial \alpha^i)$ is the gradient of $u^*$ with respect to the vector $\alpha$ in (3.4.7) (Silvey, 1980). The gradient $(\partial u^*/\partial \alpha^i)$ is a $k_d \times \left\{ \binom{k_d+2}{2} + (k_d + 1)k_r \right\}$ matrix that depends on the value $u^*$. To determine $(\partial u^*/\partial \alpha^i)$, we partition the vector $\alpha$ into five pieces

$$\alpha^i = (\alpha_0 | \alpha_d^1 | \alpha_d^{12}, \ldots, \alpha_{d_{d{k_d}}} | \alpha_{d_{d+1,k_d}}, \ldots, \alpha_{d_{d-k_d,k_d}}, \alpha_{k_d+1,k_d}, \ldots, \alpha_{k_d+k_r})$$  \hspace{1cm} (3.4.9)

then we find the gradient of $u^*$ with respect to each piece.

A.) The column with respect to $\alpha_0$ is

$$\left(\frac{\partial u^*}{\partial \alpha_0}\right) = 0_{k_d}.$$  \hspace{1cm} (3.4.10)

B.) The $k_d$ columns with respect to $\alpha_d^i$ are

$$\left(\frac{\partial u^*}{\partial \alpha_d^i}\right) = -\frac{1}{2} A_d^{-1}.$$  \hspace{1cm} (3.4.11)

C.) The $k_d$ columns with respect to the coefficients $\alpha_d^{i}, i \in \{1, \ldots, k_d\}$ are

$$\left(\frac{\partial u^*}{\partial \alpha_d^{i}}, \ldots, \frac{\partial u^*}{\partial \alpha_d^{k_d}}\right) = -A_d^{-1} \Lambda u^*$$  \hspace{1cm} (3.4.12)

where $\Lambda u^* = \text{diag}(u^*_1, \ldots, u^*_k)$ is the $k_d \times k_d$ diagonal matrix with entry $u^*_i$ in the $i^{th}$ row (column).

D.) The $\left(\frac{k_d}{2}\right)$ columns with respect to the coefficients $\alpha_d^{i,j}$ for $1 \leq i < j \leq k_d$ are

$$\left(\frac{\partial u^*}{\partial \alpha_d^{i,j}}, \ldots, \frac{\partial u^*}{\partial \alpha_d^{k_d-k_d}}\right) = -\frac{1}{2} A_d^{-1} \Gamma u^*$$  \hspace{1cm} (3.4.13)
where

\[
\Gamma_{\mathbf{u}^*} = \begin{pmatrix}
\mathbf{u}_{2}^* & \mathbf{u}_{3}^* & \ldots & \mathbf{u}_{k_d}^* & 0 & \cdots & 0 \\
\mathbf{u}_{1}^* & 0 & \cdots & 0 & \mathbf{u}_{3}^* & 0 \\
0 & \mathbf{u}_{1}^* & \cdots & 0 & \mathbf{u}_{2}^* & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 0 & \cdots & \mathbf{u}_{1}^* & 0 & \mathbf{u}_{k_d}^* \\
0 & 0 & \cdots & \mathbf{u}_{1}^* & 0 & \mathbf{u}_{k_d-1}^*
\end{pmatrix}.
\]  

(3.4.14)

E.) The \((k_d + 1)k_e\) columns with respect to the arrays \(\alpha_e\) and \(A_f\) are

\[
\begin{pmatrix}
\frac{\partial \mathbf{u}^*}{\partial \alpha_e}, \frac{\partial \mathbf{u}^*}{\partial \alpha_{f_{11}}}, \frac{\partial \mathbf{u}^*}{\partial \alpha_{f_{12}}}, \ldots, \frac{\partial \mathbf{u}^*}{\partial \alpha_{f_{k_d k_e}}} \end{pmatrix} = 0_{k_d \times (k_d + 1)k_e}.
\]  

(3.4.15)

We can factor the matrix \(-A_d^{-1}\) out of each of the partials in (3.4.10), (3.4.11), (3.4.12), (3.4.13) and (3.4.15). Thus, the gradient of \(\mathbf{u}^*\) with respect to \(\alpha\) equals

\[
\frac{\partial \mathbf{u}^*}{\partial \alpha^i} = -A_d^{-1}\mathbf{H}
\]  

(3.4.16)

where the \(k_d \times \left\{ \binom{k_d+2}{2} + (k_d + 1)k_e \right\}\) real matrix

\[
\mathbf{H} = (0_{k_d}, \frac{1}{2}I_{k_d}, A_{\mathbf{u}^*}, \frac{1}{2}\Gamma_{\mathbf{u}^*}, 0_{k_d \times (k_d + 1)k_e}).
\]  

(3.4.17)

Notice that \(\mathbf{H}\) depends on the coordinates of \(\mathbf{u}^*\) only. Using (3.4.8) and (3.4.16), the value of the asymptotic dispersion matrix for the maximum likelihood estimator \(\hat{\mathbf{u}}^*\) is proportional to

\[
\left(\frac{\partial \mathbf{u}^*}{\partial \alpha^i}\right) \mathbf{M}(\xi)^{-1} \left(\frac{\partial \mathbf{u}^*}{\partial \alpha^i}\right)^t = A_d^{-1}\mathbf{H}(\xi)^{-1}\mathbf{H}^t A_d^{-1}.
\]  

(3.4.18)

for every \(\xi \in \Xi_1\).
3.5 IavA Optimal Designs

In this section we find an experimental design, defined on \([-1, 1]^{k_x+k_r}\), that minimizes the expected value of the trace of the asymptotic variance-covariance matrix (1.2.20). The expectation of the trace is taken with respect to the prior distribution \(\pi^* = \pi_d^*\pi_n^*\) for which the hypotheses in Assumptions 3.2.1 are satisfied. The goal of this section is to find a design measure \(\xi \in \Xi_1\) that minimizes

\[
E^{\pi^*} \left[ \operatorname{tr} \left( \frac{\partial u^*}{\partial \alpha^i} M(\xi)^{-1} \left( \frac{\partial u^*}{\partial \alpha^i} \right)^\top \right) \right]
\]

over \(\Xi_1\). Using (3.4.18) and the properties of the trace operator, the design problem in (3.5.1) is equivalent to finding a design \(\xi \in \Xi_1\) that minimizes

\[
\operatorname{tr} \left[ M(\xi)^{-1} E^{\pi^*} [H^\top A_d^{-2} H] \right].
\]

over \(\Xi_1\). The design function in (3.5.2) is the convex and differentiable function of Example 1.3.1. Therefore, we can use the results of Theorem 1.3.4 to determine an optimal design.

We cannot give the IavA optimal design in a closed form, like the designs of Theorem 2.5.1, Corollary 2.5.2 and Theorem 2.6.1. Instead, we can only describe the form of its information matrix which depends on four numbers — \(r_1, r_2, t_1\) and \(t_2\). The values of these numbers are the solutions to a set of equations and cannot be given explicitly. For a given experimental situation, the values of these numbers must be obtained using numeric optimization methods to solve these equations which are given in the proof of the theorem.
Theorem 3.5.1 (IavA Optimal Design) Assume the model in (3.3.2), and let $\pi^*$ be a prior distribution for which the hypotheses in Assumptions 3.2.1 are satisfied. There exists an IavA optimal design $\xi^* \in \Xi_1$, with respect to the prior $\pi^*$, for which

i.) $E^\xi u_i q_i = 0$ where each $i_j \in \{1, 2, \ldots, k_d\}$, each $q_i_j \in \{0, 1, 2, 3\}$

with $0 \leq \sum_{j=1}^{4} q_i_j \leq 4$, and at least one $q_i_j$ is odd,

ii.) $E^\xi u_i^2 = r_1$,

iii.) $E^\xi u_i^2 = E^\xi u_i^4 = r_2$ for each $i \in \{2, \ldots, k_d\}$,

iv.) $E^\xi u_i^2 u_j^2 = t_1$ for each $i \in \{2, \ldots, k_d\}$,

v.) $E^\xi u_i^2 u_j^2 = t_2$ for each $i, j \in \{2, \ldots, k_d\}$ with $i < j$, and

vi.) $E^\xi u_i q_i q_j q_k = 0$ where each $i_j \in \{1, 2, \ldots, k_d\}$, each $q_i_j \in \{0, 1, 2, 3\}$

with $0 \leq \sum_{j=1}^{4} q_i_j \leq 3$, and $k \in \{1, 2, \ldots, k_e\}$.

where $r_1$, $r_2$, $t_1$ and $t_2$ are the solutions to a complicated set of equations given in the proof of the theorem (see (3.5.46) and the discussion that follows).

Proving Theorem 3.5.1 requires knowing $M^- (\xi^*)$ and $E^\xi [H^d A_d^{-2} H]$. The calculation of these matrices is tedious. Therefore, we only give the values of the relevant matrices in a sequence of lemmas and relegate the calculations to Appendix A.

Lemma 3.5.2 Let $\pi^*$ be a probability distribution for which the hypotheses in Assumptions 3.2.1 are satisfied, and let $H$ be defined by (3.4.17). Then, $E^\pi [H^d A_d^{-2} H]$
equals
\[
\begin{pmatrix}
0 & 0_{k_d} & 0_{k_d} & 0_{(k_d+1)k_r} \\
0_{k_d} & \frac{s_d^4}{4} \Lambda_Q & 0_{k_d \times k_d} & 0_{k_d \times (k_d+1)k_r} \\
0_{k_d} & 0_{k_d \times k_d} & s_d^{-6} \Lambda_Q & 0_{k_d \times (k_d+1)k_r} \\
0_{(k_d+1)k_r} & 0_{(k_d+1)k_r \times k_d} & 0_{(k_d+1)k_r \times k_d} & 0_{(k_d+1)k_r \times (k_d+1)k_r}
\end{pmatrix}
\] (3.5.3)

where the matrices \( \Lambda_Q \) and \( \Lambda_Q T \) on the diagonal are defined in (3.2.3) and (A.1.17), respectively. Both are diagonal; therefore, \( E^* [H' A_d^{-2} H] \) is diagonal.

Recall the form of \( \alpha \) in (3.4.7), and consider the following partition

\[
\alpha' = \begin{pmatrix} \alpha_0, \alpha_d', \alpha_{d1}, \cdots, \alpha_{d_k k_d} & \alpha_{d12}, \alpha_{d13}, \cdots, \alpha_{d_{k_d-1} k_d} & \alpha_{k_d}' \end{pmatrix}
\]

The matrices in the following lemmas are partitioned in a manner corresponding to (3.5.4).

**Lemma 3.5.3** Let \( \xi^* \) be a design, defined on \([-1, 1]^{k_d+k_r} \), for which the six conditions in Theorem 3.5.1 are satisfied. The information matrix \( M(\xi^*) \) for the vector \( \alpha \) is

\[
\begin{pmatrix}
M_d & 0_{(2k_d+1) \times \left( \frac{k_d}{2} \right)} & 0_{(2k_d+1) \times (k_d+1)k_r} \\
0_{\left( \frac{k_d}{2} \right) \times (2k_d+1)} & \Lambda_T & 0_{\left( \frac{k_d}{2} \right) \times (k_d+1)k_r} \\
0_{(k_d+1)k_r \times (2k_d+1)} & 0_{(k_d+1)k_r \times \left( \frac{k_d}{2} \right)} & \int f_3 f_3' d\xi^*
\end{pmatrix}
\] (3.5.5)

where \( M_d \) is the information matrix for the regression coefficients in \( \alpha \), the diagonal matrix \( \Lambda_T = \text{diag}(t_1 I_{k_d-1}, t_2 I_{\left( \frac{k_d}{2} \right)-1}) \) is the information matrix for the coefficients in
and \( f f_3 f'_3 d \xi^* \) is the information matrix for the coefficients in \( \alpha_3 \). The matrix \( M_d \) equals
\[
M_d = \begin{pmatrix}
M_{d11} & M_{d12} \\
M_{d12} & M_{d22}
\end{pmatrix}
\]
where
\[
M_{d11} = \begin{pmatrix}
1 & 0_{kd} \\
0_{kd} & \Lambda_R
\end{pmatrix},
\]
with \( \Lambda_R = \text{diag}(r_1, r_2 I_{k_d-1}) \),
\[
M_{d12} = \begin{pmatrix}
r_1 & r_2 1_{k_d-1} \\
0_{kd} & 0_{kd \times (k_d-1)}
\end{pmatrix},
\]
and
\[
M_{d22} = \begin{pmatrix}
r_1 & t_1 1_{k_d-1} \\
t_1 1_{k_d-1} & (r_2 - t_2) I_{k_d-1} + t_2 J_{k_d-1}
\end{pmatrix}.
\]

**Lemma 3.5.4** The Moore-Penrose generalized inverse for \( M(\xi^*) \) in Lemma 3.5.3 is
\[
M^{-1}(\xi^*) = \begin{pmatrix}
M_d^{-1} & 0_{(2k_d+1) \times (k_d^3/2)} & 0_{(2k_d+1) \times (k_d+1)k_r} \\
0_{(k_d^3/2) \times (2k_d+1)} & \Lambda_T^{-1} & 0_{(k_d^3/2) \times (k_d+1)k_r} \\
0_{(k_d+1)k_r \times (2k_d+1)} & 0_{(k_d+1)k_r \times (k_d^3/2)} & (f f_3 f'_3 d \xi^*)^{-1}
\end{pmatrix}
\]
where \( (f f_3 f'_3 d \xi^*)^{-1} \) is the Moore-Penrose generalized inverse of \( f f_3 f'_3 d \xi^* \), and the diagonal
\[
\Lambda_T^{-1} = \text{diag}(t_1^{-1} I_{k_d-1}, t_2^{-1} I_{(k_d^3/2)}}). \]
Let
\[
w = (r_1 - r_1^2)(r_2 - r_2^2 + (k_d - 2)(t_2 - r_2^2)) - (k_d - 1)(t_1 - r_1 r_2)^2,
\]
then \( M_d^{-1} \) equals
\[
\begin{pmatrix}
M_{d11}^{-1} & M_{d12}^{-1} \\
M_{d12}^{-1} & M_{d22}^{-1}
\end{pmatrix}
\]
where

\[ M^{-1}_{d11} = \begin{pmatrix} r_1r_2 + (k_d - 2)r_1t_2 - (k_d - 1)t^2_1 & 0_k^t_d \\ 0_k & \Lambda^{-1}_R \end{pmatrix}, \tag{3.5.13} \]

with \( \Lambda^{-1}_R = \text{diag}(r_1^{-1}, r_2^{-1}I_{k_d - 1}) \),

\[ M^{-1}_{d12} = \begin{pmatrix} -\frac{r_1r_2 + (k_d - 2)r_1t_2 - (k_d - 1)t^2_1}{w} & -\frac{r_1r_2 - r_1t_1}{w} 1^t_{k_d - 1} \\ 0_k & 0_k \times (k_d - 1) \end{pmatrix}, \tag{3.5.14} \]

and

\[ M^{-1}_{d22} = \begin{pmatrix} \frac{r_1 - r_1t_2}{w} & -\frac{1}{r_2 - t_2} 1^t_{k_d - 1} \\ -\frac{1}{r_2 - t_2} 1_{k_d - 1} & -\frac{(r_1 - r_1t_2)(t_2 - r_2^2) - (t_1 - r_1t_1)^2}{w} I_{k_d - 1} - \frac{(r_1 - r_1t_2)(t_2 - r_2^2) - (t_1 - r_1t_1)^2}{w} J_{k_d - 1} \end{pmatrix}. \tag{3.5.15} \]

The following Corollaries follow directly from the previous lemmas. The results are obtained by simple matrix multiplications. Therefore, each corollary is presented without proof.

**Corollary 3.5.5**: The matrix \( M^{-1}(\xi^*)E^{\star}\{H'\Lambda^{-2}H\} \) equals

\[
\begin{pmatrix}
M_{11} & M_{12} & 0_{(k_d + 1) \times \left(k_d^2\right)} & 0_{(k_d + 1) \times (k_d + 1)k_r} \\
0_{k_d \times (k_d + 1)} & M_{22} & 0_{k_d \times \left(k_d^2\right)} & 0_{k_d \times (k_d + 1)k_r} \\
0_{(k_d^2) \times (k_d + 1)} & 0_{(k_d^2) \times k_d} & \frac{3}{4} \Lambda^{-1}_Q & 0_{(k_d^2) \times (k_d + 1)k_r} \\
0_{(k_d + 1)k_r \times (k_d + 1)} & 0_{(k_d + 1)k_r \times k_d} & 0_{(k_d + 1)k_r \times \left(k_d^2\right)} & 0_{(k_d + 1)k_r \times (k_d + 1)k_r}
\end{pmatrix}.
\tag{3.5.16}
\]

where

\[
M_{11} = \begin{pmatrix} 0 & 0_{k_d} \\ 0_{k_d} & \frac{3}{4} \Lambda^{-1}_Q \end{pmatrix}, \tag{3.5.17} \]

\[
M_{12} = \begin{pmatrix} -s^2 \gamma_1r_1r_2 + (k_d - 2)r_1t_2 - (k_d - 1)t^2_1 & -s^2 \gamma_1r_1(r_2 - t_1) 1^t_{k_d - 1} \\
0_{k_d} & 0_{k_d \times (k_d - 1)} \end{pmatrix}, \tag{3.5.18} \]
and
\[
M_{22} = \begin{pmatrix}
\frac{s_d^{-6} \gamma (r_2 - r_2^2 + (k_d - 2)(t_2 - r_2^2))}{w} & -\frac{s_d^{-6} \lambda (r_1 - t_1 r_2)}{w} \\
-\frac{s_d^{-6} \lambda (r_1 - t_1 r_2)}{w^2} \mathbf{1}_{k_d-1} & \frac{s_d^{-6} \lambda (r_1 - t_1 r_2)}{w^2} \left( \mathbf{1}_{k_d-1} - \frac{(r_1 - r_2^2)(t_2 - r_2^2) - (t_1 - r_1 r_2)^2}{w} \mathbf{J}_{k_d-1} \right)
\end{pmatrix}.
\]

(3.5.19)

Corollary 3.5.6 The matrix \(M^{-}(\xi^{*})E^{*} [H'A_d^{-2} H]M^{-}(\xi^{*})\) is equal to
\[
\begin{pmatrix}
L_{11} & L_{12} & 0_{(k_d+1) \times (k_d \times (k_d+1)k_r)} & 0_{(k_d+1) \times (k_d \times (k_d+1)k_r)} \\
L_{12} & L_{22} & 0_{k_d \times (k_d \times (k_d+1)k_r)} & 0_{k_d \times (k_d \times (k_d+1)k_r)} \\
0_{(k_d+1) \times (k_d \times (k_d+1)k_r)} & 0_{(k_d \times (k_d+1)k_r \times k_d)} & \frac{s_d^{-6}}{4} A_T^{-1} \Lambda Q \Lambda^{-1} & 0_{(k_d \times (k_d+1)k_r \times (k_d \times (k_d+1)k_r)} \\
0_{(k_d+1)k_r \times (k_d \times (k_d+1)k_r)} & 0_{(k_d \times (k_d+1)k_r \times (k_d \times (k_d+1)k_r)} & 0_{(k_d \times (k_d+1)k_r \times (k_d \times (k_d+1)k_r)} \\
\end{pmatrix}
\]

(3.5.20)

where
\[
L_{11} = \begin{pmatrix}
\frac{s_d^{-6} \gamma (r_1 r_2 + (k_d - 2)r_1 t_2 - (k_d - 1)r_2 t_1)^2}{w^2} + \frac{s_d^{-6} \lambda (r_1 r_2 + (k_d - 2)r_1 t_2 - (k_d - 1)r_2 t_1)^2}{w^2} & \mathbf{0}_{k_d} \\
\mathbf{0}_{k_d} & \frac{s_d^{-6} \lambda (r_1 r_2 + (k_d - 2)r_1 t_2 - (k_d - 1)r_2 t_1)^2}{w^2} \end{pmatrix}
\]

(3.5.21)

\[
L_{12} = \begin{pmatrix}
(L_{12})_{11} & (L_{12})_{12} \\
(L_{12})_{21} & (L_{12})_{22}
\end{pmatrix}
\]

(3.5.22)

with
\[
(L_{12})_{11} = s_d^{-6} \gamma \left( \frac{r_1 r_2 + (k_d - 2)r_1 t_2 - (k_d - 1)r_2 t_1}{w} \right) \left( \frac{r_1 r_2 + (k_d - 2)r_1 t_2 - (k_d - 1)r_2 t_1}{w} \right)
\]

(3.5.23)

and
\[
(L_{12})_{12} = \left\{ \frac{s_d^{-6} \gamma \left( \frac{r_1 r_2 + (k_d - 2)r_1 t_2 - (k_d - 1)r_2 t_1}{w} \right) \left( \frac{r_1 r_2 + (k_d - 2)r_1 t_2 - (k_d - 1)r_2 t_1}{w} \right)}{w^2} \mathbf{1}_{k_d-1} \right\}
\]

(3.5.24)
and

\[ L_{22} = \begin{pmatrix} (L_{22})_{11} & (L_{22})_{12} \\ (L_{22})_{21} & (L_{22})_{22} \end{pmatrix} \] (3.5.25)

with

\[ (L_{22})_{11} = s_d^{-6} \gamma \left( \frac{r_2 - r_2^2 + (k_d - 2)(t_2 - r_2^2)}{w} \right)^2 + s_d^{-6} \lambda (k_d - 1) \left( \frac{t_1 - r_1 r_2}{w} \right)^2, \] (3.5.26)

\[ (L_{22})_{12} = \left\{ -s_d^{-6} \gamma \left( \frac{r_2 - r_2^2 + (k_d - 2)(t_2 - r_2^2)}{w} \right) \left( \frac{t_1 - r_1 r_2}{w} \right) - s_d^{-6} \lambda \left( \frac{t_1 - r_1 r_2}{w} \right)^2 \right\} \mathbf{1}_{k_d - 1}^t, \] (3.5.27)

and

\[ (L_{22})_{22} = \frac{s_d^{-6} \lambda}{(r_2 - t_2)^2} \left\{ \mathbf{1}_{k_d - 1}^t - \frac{(r_1 - r_1^2)(t_2 - r_2^2) - (t_1 - r_1 r_2)^2}{w} \right\}^2 + s_d^{-6} \gamma \left( \frac{t_1 - r_1 r_2}{w} \right)^2 \mathbf{J}_{k_d - 1}. \] (3.5.28)

**Proof of Theorem 3.5.1:** Let \( f \) be the vector of regression functions of \((u, v)\) in (3.4.6). According to Condition ii.) of Theorem 1.3.4, a design measure \( \xi \in \Xi_1 \) is \( \text{IavA} \) optimal with respect to the prior \( \pi^* \) if and only if

\[
f^t \mathbf{M}^{-\gamma}(\xi) \mathbf{E}^{\pi^*} \left[ \mathbf{H}^t \mathbf{A}^{-2} \mathbf{H} \right] \mathbf{M}^{-\gamma}(\xi) f \leq \text{tr} \left[ \mathbf{M}^{-\gamma}(\xi) \mathbf{E}^{\pi^*} \left[ \mathbf{H}^t \mathbf{A}^{-2} \mathbf{H} \right] \right] \] (3.5.29)

for every \( (u, v) \in [-1, 1]^{k_d + k_r} \) (see Example 1.3.1). We need to only verify (3.5.29) for the candidate design \( \xi^* \) in Theorem 3.5.1 to prove the theorem.

To verify (3.5.29) for the candidate design \( \xi^* \), we shall first find the values of \((u, v) \in [-1, 1]^{k_d + k_r} \) at which the left hand side of (3.5.29) is maximized for the
design $\xi^*$. Then we determine the values of the parameters $r_1$, $r_2$, $t_1$ and $t_2$, so that the left hand side evaluated at these maximizing points equals the right hand side. For the values of the numbers $r_1$, $r_2$, $t_1$ and $t_2$ so obtained, the design $\xi^*$ in Theorem 3.5.1 is IavA optimal.

The left hand side of (3.5.29) is a quadratic form with symmetric matrix $M(\xi)^{-1}E^*\left[H^tA_d^{-2}H\right]M(\xi)^{-1}$ from Corollary 3.5.6. By the positive definiteness of $A_d$, $M(\xi)^{-1}E^*\left[H^tA_d^{-2}H\right]M(\xi)^{-1}$ is nonnegative definite. Moreover, the elements in the last $(k_d + 1)k_c$ rows and columns of this matrix are all equal to zero. Since the standardized environmental variables $v$ appear in only the last $(k_d + 1)k_c$ regression functions of the regression vector $f$, the quadratic form is a function of $u$ only. Considering the form of $f$ in (3.4.6), this implies the left hand side of (3.5.29) is a quartic function of the variables $u$ only.

The quartic on the left hand side of (3.5.29) is unchanged by permutations of the indices $\{2, 3, \ldots, k_d\}$ of the variables in $u$. Therefore, we separate the standardized controllable variables $u$ into two sets, $u_1$ and $\{u_2, \ldots, u_{k_d}\}$. We shall consider the left hand side of (3.5.29) both as a function of $u_1$ for arbitrary fixed $(u_2, \ldots, u_{k_d})$ and as a function of $u_i$ for arbitrary fixed $(u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_{k_d})$. It will be shown that both of these functions can only attain their maximum values at the points $\{-1, 0, 1\}$ thus implying the left hand side of (3.5.29) can only be maximized at points in the lattice $\{-1, 0, 1\}^{k_d}$. In the sequel, let $\theta = \gamma/\lambda$ be the ratio of the eigenvalues for the completely symmetric, positive definite matrix $Q$ defined in Hypothesis iv.) of Assumptions 3.2.1.
Suppose \((u_2, \ldots, u_{kd})\) are fixed in \([-1, 1]^{kd-1}\). The left hand side of (3.5.29) scaled by \((\lambda s_d^{-4})^{-1}\) and expressed as a function of \(u_1\) only is

\[
(\lambda s_d^{-4})^{-1} \mathbf{f}^* \mathbf{M}(\xi)^{-1} \mathbf{E}^* \mathbf{H}^\top \mathbf{A}_d^{-2} \mathbf{H} \mathbf{M}(\xi)^{-1} \mathbf{f} = a_0^{(1)} + a_2^{(1)} u_1^2 + a_4^{(1)} u_1^4
\]  

(3.5.30)

where

\[
a_0^{(1)} = \frac{\theta s_d^{-2} (r_1 r_2 + (k_d - 2) r_1 t_2 - (k_d - 1) r_2 t_1)^2}{w^2} + \frac{(k_d - 1) s_d^{-2} (r_2 - t_1)^2}{w^2} + \frac{2s_d^{-2} \theta (r_1 r_2 + (k_d - 2) r_1 t_2 - (k_d - 1) r_2 t_1) (t_1 - r_1 r_2)}{w^2} \sum_{i=2}^{k_d} u_i^2 + \\
\left\{ \frac{1}{4r_2^2} - \frac{2s_d^{-2} r_1^2 (1 - r_1)(r_2 - t_1)}{w^2} \right\} \sum_{i=2}^{k_d} u_i^2 + \left\{ \frac{s_d^{-2} (r_2 - t_2)^2}{w^2} - \frac{s_d^{-2}}{4r_2^2} \right\} \sum_{i=2}^{k_d} u_i^4 - \\
s_d^{-2} \left\{ \frac{(r_1 - r_2)(t_2 - r_2^2) - (t_1 - r_1 r_2)^2}{(r_2 - t_2)^2 w} \right\} \left( 1 + \frac{(r_1 - r_1^2)(r_2 - t_2)}{w} \right) \left( \sum_{i=2}^{k_d} u_i^2 \right)^2 + \\
\left\{ \frac{s_d^{-2}}{4t_2^2} + \frac{\theta s_d^{-2} (t_1 - r_1 r_2)^2}{w^2} \right\} \left( \sum_{i=2}^{k_d} u_i^2 \right)^2,
\]  

(3.5.31)

\[
a_2^{(1)} = \frac{\theta}{4r_1^2} + \frac{2s_d^{-2} (k_d - 1) r_1 (r_2 - t_1) (t_1 - r_1 r_2)}{w^2} - \\
\frac{2s_d^{-2} \theta (r_1 r_2 + (k_d - 2) r_1 t_2 - (k_d - 1) r_2 t_1) (r_2 - r_2^2 + (k_d - 2) (t_2 - r_2^2))}{w^2} - \\
\frac{2s_d^{-2} \theta (r_2 - r_2^2 + (k_d - 2) (t_2 - r_2^2) (t_1 - r_1 r_2))}{w^2} \sum_{i=2}^{k_d} u_i^2 - \\
\left\{ \frac{2s_d^{-2} (r_1 - r_1^2) (t_1 - r_1 r_2)}{w^2} - \frac{s_d^{-2} (\theta + 1)}{4t_1^2} \right\} \sum_{i=2}^{k_d} u_i^2,
\]  

(3.5.32)

and

\[
a_4^{(1)} = \frac{s_d^{-2}}{w^2} \left\{ \theta (r_2 - r_2^2 + (k_d - 2) (t_2 - r_2^2)^2 + (k_d - 1) (t_1 - r_1 r_2)^2) \right\}.
\]  

(3.5.33)

Similarly, fix \(i_*= \{2, \ldots, k_d\}\) and suppose \((u_1, \ldots, u_{i_*-1}, u_{i_*+1}, \ldots, u_{kd})\) are fixed in \([-1, 1]^{kd-1}\). The left hand side of (3.5.29) scaled by \((\lambda s_d^{-4})^{-1}\) expressed as a function
of $u_\star$ only is

$$(\lambda s_d^{-4})^{-1} f' \mathbf{M}(\xi) - E^* \left[ \mathbf{H}^\top \mathbf{A}_d^{-2} \mathbf{H} \right] \mathbf{M}(\xi)^{-1} \mathbf{f} = a^{(\star)}_0 + a^{(\star)}_2 \sum_{i \neq 1, \star} u^2_i + a^{(\star)}_4 u^4_i.$$  

(3.5.34)

where $a^{(\star)}_0$ equals

$$a^{(\star)}_0 = \frac{\theta s_d^{-2} (r_1 r_2 + (k_d - 2) r_1 t_2 - (k_d - 1) r_2 t_1)^2}{w^2} + \frac{(k_d - 1) s_d^{-2} r_1^2 (r_2 - t_1)^2}{w^2} +$$

$$\frac{\theta}{4 r^2_1} u_1^2 + \frac{2 s_d^{-2} (k_d - 1) r_1 (r_2 - t_1) (t_1 - t_2)}{w^2} u_1^2 -$$

$$\frac{2 s_d^{-2} \theta (r_1 r_2 + (k_d - 2) r_1 t_2 - (k_d - 1) r_2 t_1) (r_2 - r_2^2 + (k_d - 2) r_2^2)}{w^2} u_1^2 -$$

$$\frac{2 s_d^{-2} \theta (r_2 - r_2^2 + (k_d - 2) (t_2 - r_2^2)) (t_1 - r_1 t_2)}{w^2} u_1^2 \sum_{i \neq 1, \star} u^2_i -$$

$$\left\{ \frac{2 s_d^{-2} (r_1 - r_1^2) (t_1 - r_1 t_2)}{w^2} \right\} u_1^2 \sum_{i \neq 1, \star} u^2_i +$$

$$\frac{2 s_d^{-2} \theta (r_1 r_2 + (k_d - 2) r_1 t_2 - (k_d - 1) r_2 t_1) (t_1 - r_1 r_2)}{w^2} \sum_{i \neq 1, \star} u^2_i +$$

$$\left\{ \frac{1}{4 r^2_2} - \frac{2 s_d^{-2} r_1^2 (1 - r_1^2) (r_2 - t_1)}{w^2} \right\} \sum_{i \neq 1, \star} u^2_i + \left\{ \frac{s_d^{-2} - s_d^{-2}}{4 t^2_1} \right\} \sum_{i \neq 1, \star} u^4_i +$$

$$\frac{s_d^{-2}}{|M_d|^2} \left\{ \theta (r_2 - r_2^2 + (k_d - 2) (t_2 - r_2^2)^2 + (k_d - 1) (t_1 - r_1 r_2)^2) u_1^4 +$$

$$s_d^{-2} \left( \frac{(r_1 - r_1^2) (t_2 - r_2^2) - (t_1 - r_1 r_2)^2}{(r_2 - t_2)^2 w} \right) \left( 1 + \frac{(r_1 - r_1^2) (r_2 - t_2)}{w} \right) \left( \sum_{i \neq 1, \star} u^2_i \right)^2 +$$

$$\left\{ \frac{s_d^{-2}}{4 t^2_2} + \frac{s_d^{-2} (r_1 - r_1 r_2)^2}{w^2} \right\} \left( \sum_{i \neq 1, \star} u^2_i \right)^2 \right\} (3.5.35)$$

$a^{(\star)}_2 = \frac{2 s_d^{-2} \theta (r_1 r_2 + (k_d - 2) r_1 t_2 - (k_d - 1) r_2 t_1) (t_1 - r_1 r_2)}{w^2} \sum_{i \neq 1, \star} u^2_i +$

$$\left\{ \frac{1}{4 r^2_2} - \frac{2 s_d^{-2} r_1^2 (1 - r_1^2) (r_2 - t_1)}{w^2} \right\} \sum_{i \neq 1, \star} u^2_i + \left\{ \frac{s_d^{-2} - s_d^{-2}}{4 t^2_1} \right\} \sum_{i \neq 1, \star} u^4_i -$$

$$\frac{2 s_d^{-2} \theta (r_2 - r_2^2 + (k_d - 2) (t_2 - r_2^2) (t_1 - r_1 r_2)}{w^2} u_1^2 -$$
\( \left\{ \frac{2s_d^{-2}(r_1 - r_2^2)(t_1 - r_1 r_2)}{w^2} + \frac{s_d^{-2}(\theta + 1)}{4t_1^2} \right\} u_1^2 + 2s_d^{-2} \left( \sum_{i \neq 1, i_*} u_i^2 \right)^2 \)

\( \left( \frac{(r_1 - r_1^2)(t_2 - r_2^2) - (t_1 - r_1 r_2)^2}{(r_2 - t_2)^2 w} \right) \left( 1 + \frac{(r_1 - r_1^2)(r_2 - t_2)}{w} \right) \) + 

\( \left\{ \frac{2s_d^{-2} + 2\theta s_d^{-2}(t_1 - r_1 r_2)^2}{w^2} \right\} \left( \sum_{i \neq 1, i_*} u_i^2 \right)^2 , \) \hspace{1cm} (3.5.36)

and

\[ a_4^{(1,*)} = \frac{s_d^{-2}}{(r_2 - t_2)^2} \left( 1 - \frac{(r_1 - r_1^2)(t_2 - r_2^2) - (t_1 - r_1 r_2)^2}{w} \right) + \frac{\theta s_d^{-2}(t_1 - r_1 r_2)^2}{w^2} - \]

\[ \frac{s_d^{-2}}{(r_2 - t_2)^2 w^2} (r_1 - r_1^2)(r_2 - t_2)((r_1 - r_1^2)(t_2 - r_2^2) - (t_1 - r_1 r_2)^2) \]. \hspace{1cm} (3.5.37)

Since \( \mathbf{M}(\xi)^{-E^*}[\mathbf{H}^t \mathbf{A}_d^{-2} \mathbf{H}] \mathbf{M}(\xi)^{-} \) is nonnegative definite, the leading coefficients for the quartic expressions in (3.5.30) and (3.5.34) are nonnegative, i.e., \( a_4^{(1)} \geq 0 \) and \( a_4^{(i_*)} \geq 0 \). Furthermore, both polynomial expressions are even functions, so they each have a critical value at zero. This implies both quartic equations can only attain their maximum value over the interval \([-1, 1]\) at the points \(-1, 0, 1\). Hence, the left hand side of (3.5.29) only attains its maximum at points of the lattice \(-1, 0, 1\)^k_d.

Now we determine values of \( r_1, r_2, t_1 \) and \( t_2 \) so that the left hand side of (3.5.29) evaluated at the points \( u \in \{-1, 0, 1\}^k_d \) equals the right hand side of (3.5.29). Equality in (3.5.29) is achieved at the points \( u \in \{-1, 0, 1\}^k_d \) if we can find values of \( r_1, r_2, t_1 \) and \( t_2 \) so that

\[ a_4^{(1)} = -a_2^{(1)} , \hspace{1cm} (3.5.38) \]

\[ a_4^{(i_*)} = -a_2^{(i_*)} , \hspace{1cm} (3.5.39) \]

and

\[ a_0^{(1)} = a_0^{(i_*)} = (\lambda s_d^{-4})^{-1} \text{tr}[\mathbf{M}(\xi)^{-E^*}[\mathbf{H}^t \mathbf{A}_d^{-2} \mathbf{H}]]. \hspace{1cm} (3.5.40) \]
Using Corollary 3.5.5, we see the rightmost quantity in (3.5.40) is

\[
(\lambda s_d^{-4})^{-1} \text{tr}[M(\xi)^{-E^\nu} [H^t A_d^{-2}H]] = \frac{\theta}{4r_1} + \frac{(k_d - 1)}{4r_2} + \\
\frac{(\theta + 1)(k_d - 1)s_d^{-2}}{4t_1} + \frac{(k_d - 1)s_d^{-2}}{r_2 - t_2} \left( 1 - \frac{(r_1 - r_2^2)}{w} (t_2 - r_2^2) - (t_1 - r_1r_2^2) \right) + \\
\frac{(k_d - 1)(k_d - 2)s_d^{-2}}{4t_2} + \frac{\theta s_d^{-2}}{w} [r_2 - r_2^2 + (k_d - 2)(t_2 - r_2^2)].
\]  

(3.5.41)

Note that when (3.5.38), (3.5.39) and (3.5.40) are satisfied, (3.5.29) regarded as a function of \( u_1 \) is

\[
a_4^{(1)} u_1^2 (u_1^2 - 1) \leq 0,
\]  

(3.5.42)

and (3.5.29) as a function of \( u_{*} \) is

\[
a_4^{(\ast)} u_{*}^2 (u_{*}^2 - 1) \leq 0.
\]  

(3.5.43)

Before determining the parameters \( r_1, r_2, t_1, \) and \( t_2 \), recall the following constraints given in Appendix A.3

\[
0 < r_2^2 \leq t_2 < r_2 \leq 1,
\]  

(3.5.44)

and

\[
0 < r_1r_2 \leq t_1 \leq r_1r_2 + \sqrt{(r_1 - r_2^2)(r_2 - r_2^2 + (k_d - 2)(t_2 - r_2^2))} \frac{1}{k_d - 1} < \min\{r_1, r_2\} \leq 1.
\]  

(3.5.45)

Finding values of \( r_1, r_2, t_1 \) and \( t_2 \) that satisfy (3.5.38), (3.5.39) and (3.5.40) subject to (3.5.44) and (3.5.45) is not analytically possible. Therefore, we use a numerical search routine to find the values \( r_1, r_2, t_1 \) and \( t_2 \) that minimize the function

\[
f_\ast (r_1, r_2, t_1, t_2) = \max_{\alpha \in \{0, 1\}, \beta \in \{1, \ldots, k_d\}} \left\{ |a_0^{(\ast)} - (\lambda s_d^{-4})^{-1} \text{tr}[M(\xi)^{-E^\nu} [H^t A_d^{-2}H]]| + \right.
\]
\[ a_0^{(1)} - (\lambda s_d^{-4})^{-1} \text{tr}[M(\xi) - E^* [H^t A_d^{-2} H]] + \left| a_4^{(1)} + a_2^{(1)} \right| + \left| a_4^{(1)} + a_2^{(1)} \right| \] 

subject to (3.5.44) and (3.5.45). If \( f_*(r_1, r_2, t_1, t_2) = 0 \) at the obtained minimizing values of the parameters, then we have determined an IAV-optimal design. □

In the numerical examples that we have encountered, there have always been values of \( r_1, r_2, t_1 \) and \( t_2 \) that satisfy the constraints (3.5.44) and (3.5.45) and for which \( f_*(r_1, r_2, t_1, t_2) = 0 \). Moreover, although we can not prove a unique set exists, we have never seen a situation where multiple sets of \( r_1, r_2, t_1 \) and \( t_2 \) values satisfy the constraints (3.5.44) and (3.5.45) and have \( f_*(r_1, r_2, t_1, t_2) = 0 \). Table 6 gives the values of \( r_1, r_2, t_1, \) and \( t_2 \) for an optimal design for given values of \( k_d, s_d, \) and \( \theta \). Notice when \( \theta = 1 \) the values of \( r_1 = r_2 \) and \( t_1 = t_2 \). This is not surprising, since \( \theta = 1 \) makes the design problem invariant to all permutations of indices of \( u \).

According to Theorem 3.5.1, when selecting the levels of the standardized environmental variables \( v \) for an optimal exact design, we need only consider Hypothesis \( vi.) \). This hypothesis requires each column for a standardized environmental variable be orthogonal to all columns for at most third-order functions of the controllable variables. Therefore, we can once again obtain efficient exact designs of size \( N \) in two stages. First, we determine a collection of \( N \) levels of \( u \) that nearly satisfy the moment conditions in Hypotheses \( i.) \) through \( vi.) \) of Theorem 3.5.1. Then, we select \( k_e \) vectors from the orthogonal complement of the space generated by the columns for the functions of the controllable variables.. Our method for obtaining the \( k_e \) columns for the variables in \( v \) is given in Section 3.7.
3.6 IavD Optimal Designs

In this section we determine sufficient conditions for an experimental design, defined on \([-1,1]^{k_\alpha+k_\varepsilon}\), to minimize the determinant of the expected value of the asymptotic variance-covariance matrix (Design Criterion 1.2.6). The expectation of the asymptotic dispersion matrix in (3.4.18) is taken with respect to the prior distribution \(\pi_d^*\) for which the hypotheses in Assumptions 3.2.2 are satisfied. The goal of this section is to find a design measure \(\xi \in \Xi_1\) that minimizes

\[
\left| E^{\pi_d^*} \left[ A_d^{-1} H M(\xi)^{-1} H^t A_d^{-1} \right] \right|
\]  

over \(\Xi_1\). Given \(\pi_d^*, A_d^{-1}\) in (3.6.1) is fixed. Combined with the monotone increasing property of the log function, this means the design problem in (3.6.1) is equivalent to finding a design \(\xi \in \Xi_1\) that minimizes

\[
\log \left| E^{\pi_d^*} \left[ H M(\xi)^{-1} H^t \right] \right|.
\]  

over \(\Xi_1\). The design function in (3.6.2) is a scaled version of the convex and differentiable function of Example 1.3.2. Therefore, we can use the results of Theorem 1.3.4 to determine an optimal design.

As was the case for the IavA optimal design, we are not able to give the IavD optimal design in a closed form. Rather, we can only define the form of the information matrix up to the value of two numbers — \(r\) and \(t\). These numbers are the solutions to the set of equations given in (3.5.38), (3.5.39) and (3.5.40); hence, their values must
be obtained using numerical optimization methods. How these equations in \( r_1, r_2, t_1 \) and \( t_2 \) are used to determine the values of \( r \) and \( t \) is described in the proof.

**Theorem 3.6.1 (IavD Optimal Design)** Assume the model in (3.3.2), and let \( \pi^*_d \) be a prior distribution for which the hypotheses in Assumptions 3.2.2 are satisfied. There exists an IavD optimal design \( \xi^* \in \Xi_1 \), with respect to the prior \( \pi^*_d \), for which

\[
\begin{align*}
\text{i.) } & \mathbb{E}^* u^q_{i_1} u^q_{i_2} u^q_{i_3} u^q_{i_4} = 0 \quad \text{where each } i_j \in \{1, 2, \ldots, k_d\}, \text{ each } q_{i_j} \in \{0, 1, 2, 3\} \\
& \text{with } 0 \leq \sum_{j=1}^{4} q_{i_j} \leq 4, \text{ and at least one } q_{i_j} \text{ is odd,} \\
\text{ii.) } & \mathbb{E}^* u^q_i = \mathbb{E}^* u^q_i = r \quad \text{for each } i \in \{1, \ldots, k_d\}, \\
\text{iii.) } & \mathbb{E}^* u^q_i u^q_j = t \quad \text{for each } i, j \in \{1, \ldots, k_d\}, \text{ and} \\
\text{iv.) } & \mathbb{E}^* u^q_{i_1} u^q_{i_2} u^q_{i_3} v_k = 0 \quad \text{where each } i_j \in \{1, 2, \ldots, k_d\}, \text{ each } q_{i_j} \in \{0, 1, 2, 3\} \\
& \text{with } 0 \leq \sum_{j=1}^{4} q_{i_j} \leq 3, \text{ and } k \in \{1, 2, \ldots, k_e\}.
\end{align*}
\]

where \( r \) and \( t \) are the solutions to a complicated set of equations described in the proof.

The proof of the theorem requires knowing the values of certain matrices. Once again, we provide the values of the relevant matrices in sequence of lemmas and exile the calculations to Appendix B. In the lemmas the matrices are partitioned in a manner corresponding to (3.5.4).

The candidate design \( \xi^* \) in Theorem 3.6.1 is identical to the candidate design \( \xi^* \) in Theorem 3.5.1 when \( (r_1, t_1) = (r_2, t_2) \). This means the form of the information matrix and its Moore-Penrose generalized inverse can be obtained from Lemmas 3.5.3
and 3.5.4 by setting \( r_1 = r_2 = r \) and \( t_1 = t_2 = t \). Therefore, the first two lemmas of this section are stated without proof.

**Lemma 3.6.2** Let \( \xi^* \) be a design, defined on \([-1, 1]^{k_d + k_r} \), for which the four conditions in Theorem 3.6.1 are satisfied. The information matrix \( M(\xi^*) \) for the vector \( \alpha \) is

\[
M_d = \begin{pmatrix}
M_{d11} & M_{d12} \\
M_{d21} & M_{d22}
\end{pmatrix}
\]

with

\[
M_{d11} = \begin{pmatrix}
1 & 0_{k_d} \\
0_{k_d} & rI_{k_d}
\end{pmatrix},
\]

\[
M_{d12} = \begin{pmatrix}
rI_{k_d} \\
0_{k_d \times k_d}
\end{pmatrix}
\]

and

\[
M_{d22} = (r - t)I_{k_d} + tJ_{k_d}.
\]
Lemma 3.6.3 The Moore-Penrose generalized inverse for $M(\xi^*)$ in Lemma 3.6.2 is given by

$$M(\xi^*)^- = \begin{pmatrix}
M_d^{-1} & 0_{(2k_d+1)\times\left(\frac{k_d}{2}\right)} & 0_{(2k_d+1)\times(k_d+1)k_r}
0_{\left(\frac{k_d}{2}\right)\times(2k_d+1)} & t^{-1}I_{\left(\frac{k_d}{2}\right)} & 0_{\left(\frac{k_d}{2}\right)\times(k_d+1)k_r}
0_{(k_d+1)k_r\times(2k_d+1)} & 0_{(k_d+1)k_r\times\left(\frac{k_d}{2}\right)} & (\int f_3f_3'd\xi^*)^{-}
\end{pmatrix}$$

(3.6.8)

where $(\int f_3f_3'd\xi^*)^{-}$ is the Moore-Penrose generalized inverse for $\int f_3f_3'd\xi^*$, and $M_d^{-1}$ equals

$$\begin{pmatrix}
M_{d11}^{-1} & M_{d12}^{-1}
M_{d21}^{-1} & M_{d22}^{-1}
\end{pmatrix}$$

(3.6.9)

with

$$M_{d11}^{-1} = \begin{pmatrix}
r + (k_d - 1) t
r - r^2 + (k_d - 1)(t - r^2)
\end{pmatrix}
0_{k_d}
r^{-1}I_{k_d} \)

(3.6.10)

and

$$M_{d12}^{-1} = \begin{pmatrix}
r
r - r^2 + (k_d - 1)(t - r^2)
\end{pmatrix}
0_{k_d \times k_d} \)

(3.6.11)

and

$$M_{d22}^{-1} = \frac{1}{r - t} \left(I_{k_d} - \frac{t - r^2}{r - r^2 + (k_d - 1)(t - r^2)} J_{k_d}\right).$$

(3.6.12)

Lemma 3.6.4 Let $H$ be defined by (3.4.17), and let $\pi_d^*$ be a prior distribution for which the hypotheses of Assumptions 3.2.2 are satisfied. The matrix $E^{\alpha_d}[H'HH]$ is
equal to

\[
\begin{pmatrix}
0 & 0 & 0^{(\frac{k_d}{2})} & 0^{(k_d+1)k_e} \\
0^{k_d} & \frac{1}{4}I_{k_d} & 0_{k_d \times k_d} & 0_{k_d \times (k_d+1)k_e} \\
0^{k_d} & 0_{k_d \times k_d} & s_d^{-2}I_{k_d} & 0_{k_d \times (k_d+1)k_e} \\
0^{(\frac{k_d}{2})} & 0^{(\frac{k_d}{2})} \times k_d & 0^{(\frac{k_d}{2})} \times k_d & \frac{s_d}{2}I^{(\frac{k_d}{2})} \\
0^{(k_d+1)k_e} & 0_{(k_d+1)k_e \times k_d} & 0_{(k_d+1)k_e \times k_d} & 0_{(k_d+1)k_e \times (k_d+1)k_e}
\end{pmatrix}.
\] (3.6.13)

Lemma 3.6.5 The matrix \( E^d [HM^-(\xi^*)H^t] \) equals

\[
\left\{ \frac{1}{4r} + \frac{s_d^{-2}}{r-t} \left( \frac{r-r^2 + (k_d-2)(t-r^2)}{r-r^2 + (k_d-1)(t-r^2)} \right) \right\} I_{k_d}. \tag{3.6.14}
\]

Lemma 3.6.6 The matrix \( M^-(\xi^*)E^d [H^tH]M^-(\xi^*) \) equals

\[
\begin{pmatrix}
L_d & 0_{(2k_d+1) \times (\frac{k_d}{2})} & 0_{(2k_d+1) \times (k_d+1)k_e} \\
0^{(\frac{k_d}{2})} \times (2k_d+1) & \frac{s_d}{2t^2}I^{(\frac{k_d}{2})} & 0^{(\frac{k_d}{2})} \times (k_d+1)k_e \\
0_{(k_d+1)k_e \times (2k_d+1)} & 0_{(k_d+1)k_e \times (\frac{k_d}{2})} & 0_{(k_d+1)k_e \times (k_d+1)k_e}
\end{pmatrix}. \quad (3.6.15)
\]

where \( L_d \) is

\[
\begin{pmatrix}
L_{d11} & L_{d12} \\
L_{d21} & L_{d22}
\end{pmatrix} \tag{3.6.16}
\]

with

\[
L_{d11} = \begin{pmatrix}
0^{k_d} & 0^{(\frac{k_d}{2})} \\
\frac{1}{4t^2}I_{k_d}
\end{pmatrix}, \quad \frac{-r_s}{(r-r^2 + (k_d-1)(t-r^2))^2}I_{k_d} \tag{3.6.17}
\]

\[
L_{d12} = \begin{pmatrix}
\frac{r}{r-r^2 + (k_d-2)(t-r^2)} & 0^{(\frac{k_d}{2})} \\
0_{k_d \times k_d}
\end{pmatrix} \tag{3.6.18}
\]
and
\[ L_{d22} = \frac{s_d^{-2}}{(r-t)^2} \left( I_{kd} - \frac{t-r^2}{r-r^2+(k_d-1)(t-r^2)} J_{kd} \right)^2. \] (3.6.19)

**Proof of Theorem 3.6.1:** Let \( \mathbf{v} \) be the vector of regression functions of \((u,v)\) in (3.4.6), and consider the form of the design function in (3.5.2) compared with (1.3.10). Condition \( \text{ii.} \) of Theorem 1.3.4, says a design measure \( \xi \in \Xi_1 \) is IavD optimal with respect to the prior \( \pi_d^* \) if and only if
\[ f^t M(\xi)^{-E^*_d} \left[ H^t \left( E^*_d [H M(\xi)^{-H^t}]^{-1} H \right) \right] M(\xi)^{-f} \leq k_d \] (3.6.20)
for every \((u,v) \in [-1,1]^{k_d+k_r}\) (see Example 1.3.2). Therefore, to prove Theorem 3.6.1, we need to verify (3.6.20) for the candidate design \( \xi^* \).

Using the result of Lemma 3.6.5, we can rewrite (3.6.20) as
\[ f^t M^{-}(\xi^*)^{E^*_d[H^tH]M^{-}(\xi^*)}f \]
\[ \leq k_d \left\{ \frac{1}{4r} + \frac{s_d^{-2}}{r-t} \left( \frac{r-r^2+(k_d-2)(t-r^2)}{r-r^2+(k_d-1)(t-r^2)} \right) + \frac{s_d^{-2}(k_d-1)}{4t} \right\}. \] (3.6.21)

From Table 6, the design \( \xi^* \) in Theorem 3.6.1 is IavA optimal when \( \theta = 1 \) and \((r_1,t_1) = (r_2,t_2)\). That is, (3.5.29) is satisfied for the design \( \xi^* \) when \( \theta = 1 \). Hence, it suffices to show that (3.6.21) is equivalent to (3.5.29) for the design \( \xi^* \) when \( \theta = 1 \) and \((r_1,t_1) = (r_2,t_2)\) to prove the theorem.

First, under the hypothesis that \( \theta = 1 \), we show the equality of the right hand side of (3.6.21) and \((\lambda s_d^{-4})^{-1} \text{tr}[M(\xi)^{-E^*}[H^tA_d^{-2}H]]\) in (3.5.41) for the design \( \xi^* \). Substituting the values \( r_1 = r_2 = r, t_1 = t_2 = t \) and \( \theta = 1 \) into (3.5.41), we obtain that
\[ w = (r-t) \left( r-r^2-(k_d-1)(t-r^2) \right); \] (3.6.22)
and \((\lambda s_d^{-4})^{-1} \text{tr}[M(\xi)^{-1}E^\pi'[H'A_d^{-2}H]]\) is
\[
\frac{k_d}{4r} + \frac{k_ds_d^{-2}}{r-t} \left( \frac{r-r^2 + (k_d-2)(t-r^2)}{r-r^2 + (k_d-1)(t-r^2)} \right) + \frac{s_d^{-2}}{2t} \left( \frac{k_d}{2} \right). \tag{3.6.23}
\]
Factoring out the common value \(k_d\), we see (3.6.23) equals the right hand side of (3.6.21).

Next, we show the equality of the matrix \(E^\pi_d[H'H]\) on the left hand side of (3.6.21) and \((\lambda s_d^{-4})^{-1}E^\pi'[H'A_d^{-2}H]\) on the left hand side of (3.5.29) for the design \(\xi^*\) when \(\theta = 1\). Using Lemmas 3.5.2 and 3.6.4 and the value \(\theta = 1\), we get
\[
(\lambda s_d^{-4})^{-1} \frac{s_d^{-4}}{4} \Lambda_Q = \frac{1}{4} I_{k_d}, \tag{3.6.24}
\]
\[
(\lambda s_d^{-4})^{-1} \frac{s_d^{-6}}{4} \Lambda_Q = s_d^{-2} I_{k_d} \tag{3.6.25}
\]
and
\[
(\lambda s_d^{-4})^{-1} \frac{s_d^{-6}}{4} \Lambda_{QI} = \frac{s_d^{-2}}{2} I(t_d^2) \tag{3.6.26}
\]
where \(\Lambda_Q\) and \(\Lambda_{QI}\) are defined in (3.2.3) and (A.1.17), respectively. The relations in (3.6.24), (3.6.25) and (3.6.26) establish the required equality of the left hand sides of (3.6.21) and (3.5.29). Therefore, the theorem is proved. \(\square\)

Under the hypothesis \(\theta = 1\), we have shown that the IavA and IavD optimal designs are the same. This means the values of \(r_1, r_2, t_1\) and \(t_2\) for \(\theta = 1\) in Table 6 are the proper design moments for an IavD optimal design. As a result, construction of optimal designs in the section can be handled in a unified manner.
3.7 Constructing Efficient Designs for the IavA and IavD Criteria

In this section we describe a method for constructing the optimal designs in Theorem 3.5.1 and Theorem 3.6.1. The properties of the optimal designs in both theorems allow a two-stage construction. First, determine a collection of $N$ levels for the standardized control variables, $\{u^{(i)}\}_{i=1}^N$, that approximate well the moment values $r_1$, $r_2$, $t_1$ and $t_2$ while maintaining marginal symmetry, i.e., for each $j \in \{1, \ldots, k_d\}$ the $N$ levels of $u_j$ are such that $\sum_{i=1}^N u_j^{(i)} = \sum_{i=1}^N (u_j^{(i)})^3 = 0$. Second, find a set of $k_e$ vectors for the standardized environmental variables $v$ that satisfy the required orthogonalities between $u$ and $v$.

In the second stage, we must find $k_e$ $N$-dimensional vectors that are orthogonal to the $\binom{k_d+3}{3} - k_d$ dimensional space spanned by the vectors of all control variable functions that are at most third-order. This means $N$ must be larger than $\binom{k_d+3}{3} - k_d + k_e$. From this orthogonal complement space, we want to select the set of $k_e$ vectors for which the $N$ levels of $v$ (resp. $x_e$) are most proportionately distributed across the space $[-1, 1]^{k_e}$ (resp. $\mathcal{X}_e$). Our method for determining the $k_e$ vectors for the standardized environmental variables is described in Section 2.8.

To determine the $N$ vectors $\{u^{(i)}\}_{i=1}^N$, recall that the support space for the optimal design in both theorems has the levels of $u$ restricted to the lattice $\{-1, 0, 1\}^{k_d}$. Even for moderate $k_d$ the number of points in this lattice can be prohibitively large. Therefore, we shall find subsets of points from the lattice that can support an optimal design.
Definition 3.7.1 For \( j \in \{0, \ldots, k\} \), the \( j^{th} \) isozero set \( \mathcal{H}^j \) in the lattice \( \{-1, 0, 1\}^k \) is the set of points that have exactly \( j \) coordinates equal to zero.

This partition of the lattice \( \{-1, 0, 1\}^k \) consists of \( k_d+1 \) sets where the \( j^{th} \) set \( \mathcal{H}^j \) has cardinality

\[
|\mathcal{H}^j| = 2^{k_d-j} \binom{k_d}{j}, \quad j \in \{0, 1, \ldots, k_d\}.
\] (3.7.1)

The main reason for partitioning the lattice into the isozero sets is that marginal symmetry is maintained on each isozero set, i.e., for each isozero set \( \mathcal{H}^j \)

\[
\sum_{\mathbf{u} \in \mathcal{H}^j} u_i = \sum_{\mathbf{u} \in \mathcal{H}^j} u_i^2 = 0 \text{ for each } i \in \{1, \ldots, k_d\}.
\] (3.7.2)

This means we can construct designs by assigning design weights to the isozero sets with equal weight given to each point in an isozero set. If we determine \( k_d+1 \) weights so that the design moments \( r_1, r_2, t_1 \) and \( t_2 \) are achieved, then we have constructed an optimal design. Of course, to have obtained an optimal design with a reduced support space, some of the \( k_d+1 \) design weights must equal zero.

Consider an integrated asymptotic variance D-optimal design (equiv. IavA optimal design with \( \theta = 1 \)). Here \( r_1 = r_2 = r \) and \( t_1 = t_2 = t \), so an optimal design must be constructed on at least three isozero sets. Let \( \mathcal{D}(i, j, k) \) be a design on the three isozero sets \( \mathcal{H}^i, \mathcal{H}^j, \) and \( \mathcal{H}^k \). From (3.7.1) we see that the smallest isozero sets are \( \mathcal{H}^{k_d} \) (the origin), \( \mathcal{H}^{k_d-1} \) (the \( 2k_d \) midfaces), and \( \mathcal{H}^0 \) (the \( 2^{k_d} \) vertices of the cube \([-1,1]^{k_d}\)), so this triple would be the preferred candidate to support the design. However, we shall show that not every one of the \( \binom{k_d+1}{3} \) combinations of three isozero sets supports an optimal design. In particular, we argue that no central composite
design $D(0, k_d - 1, k_d)$ can be an IavD optimal (resp. IavA optimal) design unless $k_d = 2$.

**Conjecture 3.7.1** When $k_d > 2$, there do not exist design weights $\omega_0, \omega_{k_d-1}, \omega_{k_d} \in (0, 1)$ with $\omega_0 + \omega_{k_d-1} + \omega_{k_d} = 1$ for the three iszero sets $\mathcal{H}^0, \mathcal{H}^{k_d-1}, \text{and } \mathcal{H}^{k_d}$ such that the design $D(0, k_d - 1, k_d; \omega_0, \omega_{k_d-1}, \omega_{k_d})$ satisfies the conditions of Theorem 3.6.1 (resp. Theorem 3.5.1).

Proving this conjecture requires having analytic expressions for the parameters $r$ and $t$. Since these expressions have not been obtained, this conjecture has not been analytically proven. However, the result has been verified numerically by the author for a variety of $k_d$ and $s_d$. In the proceeding paragraphs, we present an argument for the conjecture. In the process, we illustrate how optimal design "weights" are determined for an arbitrary three iszero set design $D(i, j, k)$. We have qualified the word weight here, because for given $i < j < k \in \{0, \ldots, k_d - 1\}$ the values of $\omega_i, \omega_j,$ and $\omega_k$ for which the design $D(i, j, k)$ matches the moments $r$ and $t$ may not be valid design weights.

To determine the design "weights" so that the design $D(i, j, k; \omega_i, \omega_j, \omega_k)$, $i < j < k \in \{0,\ldots,k_d-1\}$, matches the moments $r$ and $t$, consider the bivariate distribution of $u_m$ and $u_n$, $m, n \in \{1,\ldots,k_d\}$, in Figure 4. In this figure, we have folded the symmetric square $[-1,1]^2$ along each of the coordinate axes producing the unit square in the first quadrant. Folding along the axes places all support points of the design into four collections of points for which the amount of bivariate design mass is known.
Figure 4: Double-folding a bivariate slice of the design region \([-1, 1]^k\) to obtain the isozero set weights for the design \(D(i, j, k)\). The coordinate pairs inside the first quadrant denote the location of the points after the double fold. The values with arrows designate the proportion of the bivariate design mass on the respective points.

Using Figure 4, we can show that

\[
  r = \frac{k_d - i}{k_d} \omega_i + \frac{k_d - j}{k_d} \omega_j + \frac{k_d - k}{k_d} (1 - \omega_i - \omega_j) \tag{3.7.3}
\]

and

\[
  t = \frac{(k_d - i)(k_d - i - 1)}{k_d(k_d - 1)} \omega_i + \frac{(k_d - j)(k_d - j - 1)}{k_d(k_d - 1)} \omega_j + \frac{(k_d - k)(k_d - k - 1)}{k_d(k_d - 1)} (1 - \omega_i - \omega_j). \tag{3.7.4}
\]
To determine the "weights" for an IavD optimal design (equiv. IavA optimal design with $\theta = 1$), we simply plug the values of $r_1 = r_2 = r$ and $t_1 = t_2 = t$ from Table 6 into (3.7.3) and (3.7.4), and solve the linear system above.

For designs of the form $\mathcal{D}(0, j, k_d)$ with $j \in \{1, \ldots, k_d - 1\}$, we have

$$\omega_j = \left( \frac{k_d(k_d - 1)}{j(k_d - j)} \right) (r - t) \quad (3.7.5)$$

and

$$\omega_0 = r - \left( \frac{k_d - 1}{j} \right) (r - t). \quad (3.7.6)$$

Clearly, $\omega_j > 0$, $\omega_0 < 1$, and $\omega_0 + \omega_j > 0$. If $\omega_0 \geq 0$, and $\omega_0 + \omega_j \leq 1$ for some value $j \in \{1, \ldots, k_d - 1\}$, then $\mathcal{D}(0, j, k_d)$ is an optimal design for that value of $j$. The statement $\omega_0 \geq 0$ implies

$$j \geq (k_d - 1) \left( \frac{r - t}{r} \right) \quad (3.7.7)$$

while $\omega_0 + \omega_j \leq 1$ implies

$$j \leq k_d - (k_d - 1) \left( \frac{r - t}{1 - r} \right). \quad (3.7.8)$$

Using (3.7.7) and (3.7.8), it is easy to show that

$$k_d - (k_d - 1) \left( \frac{r - t}{1 - r} \right) \geq (k_d - 1) \left( \frac{r - t}{r} \right) + 1. \quad (3.7.9)$$

This implies there exists a $j \in \{1, \ldots, k_d - 1\}$ that satisfies (3.7.7) and (3.7.8), so $\mathcal{D}(0, j, k_d)$ is an optimal design for at least one $j \in \{1, \ldots, k_d - 1\}$. However, using the values of $r$ and $t$ ($\theta = 1$) from Table 6, we see that

$$k_d - (k_d - 1) \left( \frac{r - t}{1 - r} \right) < k_d - 1 \quad (3.7.10)$$
Table 4: All designs $D(i, j, k)$ supported on three isozero sets that are integrated asymptotic variance D-optimal designs for estimating $u^*$ (resp. $x^*_d$) when $k_d = 4$ and $s_d = 1$.

<table>
<thead>
<tr>
<th>Three Isozero set</th>
<th>Design</th>
<th>Cardinality of Support Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$D(0, 1, 2; .247, .388, .365)$</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>$D(0, 1, 3; .125, .612, .263)$</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>$D(0, 2, 3; .329, .612, .059)$</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>$D(0, 2, 4; .321, .657, .022)$</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>$D(1, 2, 4; .857, .014, .129)$</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>$D(1, 3, 4; .861, .014, .125)$</td>
<td>41</td>
</tr>
</tbody>
</table>

unless $k_d = 2$. This means for these values of $r$ and $t$ the value of $j$ is less than $k_d - 1$ for $k_d \geq 3$, so no central composite design $D(0, k_d - 1, k_d)$ is optimal if $k_d > 2$.

Table 4 contains all possible IavD optimal designs (equiv. IavA optimal designs for $\theta = 1$) constructed on three isozero sets for $k_d = 4$ and $s_d = 1$. In this case, $D(0, 2, 4)$ is the only optimal design of the form $D(0, j, k_d)$. This design has 2.2% of the design mass at the origin, i.e. $H^4$, 2.7% of the mass on each of the 24 points of $H^2$, and 2.0% of its mass on each of the 16 vertices in $H^0$. This design is one of the optimal designs with the smallest number of support points.

In some situations, it is possible to reduce the number of support points to a value less than the totality of the isozero sets with positive support. For $j \in \{0, 1, \ldots, k_d\}$, we can divide the isozero set $H^j$ into $\binom{k_d}{j}$ blocks where each block contains a $2^{k_d-j}$ factorial arrangement that is augmented with $j$ columns of zeroes. If $k_d - j$ is large
enough, we can take only a $2^{-f}$ fraction of the $2^{k_d-j}$ factorial in each block without confounding the coefficients in (2.3.1). This reduces the support set for the optimal design by

$$
2^{k_d-j} \binom{k_d}{j} \left( \frac{2^f - 1}{2^f} \right)
$$

(points. The model in (2.3.1) contains a full second-order model in the control variables, so two-way interactions must be aliased with three-way or higher interactions, i.e., the design must be at least Resolution V. This means we can construct an optimal design using a fraction of the isozero set $\mathcal{H}^j$ only if $k_d - j \geq 5$. Examples of aliasing schemes for Resolution V design are given in Box, Hunter and Hunter (1978).

IavA optimal designs for $\theta \neq 1$ have four parameters $r_1 \neq r_2$ and $t_1 \neq t_2$, so construction of an optimal design requires having at least five different subsets of points in the lattice with positive design mass. We could use five isozero sets for the subsets, but this makes the experimental design prohibitively large. Instead, we split two of isozero sets from the design $\mathcal{D}(i,j,k)$ into four smaller sets. Since neither $\mathcal{H}^0$ nor $\mathcal{H}^{k_d}$ can be split in this manner, we do not consider designs of the form $\mathcal{D}(0,j,k_d)$.

**Definition 3.7.2** For $j \in \{1, \ldots, k_d - 1\}$, the subisozero set $\mathcal{H}^{(0)}_j$ is the subset of points of $\mathcal{H}^j$ whose first coordinate is 0, and the subisozero set $\mathcal{H}^{(1)}_j$ is the subset of points of $\mathcal{H}^j$ whose first coordinate is $\pm 1$.

Splitting an isozero set in this fashion results in two subisozero sets that retain the marginal symmetry property of the original isozero set. The cardinalities of the
two subisozero sets $H^{(0)}$ and $H^{(1)}$ are

$$|H^{(0)}| = 2^{k_d - j} \binom{k_d - 1}{j - 1}$$  \hspace{1cm} (3.7.12)

and

$$|H^{(1)}| = 2^{k_d - j} \binom{k_d - 1}{j}$$  \hspace{1cm} (3.7.13)

Note that we can only fractionate the subisozero sets of $H^j$ if $k_d - j - 1 \geq 5$.

Figure 5: Double-folding two bivariate slices of the design region $[-1, 1]^d$ to obtain the isozero set weights for the designs $D(0, j_1^{(0)}, j_1^{(1)}, j_2^{(0)}, j_2^{(1)})$ and $D(j_1^{(0)}, j_1^{(1)}, j_2^{(0)}, j_2^{(1)}, k_d)$. The coordinate pairs inside the first quadrants denote the location of the points after the double fold. The values with arrows designate the proportion of the bivariate design mass on the respective points.

To construct IavA optimal designs, we consider only designs of the form $D(0, j_1^{(0)}, j_1^{(1)}, j_2^{(0)}, j_2^{(1)})$ and $D(j_1^{(0)}, j_1^{(1)}, j_2^{(0)}, j_2^{(1)}, k_d)$. To determine the design “weights” that make these designs optimal, consider the bivariate distributions in Figure 5 for
\( (u_1, u_n) \) and \( (u_m, u_n) \) with \( m, n \in \{2, \ldots, k_d\} \). For the design \( \mathcal{D}(0,j_1(0), j_1(1), j_2(0), j_2(1)) \) we derive the following system of equations

\[
t_1 = \omega_0 + \frac{k_d - 1 - j_1}{k_d - 1} \omega_{j_1}^{(1)} + \frac{k_d - 1 - j_2}{k_d - 1} \omega_{j_2}^{(1)}, \tag{3.7.14}
\]

\[
r_1 - t_1 = \frac{j_1}{k_d - 1} \omega_{j_1}^{(1)} + \frac{j_2}{k_d - 1} \omega_{j_2}^{(1)}, \tag{3.7.15}
\]

\[
r_2 - t_1 = \frac{k_d - j_1}{k_d - 1} \omega_{j_1}^{(0)} + \frac{k_d - j_2}{k_d - 1} \omega_{j_2}^{(0)}, \tag{3.7.16}
\]

and

\[
r_2 - t_2 = \frac{(k_d - j_1)(j_1 - 1)}{(k_d - 1)(k_d - 2)} \omega_{j_1}^{(0)} + \frac{j_1(k_d - j_1 - 1)}{(k_d - 1)(k_d - 2)} \omega_{j_1}^{(1)} + \frac{j_1(k_d - j_1 - 1)}{(k_d - 1)(k_d - 2)} \omega_{j_1}^{(0)} + \frac{j_2(k_d - j_2 - 1)}{(k_d - 1)(k_d - 2)} \omega_{j_2}^{(1)} - \frac{j_2(k_d - j_2 - 1)}{(k_d - 1)(k_d - 2)} \omega_{j_2}^{(0)}. \tag{3.7.17}
\]

Similarly, for the design \( \mathcal{D}(j_1(0), j_1(1), j_2(0), j_2(1), k_d) \) the linear system for the optimal design weights is

\[
t_1 = \frac{k_d - 1 - j_1}{k_d - 1} \omega_{j_1}^{(1)} + \frac{k_d - 1 - j_2}{k_d - 1} \omega_{j_2}^{(1)}, \tag{3.7.18}
\]

\[
r_1 = \omega_{j_1}^{(1)} + \omega_{j_2}^{(1)}, \tag{3.7.19}
\]

\[
r_2 = \frac{k_d - j_1}{k_d - 1} (\omega_{j_1}^{(0)} + \omega_{j_1}^{(1)}) + \frac{k_d - j_2}{k_d - 1} (\omega_{j_2}^{(0)} + \omega_{j_2}^{(1)}) + \frac{\omega_{j_1}^{(1)} + \omega_{j_2}^{(1)}}{k_d - 1}, \tag{3.7.20}
\]

and

\[
t_2 = \frac{(k_d - j_1)(k_d - j_1 - 1)}{(k_d - 1)(k_d - 2)} \omega_{j_1}^{(0)} + \frac{(k_d - j_1 - 1)(k_d - j_1 - 2)}{(k_d - 1)(k_d - 2)} \omega_{j_1}^{(1)} + \frac{(k_d - j_2)(k_d - j_2 - 1)}{(k_d - 1)(k_d - 2)} \omega_{j_2}^{(0)} + \frac{(k_d - j_2 - 1)(k_d - j_2 - 2)}{(k_d - 1)(k_d - 2)} \omega_{j_2}^{(1)}. \tag{3.7.21}
\]

Using the values of \( r_1, r_2, t_1 \) and \( t_2 \) from Table 6, we can solve the linear system for both designs to determine if an optimal design can be supported on either of the five
subsets. In examples that we have tried there has always been an optimal design of this form.

An exact design $\xi_N$ that is constructed by the method described in this section has an information matrix $M(\xi_N)$ of the form given by Lemma 3.5.3. However, in most cases the values of the moment parameters $r_1, r_2, t_1$ and $t_2$ will not be achieved by the exact design. To determine the efficiency of an exact design for the IavA criterion, we use (3.5.41) to calculate the ratio

$$\frac{\text{tr} \left[ M(\xi^*) - E^{\xi^*} [H'A_d^2H] \right]}{\text{tr} \left[ M(\xi_N) - E^{\xi_N} [H'A_d^2H] \right]}.$$  

(3.7.22)

The ratio in (3.7.22) is less than or equal to 1 for every exact design, and we want to select an exact design for which the ratio is nearest to 1. Similarly, for the IavD criterion, we use Lemma 3.6.5 to determine the efficiency of the exact design $\xi_N$, i.e.,

$$\frac{E^{\xi^*} [HM(\xi^*) - H']}{E^{\xi_N} [HM(\xi_N) - H']}.$$  

(3.7.23)

Computer code that constructs designs on $[-1, 1]^{k_d+k_e}$ for Theorem 3.5.1 and Theorem 3.6.1 is available from the author. The program first determines the values of $r_1, r_2, t_1$ and $t_2$ given user-supplied values of $k_d, k_e, s_d$, and $\theta$. Once the values of the design moments are determined, the program suggests a near-optimal exact design of the form $D(0, j_1^{(0)}, j_1^{(1)}, j_2^{(0)}, j_2^{(1)})$ or $D(j_1^{(0)}, j_1^{(1)}, j_2^{(0)}, j_2^{(1)}, k_d)$. Along with the suggested design are the values of the design's IavA and, if appropriate, IavD efficiencies. At this point the user can change the number of design points, alter the fractionating scheme (if appropriate), change the number of replications on each isozero set (sub-isozero set), or change the isozero set triple ($\{0, j, k_d\}$ is not allowed). After each change, the
requested design is constructed and efficiencies calculated. Design examples that are constructed using the code are given in Section 3.8.

3.8 Examples of IavA and IavD Optimal Designs

In this section we construct efficient exact designs for estimating the optimal operating conditions $x_d^*$ in three separate quality improvement computer experiments. The designs in the first two examples are used in the simulation experiments of Chapter IV. The third example is presented to illustrate different features of the author's experimental design program. For each of these examples, we assume the vector of environmental variables $x_e$ is distributed according to the known distribution $\pi_e$ for which

$$\mu_e = E^{\pi_e}[x_e] = 0, \quad \Sigma_e = E^{\pi_e}[x_e][x_e]^t = I_{k_e}. \quad (3.8.1)$$

In addition, we assume $x_e = [-1,1]^{k_e}$; thus, the sizing scalar $s_e = 1$. This means no scaling or transforming of the levels of the environmental variables in the design constructed by the author's program is required to obtain the optimal design on $X$.

Example 3.8.1 (IavAD Optimal Design for $(k_d, k_e) = (2, 4)$) Suppose a computer code depends on six inputs: two that can be controlled by the manufacturer and four that are determined by the distribution $\pi_e$. In addition, assume the scalar output $y(x_d, x_e)$ is a realization of the stochastic process $Y(x_d, x_e)$ in (1.2.3), so the
optimal operating conditions are $\mathbf{x}^*_d$ given by (3.1.1). In this example we construct an efficient 25 point design for estimating $\mathbf{x}^*_d$ using the IavD design criterion (resp. IavA design criterion with $\theta = 1$). Since we give joint consideration to both criteria, we shall refer to this design as the IavAD design.

Let $\pi_d^*$ be a prior for which the hypotheses in Assumptions 3.2.2 are given by

$$\mu_d^* = E^n_d[\mathbf{x}^*_d] = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and

$$\Sigma_d = E^n_d[(\mathbf{x}^*_d)(\mathbf{x}^*_d)^t] = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$  

(3.8.3) (3.8.4)

In addition, assume the design space for the controllable variables is $\mathcal{X}_d = [-1, 1]^2$. Hence, $s_d = 1$.

Using the above information, we can use the author's experimental design program to construct the design. The session of the program that constructed the efficient 25 point IavAD design for this example is given below. Since $k_d = 2$, there are only three design parameters: $r_1$, $r_2$ and $t_1$, and whereas $\theta = 1$, $r_1 = r_2$. For the values of these parameters given by the program, the function $f_*(r_1, r_2, t_1, t_2)$ in (3.5.46) equals zero. These values were determined using a public domain computer code implementation of the global search method simulated annealing.

Enter the number of controllable variates -- Kd 2
Enter the number of environmental variates -- Ke 4
Enter the scaling factor of Xd -- Sd 1
Enter the ratio -- theta = gamma/lambda
For the D-criterion the ratio = 1.0

Please wait a couple minutes.

Optimal values for r1, r2, and t1

0.58743 0.58743 0.35881

Here is a sample design
A 10 point optimal design requires
1 reps of a 1/1 fraction of the 0th isozero set (4 points)
1 reps of a 1/1 fraction of the 1th (0) isozero set (2 points)
1 reps of a 1/1 fraction of the 1th (1) isozero set (2 points)
2 reps of a 1/1 fraction of the 2th isozero set (1 points)

IavA criterion efficiency is 0.98769
IavD criterion efficiency is 0.97553

Enter the desired number of points in the design.
25

A 25 point optimal design requires
2 reps of a 1/1 fraction of the 0th isozero set (4 points)
3 reps of a 1/1 fraction of the 1th (0) isozero set (2 points)
3 reps of a 1/1 fraction of the 1th (1) isozero set (2 points)
5 reps of a 1/1 fraction of the 2th isozero set (1 points)

IavA criterion efficiency is 0.99488
IavD criterion efficiency is 0.98979

Is this design okay? (y or n)
y

Divide the noise margins into how many regions
5

The multivariate entropy efficiency is
1.00000
The univariate entropy efficiencies are
1.000 1.000 1.000 1.000
The product correlations between columns are
0.000 0.000 0.000
0.000
0.000
Are these efficiencies okay? (y or n)
y
filename for the output: TEMP puts output to the screen

Table 5: Reproduction of the 10 point sample design in Example 3.8.1

<table>
<thead>
<tr>
<th>Replications</th>
<th>Fraction $2^f$</th>
<th>Isozero Set $\mathcal{H}_j$</th>
<th>Cardinality of $\mathcal{H}_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$2^0 = 1$</td>
<td>$\mathcal{H}_0$</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>$2^0 = 1$</td>
<td>$\mathcal{H}_1^{(0)}$</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>$2^0 = 1$</td>
<td>$\mathcal{H}_1^{(1)}$</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>$2^0 = 1$</td>
<td>$\mathcal{H}_2$</td>
<td>1</td>
</tr>
</tbody>
</table>

After determining the values of $r_1$, $r_2$ and $t_1$, the program constructs a sample design of sufficient size to be efficient, i.e., $(\geq 99\%)$, with respect to the IavA criterion. We reproduced the 10 point sample design in Table 5 using the notation of this dissertation to clarify the output. We can manipulate the replication of the iszero sets, but the 25 point design given by the program is already highly efficient for both criteria. For each noise variable, we partitioned the margin $[-1,1]$ into five equal subintervals. This subdivides the space $\mathcal{X}_r = [-1,1]^4$ into $5^4 = 625$ equal cuboidal regions. The multivariate entropy efficiency of 1.000 means each cuboidal region has at most one design point in it, and each marginal entropy efficiency equal to 1.000 means each of the five subintervals in the margin for each of the four variables has five design points in it. The product correlations provide a check that the four columns for the environmental variables are mutually orthogonal. The design is given in Table 7. □
Example 3.8.2 (IavAD Optimal Design for \((k_d, k_e) = (4, 5)\)) Suppose a computer code depends on nine inputs: four variables that can be controlled by the manufacturer and five environmental variables determined by the distribution \(\pi_e\). In addition, assume the scalar output \(y(x_d, x_e)\) is a realization of the stochastic process \(Y(x_d, x_e)\) in (1.2.3), so the optimal operating conditions are \(x_d^*\) given by (3.1.1). In this example we construct an efficient 64 point design for estimating \(x_d^*\) using the IavD design criterion (resp. IavA design criterion with \(\theta = 1\)).

We assume \(\pi_d^*\) is a prior for which the hypotheses in Assumptions 3.2.2 are given by

\[
\mu_d^* = \mathbb{E}^{\pi_d^*}[x_d^*] = 0_d, \tag{3.8.5}
\]

and

\[
\Sigma_d = \mathbb{E}^{\pi_d^*}[(x_d^*) (x_d^*)^t] = I_d. \tag{3.8.6}
\]

In addition, we assume the design space for the controllable variables is \(X_d = [-1, 1]^d\). This implies \(s_d = 1\).

The following contains portions of the program session that constructed the efficient 64 point design on \([-1, 1]^9\) for the IavAD design criterion in this example. Since \(\theta = 1\), there are only two design parameters: \(r_1 = r_2 = r\) and \(t_1 = t_2 = t\).

Enter the number of controllable variates -- Kd 4
Enter the number of environmental variates -- Ke 5
Enter the scaling factor of Xd -- Sd 1
Enter the ratio -- theta = gamma/lambda
For the D-criterion the ratio = 1.0 1
Please wait a couple minutes.

Optimal values for r1, r2, t1, and t2
.64957 .64957 .43074 .43074

Here is a sample design
A 72 point optimal design requires
1 reps of a 1/1 fraction of the 0th isozero set (16 points)
2 reps of a 1/1 fraction of the 2th (0) isozero set (12 points)
2 reps of a 1/1 fraction of the 2th (1) isozero set (12 points)
1 reps of a 1/1 fraction of the 3th (0) isozero set (6 points)
1 reps of a 1/1 fraction of the 3th (1) isozero set (2 points)

IavA criterion efficiency is .96090
IavD criterion efficiency is .85253

Enter the desired number of points in the design.
64

A 72 point optimal design requires
1 reps of a 1/1 fraction of the 0th isozero set (16 points)
2 reps of a 1/1 fraction of the 2th (0) isozero set (12 points)
2 reps of a 1/1 fraction of the 2th (1) isozero set (12 points)
1 reps of a 1/1 fraction of the 3th (0) isozero set (6 points)
1 reps of a 1/1 fraction of the 3th (1) isozero set (2 points)

IavA criterion efficiency is .96090
IavD criterion efficiency is .85253

Is this design okay? (y or n)
n
Change the isozero sets, fractions or replications (i, f or r)
r
Enter the number of replications for the 5 (or 4) sets
2 1 1 1 1

A 64 point optimal design requires
2 reps of a 1/1 fraction of the 0th isozero set (16 points)
1 reps of a 1/1 fraction of the 2th (0) isozero set (12 points)
1 reps of a 1/1 fraction of the 2th (1) isozero set (12 points)
1 reps of a 1/1 fraction of the 3th (0) isozero set ( 6 points)
1 reps of a 1/1 fraction of the 3th (1) isozero set ( 2 points)

IavA criterion efficiency is .90588
IavD criterion efficiency is .67341

Is this design okay? (y or n)

n
Change the isozero sets, fractions or replications (i, f or r)
i
Enter the three isozero sets -- Must include either 0 or Kd
0 1 3

A 64 point optimal design requires
1 reps of a 1/1 fraction of the 0th isozero set ( 16 points)
1 reps of a 1/1 fraction of the 1th (0) isozero set ( 8 points)
1 reps of a 1/1 fraction of the 1th (1) isozero set ( 24 points)
2 reps of a 1/1 fraction of the 3th (0) isozero set ( 6 points)
2 reps of a 1/1 fraction of the 3th (1) isozero set ( 2 points)

IavA criterion efficiency is .97492
IavD criterion efficiency is .90337

Is this design okay? (y or n)

y

Divide the noise margins into how many regions

5

The multivariate entropy efficiency is
1.00000
The univariate entropy efficiencies are
.997 .996 .993 .994 .999
The product correlations between columns are
.000 .000 .000 .000
.000 .000 .000
.000 .000
.000

Are these efficiencies okay? (y or n)
y
filename for the output : TEMP puts output to the screen
Although the program finds the optimal values for $r_1$, $r_2$, $t_1$, and $t_2$, it does not determine the best exact design for a given $N$. Instead, it allows the experimenter to alter the triple of isozero sets and their associated replications to search for an acceptable design. In this example, the fractionating scheme could not be modified, since $k_d < 5$.

The suggested design is obtained by considering successively larger sample sizes $N$ and searching through all possible triples of isozero sets, that support an optimal design and for which a design of size $N$ exists. When a triple is found that supports an $N$ point highly efficient design, then resulting design becomes the starting sample design in the program.

To determine the levels of the noise variables, we partitioned each margin $[-1, 1]$ into five equal subintervals subdividing the space $\mathcal{X}_c = [-1, 1]^4$ into $5^4 = 3125$ equal cuboidal regions. It should be no surprise that each cuboidal region has at most one design point in it making the multivariate entropy efficiency equal 1.000. However, notice that each marginal entropy efficiency is less than 1.000 which means no environmental variable has its levels equally divided among the five equal intervals of $[-1, 1]$. This can be attributed to the random manner with which environmental variables are chosen. As $N$ increases, it becomes more unlikely that a randomly generated $N$-dimensional vector will have coordinates evenly spread across the margin. The 64 point exact design constructed in this example is given in Table 8. □
Example 3.8.3 (IavA Optimal Design for \((k_d, k_c) = (7, 5)\)) This example considers a quality improvement experiment with enough controllable variables to fractionate an isozero set. Suppose a computer code depends on twelve inputs: seven that are controllable and five that are determined by the distribution \(\pi_c\). In addition, assume the scalar output \(y(x_d, x_c)\) is a realization of the stochastic process \(Y(x_d, x_c)\) in (1.2.3), so the optimal operating conditions are \(x_d^*\) given by (3.1.1). The number of parameters in (1.2.3) that must estimated is 76. In this example we construct an efficient 162 point design for estimating \(x_d^*\) using the IavA design criterion with \(\theta = 0.5\).

Let \(\pi^*\) be a prior for which the hypotheses in Assumptions 3.2.1 are given by

\[
\mu^*_d = E^{x_d^*}[x_d^*] = 0_7, \tag{3.8.7}
\]

\[
\Sigma_d = E^{x_d^*}[(x_d^*)^T(x_d^*)^T] = (1.5)^{-2}I_7, \tag{3.8.8}
\]

the distribution of \(x_d^*\) under \(x_d^*\) is exchangeable, and the completely symmetric positive definite matrix \(Q\) has eigenvalues \(\lambda > 0\) with multiplicity six and \(\lambda/2\) with multiplicity one. In addition, we assume the design region \(X_d = \{1.5R'u : u \in [-1, 1]^7\}\) where \(R\) is the matrix of eigenvectors for the set of real \(7 \times 7\) completely symmetric matrices (see Section 3.2.1). Note that \(s_d = 1.5\).

The following program session constructs a 162 point design on \(X\) that has an approximate efficiency of 78% for the IavA design criterion. Because of its size the design, we present only the program session. Moreover, to reduce the volume of the discussion, the session has been edited.
Enter the number of controllable variates -- $K_d$ 7
Enter the number of environmental variates -- $K_e$ 5
Enter the scaling factor of $X_d$ -- $S_d$ 1.5
Enter the ratio -- $\theta = \gamma / \lambda$
For the $D$-criterion the ratio = 1.0 .5

Please wait a couple minutes.

Optimal values for $r_1$, $r_2$, $t_1$, and $t_2$
.73630 .70116 .52092 .49711

Here is a sample design
A 430 point optimal design requires
1 reps of a 1/2 fraction of the 0th isozero set (64 points)
1 reps of a 1/2 fraction of the 2th (0) isozero set (96 points)
1 reps of a 1/2 fraction of the 2th (1) isozero set (240 points)
2 reps of a 1/1 fraction of the 6th (0) isozero set (12 points)
3 reps of a 1/1 fraction of the 6th (1) isozero set (2 points)

$I_{avA}$ criterion efficiency is .99790

Is this design okay? (y or n) n

Change the isozero sets, fractions or replications (i, f or r) i
Enter the three isozero sets -- Must include either 0 or $K_d$
5 6 7
An optimal design is not possible. Proceed anyway? (y or n) y

A 99 point optimal design requires
1 reps of a 1/1 fraction of the 7th isozero set (1 points)
1 reps of a 1/1 fraction of the 5th (0) isozero set (60 points)
1 reps of a 1/1 fraction of the 5th (1) isozero set (24 points)
1 reps of a 1/1 fraction of the 6th (0) isozero set (12 points)
1 reps of a 1/1 fraction of the 6th (1) isozero set (2 points)
IavA criterion efficiency is .17364

Is this design okay? (y or n)

n

Change the isozero sets, fractions or replications (i, f or r)

i

Enter the three isozero sets -- Must include either 0 or Kd

0 5 6

An optimal design is not possible. Proceed anyway? (y or n)

y

A 162 point optimal design requires

1 reps of a 1/2 fraction of the 0th isozero set (64 points)
1 reps of a 1/1 fraction of the 5th (0) isozero set (60 points)
1 reps of a 1/1 fraction of the 5th (1) isozero set (24 points)
1 reps of a 1/1 fraction of the 6th (0) isozero set (12 points)
1 reps of a 1/1 fraction of the 6th (1) isozero set (2 points)

IavA criterion efficiency is .77828

Is this design okay? (y or n)

y

Divide the noise margins into how many regions

6

The multivariate entropy efficiency is

.93100

The univariate entropy efficiencies are

.909 .914 .890 .891 .904

The product correlations between columns are

.000 .000 .000 .000

.000 .000 .000

.000 .000

.000

Are these efficiencies okay? (y or n)

y

filename for the output : TEMP puts output to the screen
For this example, we clearly had to choose a different triple of isozero sets to keep the size of the design reasonable. The two collections of three isozero sets (not of the form $\mathcal{D}(0, j, 7)$) that have the smallest cardinalities, after fractionating, are $\mathcal{D}(5, 6, 7)$ with 99 points and $\mathcal{D}(0, 5, 6)$ with 162 points. Even though neither of these triples can support an optimal design, the program does allow the examination of designs on these support sets. Clearly, the single replicate design on $\mathcal{D}(5, 6, 7)$ is not sufficiently efficient: subsequent manipulations of the replication sizes, not shown, did not produce an efficiency greater than .175. Switching to the design on $\mathcal{D}(0, 5, 6)$ we see that we have dramatically increased the efficiency of the design. Subsequent manipulations of the replication sizes, not shown, produced designs with smaller efficiencies. Therefore, we accepted this collection of inputs for the control variables.

For the noise variables we divided each of the five univariate margins into six equal subintervals. This partitioned the space $\mathcal{X}_e$ into $5^6$ equal regions. For this large sample size, neither the multivariate entropy efficiency nor any of the univariate efficiencies equaled one. Several attempts were made, and the best collection is shown. □

Since $\theta = 0.5 \neq 1$, $r_1 \neq r_2$ and $t_1 \neq t_2$, but the differences between the both sets of parameters are small (approx. 0.03). For the 430 point suggested design, the difference in replications between the two sets $\mathcal{H}^{6(0)}$ and $\mathcal{H}^{6(1)}$ nearly accounted for both moment differences without heavier replication of the set $\mathcal{H}^{2(1)}$ in comparison to $\mathcal{H}^{2(0)}$. This suggests it may not be necessary to split two isozero sets to obtain an efficient design. If this is the case, then we can consider designs of the form $\mathcal{D}(0, j, k_d)$. In addition, this example indicates our method for generating the $k_e$ columns of noise
variables does not adequately spread the design points over the space $X_r$. We shall address both of these issues further in Chapter VI.
Table 6: Values of $r_1, r_2, t_1$ and $t_2$ in Conditions ii.) through v.) of Theorem 3.5.1 for InvA-optimal designs with respect to a prior $\pi^*$ for which the hypotheses in Assumptions 3.2.1 are satisfied. The highlighted designs are constructed in Section 3.8.

<table>
<thead>
<tr>
<th>$k_d$</th>
<th>$\theta$</th>
<th>$s_d$</th>
<th>$r_1$</th>
<th>$r_2$</th>
<th>$t_1$</th>
<th>$t_2$</th>
<th>$f_*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td></td>
<td>0.5</td>
<td>0.601</td>
<td>0.561</td>
<td>0.352</td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.610</td>
<td>0.574</td>
<td>0.363</td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>0.623</td>
<td>0.593</td>
<td>0.381</td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>1.0</td>
<td>0.5</td>
<td>0.576</td>
<td>0.576</td>
<td>0.347</td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>$\textbf{0.587}$</td>
<td>$\textbf{0.587}$</td>
<td>$\textbf{0.359}$</td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>0.604</td>
<td>0.604</td>
<td>0.377</td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td>2.0</td>
<td></td>
<td>0.5</td>
<td>0.561</td>
<td>0.601</td>
<td>0.352</td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.574</td>
<td>0.610</td>
<td>0.363</td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>0.593</td>
<td>0.623</td>
<td>0.381</td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>1.0</td>
<td>0.5</td>
<td>0.678</td>
<td>0.637</td>
<td>0.440</td>
<td>0.415</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.682</td>
<td>0.643</td>
<td>0.446</td>
<td>0.423</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>0.688</td>
<td>0.653</td>
<td>0.456</td>
<td>0.435</td>
<td>0.000</td>
</tr>
<tr>
<td>2.0</td>
<td></td>
<td>0.5</td>
<td>0.644</td>
<td>0.644</td>
<td>0.424</td>
<td>0.424</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>$\textbf{0.650}$</td>
<td>$\textbf{0.650}$</td>
<td>$\textbf{0.431}$</td>
<td>$\textbf{0.431}$</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>0.659</td>
<td>0.659</td>
<td>0.442</td>
<td>0.442</td>
<td>0.000</td>
</tr>
<tr>
<td>10</td>
<td>1.0</td>
<td>0.5</td>
<td>0.621</td>
<td>0.656</td>
<td>0.417</td>
<td>0.439</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.628</td>
<td>0.661</td>
<td>0.425</td>
<td>0.446</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>0.640</td>
<td>0.670</td>
<td>0.437</td>
<td>0.456</td>
<td>0.000</td>
</tr>
<tr>
<td>2.0</td>
<td></td>
<td>0.5</td>
<td>0.761</td>
<td>0.725</td>
<td>0.555</td>
<td>0.529</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.763</td>
<td>0.727</td>
<td>0.558</td>
<td>0.533</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>0.765</td>
<td>0.731</td>
<td>0.563</td>
<td>0.538</td>
<td>0.000</td>
</tr>
<tr>
<td>10</td>
<td>1.0</td>
<td>0.5</td>
<td>0.727</td>
<td>0.727</td>
<td>0.533</td>
<td>0.533</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.729</td>
<td>0.729</td>
<td>0.536</td>
<td>0.536</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>0.733</td>
<td>0.733</td>
<td>0.542</td>
<td>0.542</td>
<td>0.000</td>
</tr>
<tr>
<td>2.0</td>
<td></td>
<td>0.5</td>
<td>0.702</td>
<td>0.732</td>
<td>0.518</td>
<td>0.539</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0</td>
<td>0.705</td>
<td>0.734</td>
<td>0.522</td>
<td>0.543</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.5</td>
<td>0.710</td>
<td>0.737</td>
<td>0.528</td>
<td>0.548</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Table 7: A 25 point integrated asymptotic variance D-optimal (resp. A-optimal with \( \theta = 1 \)) design on \([-1, 1]^6\) for \((k_d, k_e) = (2, 4)\).

<table>
<thead>
<tr>
<th>(x_{d1})</th>
<th>(x_{d2})</th>
<th>(x_{e1})</th>
<th>(x_{e2})</th>
<th>(x_{e3})</th>
<th>(x_{e4})</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>-0.24</td>
<td>0.67</td>
<td>0.60</td>
<td>0.15</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>0.28</td>
<td>0.08</td>
<td>-0.73</td>
<td>0.81</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>-0.59</td>
<td>-0.85</td>
<td>0.39</td>
<td>0.98</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>-0.01</td>
<td>-0.70</td>
<td>0.40</td>
<td>-0.73</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>0.39</td>
<td>-0.82</td>
<td>-0.37</td>
<td>-0.17</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>0.13</td>
<td>-0.23</td>
<td>0.96</td>
<td>0.84</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>0.74</td>
<td>0.70</td>
<td>-0.16</td>
<td>1.00</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>0.16</td>
<td>0.55</td>
<td>-0.17</td>
<td>0.71</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>0.97</td>
<td>-0.31</td>
<td>0.26</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>0.67</td>
<td>-0.16</td>
<td>-0.17</td>
<td>-0.04</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>0.50</td>
<td>-1.00</td>
<td>0.57</td>
<td>0.67</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>-0.85</td>
<td>-0.42</td>
<td>0.03</td>
<td>0.39</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>0.19</td>
<td>0.33</td>
<td>-0.71</td>
<td>-0.88</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>-0.12</td>
<td>0.89</td>
<td>-0.32</td>
<td>-0.31</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>-0.34</td>
<td>-0.47</td>
<td>0.06</td>
<td>-0.86</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>0.45</td>
<td>0.51</td>
<td>-0.26</td>
<td>0.49</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>0.75</td>
<td>0.81</td>
<td>0.43</td>
<td>0.48</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>-0.44</td>
<td>0.53</td>
<td>0.80</td>
<td>-0.21</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>-0.72</td>
<td>-0.04</td>
<td>-0.95</td>
<td>0.42</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>-0.32</td>
<td>-0.74</td>
<td>-1.00</td>
<td>-0.23</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.47</td>
<td>0.25</td>
<td>0.73</td>
<td>0.62</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.78</td>
<td>-0.34</td>
<td>-0.19</td>
<td>-0.13</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>-0.64</td>
<td>0.12</td>
<td>0.77</td>
<td>-0.35</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.77</td>
<td>-0.46</td>
<td>0.47</td>
<td>0.03</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>-0.76</td>
<td>-0.17</td>
<td>-0.26</td>
<td>-0.26</td>
</tr>
</tbody>
</table>
Table 8: A 64 point integrated asymptotic variance D-optimal (resp. A-optimal with $\theta = 1$) design on $[-1, 1]^9$ for $(k_d, k_e) = (4, 5)$.

<table>
<thead>
<tr>
<th>$x_{d1}$</th>
<th>$x_{d2}$</th>
<th>$x_{d3}$</th>
<th>$x_{d4}$</th>
<th>$x_{e1}$</th>
<th>$x_{e2}$</th>
<th>$x_{e3}$</th>
<th>$x_{e4}$</th>
<th>$x_{e5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>0.23</td>
<td>-0.07</td>
<td>0.46</td>
<td>-0.67</td>
<td>0.95</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.10</td>
<td>0.76</td>
<td>-0.20</td>
<td>0.22</td>
<td>0.00</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-0.95</td>
<td>-0.39</td>
<td>0.86</td>
<td>-0.31</td>
<td>-0.41</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-0.14</td>
<td>0.63</td>
<td>0.55</td>
<td>-0.21</td>
<td>-0.30</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>0.45</td>
<td>-0.47</td>
<td>0.05</td>
<td>0.47</td>
<td>-0.77</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-0.03</td>
<td>-0.08</td>
<td>0.44</td>
<td>0.55</td>
<td>0.68</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-0.19</td>
<td>0.19</td>
<td>-0.56</td>
<td>-0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.36</td>
<td>-0.55</td>
<td>0.15</td>
<td>-0.84</td>
<td>-0.57</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>0.15</td>
<td>-0.25</td>
<td>0.08</td>
<td>-0.12</td>
<td>-0.79</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.40</td>
<td>-0.97</td>
<td>-0.63</td>
<td>0.63</td>
<td>-0.06</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-0.46</td>
<td>-0.12</td>
<td>-0.05</td>
<td>0.77</td>
<td>-0.25</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-0.91</td>
<td>0.10</td>
<td>-0.20</td>
<td>0.07</td>
<td>-0.71</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-0.26</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.03</td>
<td>0.53</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.63</td>
<td>0.81</td>
<td>0.81</td>
<td>-0.72</td>
<td>-0.03</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-0.89</td>
<td>0.26</td>
<td>0.66</td>
<td>-0.65</td>
<td>0.32</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.15</td>
<td>-0.90</td>
<td>0.27</td>
<td>-0.17</td>
<td>0.30</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-0.43</td>
<td>-0.26</td>
<td>-0.44</td>
<td>0.70</td>
<td>0.24</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.07</td>
<td>0.48</td>
<td>0.52</td>
<td>-0.81</td>
<td>-0.56</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.97</td>
<td>0.38</td>
<td>-0.33</td>
<td>-0.81</td>
<td>0.40</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.26</td>
<td>-0.44</td>
<td>-0.64</td>
<td>0.25</td>
<td>0.80</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-0.47</td>
<td>0.69</td>
<td>-0.26</td>
<td>-0.37</td>
<td>0.60</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.08</td>
<td>-0.77</td>
<td>-0.77</td>
<td>-0.17</td>
<td>-0.43</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.97</td>
<td>-0.37</td>
<td>-0.17</td>
<td>0.32</td>
<td>-1.00</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-0.62</td>
<td>0.84</td>
<td>-0.15</td>
<td>0.49</td>
<td>0.64</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-0.24</td>
<td>0.11</td>
<td>-0.58</td>
<td>-0.24</td>
<td>-0.01</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-0.63</td>
<td>-0.88</td>
<td>-0.30</td>
<td>-0.75</td>
<td>0.31</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.67</td>
<td>0.41</td>
<td>0.34</td>
<td>0.70</td>
<td>0.13</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>-0.20</td>
<td>0.32</td>
<td>0.53</td>
<td>-0.40</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>0.48</td>
<td>-0.01</td>
<td>0.80</td>
<td>-0.09</td>
<td>-0.76</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-0.72</td>
<td>0.37</td>
<td>0.25</td>
<td>0.70</td>
<td>0.09</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-0.23</td>
<td>-0.63</td>
<td>-0.64</td>
<td>-0.26</td>
<td>0.67</td>
</tr>
</tbody>
</table>
Table 8 (continued)

<table>
<thead>
<tr>
<th>$x_{d_1}$</th>
<th>$x_{d_2}$</th>
<th>$x_{d_3}$</th>
<th>$x_{d_4}$</th>
<th>$x_{e_1}$</th>
<th>$x_{e_2}$</th>
<th>$x_{e_3}$</th>
<th>$x_{e_4}$</th>
<th>$x_{e_5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.79</td>
<td>0.26</td>
<td>-0.19</td>
<td>-0.46</td>
<td>0.24</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>0.09</td>
<td>0.42</td>
<td>-0.75</td>
<td>0.70</td>
<td>-0.59</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>-0.27</td>
<td>-0.68</td>
<td>-0.20</td>
<td>0.25</td>
<td>0.61</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>-0.12</td>
<td>-0.34</td>
<td>0.63</td>
<td>-0.62</td>
<td>0.61</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.54</td>
<td>0.85</td>
<td>-0.31</td>
<td>0.16</td>
<td>-0.29</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>0.52</td>
<td>0.43</td>
<td>0.07</td>
<td>-0.48</td>
<td>0.59</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.61</td>
<td>0.25</td>
<td>0.70</td>
<td>-0.61</td>
<td>0.36</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>0.63</td>
<td>0.19</td>
<td>-0.44</td>
<td>0.77</td>
<td>-0.27</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-0.06</td>
<td>-0.53</td>
<td>0.48</td>
<td>-0.14</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>0.21</td>
<td>-0.35</td>
<td>0.49</td>
<td>0.40</td>
<td>-0.91</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.65</td>
<td>-0.19</td>
<td>-1.00</td>
<td>-0.18</td>
<td>0.52</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>-0.39</td>
<td>0.71</td>
<td>0.06</td>
<td>0.09</td>
<td>-0.53</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>-0.70</td>
<td>-0.18</td>
<td>-0.25</td>
<td>0.50</td>
<td>0.62</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>-0.84</td>
<td>0.66</td>
<td>-0.41</td>
<td>0.42</td>
<td>0.96</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.46</td>
<td>0.04</td>
<td>0.67</td>
<td>0.37</td>
<td>-0.24</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>-0.12</td>
<td>-0.88</td>
<td>-0.45</td>
<td>0.12</td>
<td>-0.22</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.79</td>
<td>0.73</td>
<td>-0.08</td>
<td>0.40</td>
<td>-0.61</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>-0.44</td>
<td>-0.57</td>
<td>0.84</td>
<td>0.98</td>
<td>0.44</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.26</td>
<td>0.99</td>
<td>-0.76</td>
<td>-0.59</td>
<td>-0.76</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>-0.48</td>
<td>0.42</td>
<td>-0.57</td>
<td>-0.67</td>
<td>-0.78</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.06</td>
<td>-0.81</td>
<td>0.87</td>
<td>1.00</td>
<td>0.21</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>0.79</td>
<td>-0.35</td>
<td>-0.61</td>
<td>-0.16</td>
<td>-0.20</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>-0.52</td>
<td>0.71</td>
<td>-0.39</td>
<td>-0.01</td>
<td>0.57</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>0.54</td>
<td>0.74</td>
<td>0.69</td>
<td>0.66</td>
<td>0.90</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>-0.49</td>
<td>-0.56</td>
<td>0.82</td>
<td>-0.39</td>
<td>-0.88</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.89</td>
<td>-0.77</td>
<td>0.21</td>
<td>-0.73</td>
<td>-0.61</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.59</td>
<td>-0.22</td>
<td>0.58</td>
<td>-0.70</td>
<td>-0.46</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.16</td>
<td>0.33</td>
<td>0.22</td>
<td>0.98</td>
<td>-0.03</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.77</td>
<td>-0.44</td>
<td>0.33</td>
<td>0.02</td>
<td>0.89</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.45</td>
<td>-0.71</td>
<td>-0.79</td>
<td>-0.73</td>
<td>0.04</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.82</td>
<td>0.60</td>
<td>-0.77</td>
<td>0.10</td>
<td>-0.62</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.12</td>
<td>1.00</td>
<td>0.56</td>
<td>0.08</td>
<td>0.16</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.65</td>
<td>-0.87</td>
<td>0.11</td>
<td>-0.40</td>
<td>0.63</td>
</tr>
</tbody>
</table>
CHAPTER IV

A Comparison of Parameter Design Methodologies

4.1 Introduction

In this chapter we perform two simulation experiments to compare our parameter design methodology for determining the optimal operating conditions with four competing methodologies that also use the response model approach described in Section 1.2.2. In the simulation experiments we define a function $y(x|x_d^*)$ of a vector $x \in X$. The function is parameterized by the unique optimal operating condition, $x_d^* \in X$, for the integrated loss function $L(x_d|x_d^*)$, i.e., $x_d^* = \text{argmin } L(x_d|x_d^*)$. For each simulation run, we randomly generate the true value of $x_d^*$. Thus, the response function, $y(\cdot|x_d^*)$, and the loss function, $L(\cdot|x_d^*)$ are different for each run. During each run, we record the estimates $\hat{x}_d^*$ of the true optimum operating conditions $x_d^*$ and $L(\hat{x}_d^*; x_d^*)$ of the true minimum value of the loss function $L(x_d^*|x_d^*)$ for each of the five methodologies. Comparisons of the methodologies are then made in terms of $L(x_d^*; x_d^*)/L(\hat{x}_d^*; x_d^*)$ and the distance between $\hat{x}_d^*$ and $x_d^*$.

A parameter design methodology based on the response model approach has three components: specification of a model for the quality characteristic, selection of an
experimental design based on this model and identification of a method for estimating the minimizing value of the integrated loss function $L(\cdot|x_d^*)$. In Section 4.2, we present the respective models ($Q$-LM or $\mu$-LM), designs (IavAD, AMSE$_Q$, AMSE$_\mu$ or OA$_2$-LHS), and methods of estimation (MLE, MinUP or MC-BLUP) used by the methodologies in the simulation experiments. We designate each of the methodologies with a three-level name that describes the model, the experimental design, and the method of estimation that are used. The description of the five methodologies along with their three-level names is given in Section 4.2.4. Section 4.3 presents the first simulation experiment. It compares the methodologies for a $(k_d, k_r) = (2,4)$ parameter design problem using 25 point experimental designs. The second simulation experiment, which compares the methodologies for a $(k_d, k_r) = (4,5)$ parameter design problem using 64 point experimental designs, is given in Section 4.4.

4.2 Defining the Methodology Components

This section presents the competing parameter design methodologies that are compared in the simulation experiments. A parameter design methodology based on the response model approach has three components: the model, the design and the method of estimation. Section 4.2.1 presents the two models for the quality characteristic that are encountered in the simulations. Section 4.2.2 gives the experimental designs used by the competing methodologies, and Section 4.2.3 describes the three methods of estimation of $x_d^*$ used in the simulation experiment. In Section 4.2.4, we define the competing methodologies.
4.2.1 Models

There are two models for the quality characteristic that are used in the competing methodologies. Both have been defined and discussed previously in this dissertation. Readers can refer back to the discussions in Section 1.1.1 and Section 1.2.3.

There are three assumptions that we make for both models. First, suppose $X = \mathcal{X}_d \times \mathcal{X}_r$ where $\mathcal{X}_d = [-1, 1]^k_d$ and $\mathcal{X}_r = [-1, 1]^k_r$. Second, assume that the random environmental variables $\mathbf{x}_r$ are distributed according to $\pi_r$ for which $\mu_r = \mathbf{E}_r[\mathbf{x}_r] = \mathbf{0}_{k_r}$ and $\Sigma_r = \mathbf{E}_r[\mathbf{x}_r\mathbf{x}_r'] = \mathbf{I}_{k_r}$. Third, regard $y(\mathbf{x}_d, \mathbf{x}_r|\mathbf{x}_d^*)$ as a realization of the stochastic process $Y(\mathbf{x}_d, \mathbf{x}_r)$ that includes a regression model.

The first stochastic process model used in the simulation experiments contains a second-order model for the controllable factors, i.e.,

**Q-LM**

$$Y(\mathbf{x}_d, \mathbf{x}_r) = \beta_0 + \beta'_d \mathbf{x}_d + \mathbf{x}_d' \mathbf{B}_d \mathbf{x}_d + \beta'_r \mathbf{x}_r + \mathbf{x}_r' \mathbf{B}_r \mathbf{x}_r + Z(\mathbf{x}_d, \mathbf{x}_r) \quad (4.2.1)$$

where the regression coefficient arrays are defined in (1.2.7) through (1.2.10) with $\mathbf{B}_d$ assumed to be positive definite. The error term $Z(\cdot, \cdot)$ in (4.2.1) is a mean zero Gaussian process with constant variance $\sigma^2$ and zero correlations between pairs of distinct input sites. Note that the number of regression parameters in (4.2.1) is $\binom{k_d + 2}{2} + (k_d + 1)k_r$. We require the experimental designs in Section 4.2.2 to have at least this many points.
The second stochastic process model employed in the simulation experiments is
the constant mean model. That is,

\[ \mu-LM \]

\[ Y(\mathbf{x}_d, \mathbf{x}_e) = \mu + Z(\mathbf{x}_d, \mathbf{x}_e). \]

(4.2.2)

where \( Z(\cdot, \cdot) \) is a mean zero Gaussian stochastic process with constant variance \( \sigma_Z^2 \) and correlation between observations at the input sites \((\mathbf{x}_d^{(1)}, \mathbf{x}_e^{(1)})\) and \((\mathbf{x}_d^{(2)}, \mathbf{x}_e^{(2)})\) given by

\[ \prod_{i=1}^{k_d} \exp\{-\theta_i |x_d^{(1)} - x_d^{(2)}|^2\} \prod_{j=1}^{k_e} \exp\{-\theta_j |x_e^{(1)} - x_e^{(2)}|^2\}. \]

(4.2.3)

In (4.2.3) we defined the smoothness parameters of (1.1.2) to equal two. Fixing the values of these parameters halves the number of correlation parameters that must be estimated (see Example 1.1.1). Moreover, a draw from the resulting process is very smooth: it has infinitely many derivatives.

4.2.2 Design Criteria

In this section we present the four design strategies used in each simulation and discuss the methods used to derive them. Each design strategy has already been described in this dissertation (see Section 1.1.2 and Chapter III). Therefore, the discussion in this section primarily entails a presentation of the input values for computer programs that generate the optimal designs.

The size of the four experimental designs for the \((k_d, k_e) = (2, 4)\) problem is 25, while the four designs for the \((k_d, k_e) = (4, 5)\) problem have 64 points. These sample
sizes arise from two requirements that are imposed by the competing methodologies. First, there must be at least \( \left( \frac{k_d+2}{2} \right) + (k_d + 1)k_e \), i.e., 18 and 40, points in the experimental design to allow estimation of all the regression parameters in (4.2.1). Second, for the orthogonal based Latin hypercube designs, there must exist strength two orthogonal arrays with at least \( k_d + k_e \) columns. The smallest sample sizes for which both requirements are met are 25 and 64, respectively. The 25 point orthogonal array has five levels for each of the six variables in the \((k_d, k_e) = (2, 4)\) problem, and the 64 point orthogonal array has eight levels for each of the nine variables in the \((k_d, k_e) = (4, 5)\) problem.

The first experimental design strategy used in the simulations was developed in Chapter III. It is based on the Q-LM model in (4.2.1). In the simulation experiments, we assume that the ratio of the eigenvalues of the completely symmetric, positive definite matrix \( Q \) in Hypothesis iii.) of Assumptions 3.2.1 is equal to one. This makes the designs in Theorem 3.5.1 and Theorem 3.6.1 identical. We call this design strategy lavAD. We formulated the first two examples of Section 3.8 in a way such that we could use the designs for the simulation experiments. The design constructed in Example 3.8.1 and given in Table 7 is used in the \((k_d, k_e) = (2, 4)\) problem, and the design constructed in Example 3.8.2 and given in Table 8 is used in the \((k_d, k_e) = (4, 5)\) problem.

The next two design strategies are based on different models, but they both use the average mean square error (AMSE) design criterion (Welch, 1983) described in
Section 1.1.2. The first design strategy, based on the Q-LM model in (4.2.1), is denoted by $\text{AMSE}_Q$. The second design strategy is based on the $\mu$-LM model in (4.2.2), and it is denoted $\text{AMSE}_\mu$.

The AMSE criterion considers both the error that arises from model inadequacy and random error. The ACED experimental design software package (Welch, 1985) determines optimal designs with respect to this criterion. To construct AMSE optimal designs using ACED, one must specify the model, define a list of candidate points, and quantify the relative magnitude of the bias error to the variance error. For computer experiments the error arises solely from model inadequacy, so we weighed the bias portion of the AMSE heavily in the ACED program. In both simulation experiments, and for both models, we defined the ratio between the absolute worst-case departure and the standard deviation to equal 100.

For the Q-LM model in (4.2.1), we restricted the candidate levels of the controllable variables to the lattice $\{-1, 0, 1\}^{k_d}$, and we desired the corresponding levels of the environmental variables $x_e$ to be uniformly distributed over $X_e = [-1, 1]^{k_e}$. To construct the set of candidate points, we generated five 10 point Latin hypercube samples on $[-1, 1]^{k_e}$ for each of the $3^{k_d}$ points the lattice. This created a set of $3^{k_d} \cdot 50$ candidate design points. For both simulations we determined the $\text{AMSE}_Q$-optimal design by executing the excursion algorithm in ACED five times (starting with a different design each time) and then selecting the design that performed the best. The 25 point $\text{AMSE}_Q$-optimal design for the $(k_d, k_e) = (2, 4)$ problem is given in Table 16. Table 17 gives the 64 point $\text{AMSE}_Q$-optimal design for the $(k_d, k_e) = (4, 5)$ problem.
For the $\mu$-LM model in (4.2.2), we wanted the set of candidate points to be uniformly distributed over $\mathcal{X} = [-1, 1]^{k_d+k_e}$. To construct the list of candidate points, we generated ten 20 point Latin hypercube samples from $\mathcal{X} = [-1, 1]^{k_d+k_e}$. This means there was a list of 200 candidate points for each simulation. For each simulation experiment, we determined the AMSE$_\mu$-optimal design by executing the excursion algorithm five times (starting with a different design each time) and selecting the best design. In Table 18 we give the 25 point AMSE$_\mu$-optimal design for the $(k_d, k_e) = (2, 4)$ problem. Table 19 gives the 64 point AMSE$_\mu$-optimal design for the $(k_d, k_e) = (4, 5)$ problem.

The fourth and final design strategy uses strength two orthogonal array based Latin hypercube designs. We denote this strategy by OA$_{2}$-LHS. The strength two property provides uniform distributions of design points on each of the $\binom{k_d+k_e}{2}$ bivariate margins as well as the $k_d + k_e$ univariate margins. The designs for the simulation experiments are constructed using a public-domain software package developed by Owen (1994). The algorithm in Owen's package that we used is described in Bose (1938). The program produces a $(k_d+k_e-1)^2 \times (k_d+k_e)$ orthogonal array, $O_A$, with levels $\{0, \ldots, k_d+k_e-2\}$. We converted $O_A$ to a Latin hypercube design $X_{LHS}$, on $\mathcal{X}$, using the transformation

$$(X_{LHS})_{i,j} = \frac{(O_A)_{i,j} + 1 - U_{i,j}}{k_d + k_e - 1}$$

for $i = 1, 2, \ldots, (k_d+k_e-1)^2$, $j = 1, 2, \ldots, k_d+k_e$ and where $\{U_{i,j}\}$ is a collection of $(k_d+k_e-1)^2(k_d+k_e)$ independent Uniform$(0,1)$ random variables. In Table 20 we
give the 25 point OA2-LHS design for the \((k_d, k_c) = (2, 4)\) problem. Table 21 gives
the 64 point OA2-LHS design for the \((k_d, k_c) = (4, 5)\) problem.

4.2.3 Estimating the Optimal Operating Levels

Three methods of estimating \(x_d^*\) are used in the simulation experiments. Two of these
methods are applied with the model in (4.2.1), while the final method is used with
the model in (4.2.2).

For the quadratic response model (4.2.1), the expected integrated loss function
in (1.2.4) equals

\[
L(x_d|x_d^*) = \beta_0 + \beta_d'x_d + x_d'B_dx_d. \tag{4.2.5}
\]

Since we assume \(B_d\) is positive definite, the function \(L(\cdot|x_d^*)\) is uniquely minimized
at the point \(x_d^* = (1/2)B_d^{-1}\beta_d\).

The first estimation scheme in the simulation experiments uses the estimator

\[
\hat{x}_d^* = -\frac{1}{2} \hat{B}_d^{-1}\hat{\beta}_d \tag{4.2.6}
\]

of \(x_d^*\) where the matrices \(\hat{B}_d\) and \(\hat{\beta}_d\) contain the least-squares estimators of the regression coefficients in \(B_d\) and \(\beta_d\). Under homogeneity and independence assumptions for
the error term \(Z(\cdot, \cdot)\), the least-squares estimators are also the maximum likelihood estimators. By invariance, \(\hat{x}_d^*\) in (4.2.6) is the maximum likelihood estimator of \(x_d^*\).

For each run in the simulation experiments the true optimal operating level \(x_d^*\)
is randomly generated from \([-1, 1]^{k_d}\), yet estimates of \(x_d^*\) provided by the m.l.e. \(\hat{x}_d^*\)
may not be in \([-1, 1]^{k_d}\). The remaining estimation procedures used in the simulation experiments always provide an estimate of \(x_d^*\) in \([-1, 1]^{k_d}\), since they search for the estimate over the region \([-1, 1]^{k_d}\). To balance the comparison of the methodologies, we truncate (coordinate-wise) the estimate given by \(\hat{x}_d^*\) whenever its value lies outside \([-1, 1]^{k_d}\). For example, when the value of the estimate is \((-1.4, .23, 1.8, 2.3)\), we truncate it to \((-1.0, .23, 1.0, 1.0)\).

To construct the maximum likelihood estimator, we assumed the matrix \(B_d\) in Q-LM model is positive definite. When this assumption is not appropriate, the estimator \(\hat{x}_d^*\) in (4.2.6) is not estimating the optimal operating conditions. Instead, it is estimating some saddlepoint of the function \(L(\cdot | x_d^*)\) in (4.2.5). However, even if \(B_d\) in the Q-LM model is not positive definite, the function

\[
\hat{L}(x_d | x_d^*) = \hat{\beta}_0 + \hat{\beta}_d^t x_d + x_d^t \hat{B}_d x_d.
\]

is still an unbiased predictor of \(L(\cdot | x_d^*)\).

The second method of estimation in the simulation experiments estimates the value of \(x_d^*\) by minimizing the predictor \(\hat{L}(x_d | x_d^*)\) in (4.2.7) over \([-1, 1]^{k_d}\), i.e.,

\[
\text{MinUP}
\]

\[
\hat{x}_d^* = \underset{\{-1,1\}^{k_d}}{\arg \min} \; \hat{\beta}_0 + \hat{\beta}_d^t x_d + x_d^t \hat{B}_d x_d
\]

where \(\hat{\beta}_0\), \(\hat{B}_d\) and \(\hat{\beta}_d\) contain the least-squares estimators of the regression coefficients in \(\beta_0\), \(B_d\) and \(\beta_d\). To determine \(x_d^*\) in (4.2.8), we find the minimum of the right hand side using simulated annealing. In the simulation experiments, the starting point for \(x_d^*\) in the simulated annealing algorithm is selected randomly from \([-1, 1]^{k_d}\).
For the $\mu$-LM model in (4.2.2) the best linear unbiased predictor of $y(x_d, x_e|x_d^*)$ is

$$
\hat{y}(x_d, x_e) = \mu + r'(x_d, x_e)R^{-1}(y - 1_n\hat{\mu}).
$$

(4.2.9)

where $\hat{\mu} = (1_n^tR^{-1}1_n)1_n^tR^{-1}y$ is the generalized least-squares estimate of $\mu$ (see Example 1.1.1). The predictor of the integrated loss function $L(\cdot|\mathbf{x}_d^*)$ is then given by

$$
\hat{L}(x_d|x_d^*) = \int_{x_e} \hat{\mu} + r'(x_d, x_e)R^{-1}(y - 1_n\hat{\mu})d\pi_e(x_e),
$$

(4.2.10)

and the optimal operating conditions $x_d^*$ can be determined by minimizing the predictor in (4.2.10). However, Monte Carlo methods must be used to estimate the integral in (4.2.10). For the simulation experiments, we generated a 50 point strength two orthogonal array based Latin hypercube sample for the environmental variables, i.e.,

$\{x_e^{(i)}\}_{i=1}^{50}$, and we estimated the integral in (4.2.10) by the sample average, i.e.,

$$
\hat{L}(x_d|x_d^*) = \frac{1}{50} \sum_{i=1}^{50} \left[ \hat{\mu} + r'(x_d, x_e^{(i)})R^{-1}(y_n - 1_n\hat{\mu}) \right].
$$

(4.2.11)

Thus, the third method of estimation in the simulation experiments estimates $x_d^*$ by

**MC-BLUP**

$$
x_d^* = \arg\min_{(-1,1)^k} \frac{1}{50} \sum_{i=1}^{50} \left[ \hat{\mu} + r'(x_d, x_e^{(i)})R^{-1}(y_n - 1_n\hat{\mu}) \right].
$$

(4.2.12)

To determine the minimizing value $x_d^*$ in (4.2.12), we used simulated annealing. The starting point for the algorithm was selected at random from $[-1,1]^k$. Further, note the 50 point orthogonal array based Latin hypercube sample was constructed by the software of Owen (1994) using the algorithm in Bose and Bush (1952).
4.2.4 Candidate Methodologies

The methodology developed in this dissertation is called the Q-LM/IavAD/MLE methodology. We have noted that the performance of this methodology relies heavily on correctness of the assumption of positive definiteness of the matrix $B_d$. As a check, we include the Q-LM/IavAD/MinUP methodology in the simulation experiments. We should expect the Q-LM/IavAD/MinUP to perform better than Q-LM/IavAD/MLE, since the MinUP method of estimation of $x_d^*$ should do as well as the MLE method for $B_d$ positive definite and better otherwise. For the Q-LM model in (4.2.1), we also consider the Q-LM/AMSE$_Q$/MinUP methodology. Poorer performance of this methodology compared to the Q-LM/IavAD/MinUP methodology would support our hypothesis that better estimates of $x_d^*$ can be obtained if the design is determined using a criterion that directly considers the estimation of $x_d^*$.

The last two methodologies in the simulation experiments are based on the $\mu$-LM model. The first, $\mu$-LM/AMSE$_{\mu}$/MC-BLUP, uses the all-bias optimal designs from ACED (Welch, 1985), while the second, $\mu$-LM/OA$_2$-LHS/MC-BLUP, employs strength two orthogonal array based Latin hypercube designs. Comparisons of the $\mu$-LM based methodologies with the Q-LM based methodologies provide evaluation of the mean-dominant Q-LM model against the correlation-dominant $\mu$-LM model for the parameter design problem. In addition, comparisons between the $\mu$-LM based methodologies may indicate a preferable design strategy for the $\mu$-LM model. Table 9 gives the five methodologies used in both simulation experiments.
Table 9: The five competing methodologies in the two simulated parameter design computer experiments.

<table>
<thead>
<tr>
<th>Model:</th>
<th>Q-LM</th>
<th>( \mu )-LM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimator</td>
<td>IavAD</td>
<td>AMSE_Q</td>
</tr>
<tr>
<td>MLE</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>MinUP</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>MC-BLUP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.3 The \((k_d, k_e) = (2, 4)\) Problem

In this section we compare the five competing parameter design methodologies using 25 point designs for the response function

\[
y(x_d, x_e | x_d^*) = 8x_d x_e_2 - \sqrt{3} x_d x_e_3 + 6.5 x_e_4 + \left\{ 10 - 3.5e^{-(x_d - x_d^*)^2} \right\} \times 
\left( 3 - 0.7 x_e_1 + 5.5(x_d - x_d^*) \sin \left( \frac{\pi}{2} (x_d - x_d^*) \right) \right) .
\]  

(4.3.1)

where \(x_d = (x_{d1}, x_{d2})^t \in [-1, 1]^2\), \(x_e = (x_{e1}, x_{e2}, x_{e3}, x_{e4})^t \in [-1, 1]^4\) and \(x_d^* = (x_d^*, x_d^*)^t \in [-1, 1]^2\). Since \(E^n[x_e] = 0_4\), the true expected integrated loss function equals

\[
L(x_d | x_d^*) = (3 + 5.5(x_d - x_d^*) \sin \left( \frac{\pi}{2} (x_d - x_d^*) \right) \right) \left( 10 - 3.5e^{-(x_d - x_d^*)^2} \right) .
\]  

(4.3.2)

The loss function in (4.3.2) is minimized, uniquely, at the point \(x_d^*\). Figure 6 shows two pictures of the loss function \(L(\cdot | x_d^*)\) for different values for \(x_d^*\). These pictures suggest that the value of \(x_d^*\) is easier to determine than the value of \(x_d^*\) with the
difficulty of determining either coordinate increasing as the true optimal point $x_d^*$ moves away from $(0,0)$.

![Figure 6: Perspective plots of the loss function $L(x_d^*)$ for the $(k_d, k_c) = (2,4)$ parameter design problem for (a) $x_d^* = (0,0)^t$ and (b) $x_d^* = (-0.75, 0.50)^t$.](image)

For the simulation experiment, we divide each margin $[-1,1]$ for the controllable variables $x_d$ into ten equal intervals. This partitions the whole space $X_d = [-1,1]^2$ into 100 equal squares. We randomly sample one value of $x_d^*$ from each square, so that there are 100 runs in the simulation.

Comparisons of the methodologies are made in terms of the value of loss efficiency, i.e.,

$$\frac{L(x_d^*|x_d^*)}{L(\tilde{x}_d^*|x_d^*)}$$

(4.3.3)

The function in (4.3.3) is always less than one, and the closer its value is to one, the better the methodology performed. Table 10 presents the proportion of the 100 simulation runs that achieved at least the given loss efficiency levels for each of the methodologies. Clearly, the methodologies that use the $\mu$-LM model perform the best, while the methodology developed in this dissertation lags far behind the other
competitors. We also notice that the methodologies Q-LM/IavAD/MinUP and Q-LM/AMSE_Q/MinUP perform equally well. This implies that using a design criterion that explicitly considers the estimation of \( x_d^* \) provided little benefit in this problem.

Table 10: Proportion of the 100 simulation runs greater than or equal to the given loss efficiency levels for the \((k_d, k_r) = (2, 4)\) parameter design problem. Highlighted values in the table designate the optimal methodology or methodologies.

<table>
<thead>
<tr>
<th>Loss Efficiency Level</th>
<th>Q-LM IavAD</th>
<th>Q-LM IavAD</th>
<th>Q-LM AMSE_Q</th>
<th>Q-LM AMSE_Q</th>
<th>( \mu )-LM AMSE_( \mu )</th>
<th>( \mu )-LM OA_2-LHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.995</td>
<td>0.03</td>
<td>0.07</td>
<td>0.08</td>
<td>0.35</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>0.990</td>
<td>0.07</td>
<td>0.13</td>
<td>0.11</td>
<td>0.54</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>0.975</td>
<td>0.17</td>
<td>0.24</td>
<td>0.25</td>
<td>0.69</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>0.950</td>
<td>0.33</td>
<td>0.48</td>
<td>0.47</td>
<td>0.79</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>0.900</td>
<td>0.40</td>
<td>0.76</td>
<td>0.76</td>
<td>0.93</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>0.850</td>
<td>0.46</td>
<td>0.84</td>
<td>0.84</td>
<td>0.96</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>0.750</td>
<td>0.47</td>
<td>0.92</td>
<td>0.93</td>
<td>0.99</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>0.500</td>
<td>0.61</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

Before drawing conclusions based on Table 10, note that the response function \( y(\cdot, \cdot | x_d^* \) in (4.3.1) is an adaptation of an example in Welch et al (1992). Their article investigated using the \( \mu \)-LM model to predict deterministic computer data, so the methodologies that use the \( \mu \)-LM model should predict the response function \( y(\cdot, \cdot | x_d^* \) well and hence perform well in the simulation. Presumably, if the test model is a second-order (or near second-order) response, our methodology will perform better. In Figure 6, we see that when \( x_d = (0, 0) \), the test model (4.3.1) looks like a second-order response surface, but at \( x_d = (-.75, .5) \) it does not. Moreover, the design
region $X_d$ was centered by assuming $x_d^* = (0, 0)$. This means that as $x_d^*$ moves away from $(0, 0)$, not only does the initial guess for $x_d^*$ get worse, but the test model is less well approximated by a second order model. In Table 11 the loss efficiencies of the competing methodologies are retabulated using only the 64 runs of the simulation for which the true $x_d^*$ is in $[-0.8, 0.8]^2$. In comparison with Table 10, this table shows that the Q-LM/IAvAD/MLE methodology performs better when the true $x_d^*$ is close to $(0, 0)$. However, there is a similar improvement in each of the methodologies, and the $\mu$-LM model based methodologies are still superior. The same holds true if we consider further reductions of the space $[-1, 1]^2$ for $x_d^*$ (e.g., $[-0.6, 0.6]^2$ or $[-0.4, 0.4]^2$).

Table 11: Proportion of the 64 simulation runs, with $x_d^* \in [-0.8, 0.8]^2$, greater than or equal to the given loss efficiency levels for the $(k_d, k_c) = (2, 4)$ parameter design problem. Highlighted values in the table designate the optimal methodology or methodologies.

<table>
<thead>
<tr>
<th>Loss Efficiency Level</th>
<th>Q-LM</th>
<th>Q-LM</th>
<th>Q-LM</th>
<th>$\mu$-LM</th>
<th>$\mu$-LM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IavAD</td>
<td>IavAD</td>
<td>AMSE</td>
<td>AMSE</td>
<td>OA2-LHS</td>
</tr>
<tr>
<td></td>
<td>MLE</td>
<td>MinUP</td>
<td>MinUP</td>
<td>$\mu$</td>
<td>MC-BLUP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LM</td>
<td></td>
</tr>
<tr>
<td>0.995</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
<td>0.39</td>
<td>0.42</td>
</tr>
<tr>
<td>0.990</td>
<td>0.11</td>
<td>0.09</td>
<td>0.09</td>
<td>0.56</td>
<td>0.53</td>
</tr>
<tr>
<td>0.975</td>
<td>0.27</td>
<td>0.25</td>
<td>0.27</td>
<td>0.72</td>
<td>0.61</td>
</tr>
<tr>
<td>0.950</td>
<td>0.52</td>
<td>0.52</td>
<td>0.52</td>
<td>0.83</td>
<td>0.72</td>
</tr>
<tr>
<td>0.900</td>
<td>0.63</td>
<td>0.70</td>
<td>0.70</td>
<td>0.98</td>
<td>0.81</td>
</tr>
<tr>
<td>0.850</td>
<td>0.70</td>
<td>0.78</td>
<td>0.78</td>
<td>0.98</td>
<td>0.91</td>
</tr>
<tr>
<td>0.750</td>
<td>0.72</td>
<td>0.88</td>
<td>0.89</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>0.500</td>
<td>0.80</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 12: Average and (median) coordinate-wise ranks of the five competing methodologies in the 100 simulation runs for the \((k_d, k_e) = (2, 4)\) parameter design problem.

<table>
<thead>
<tr>
<th>Control Variable</th>
<th>Q-LM</th>
<th>Q-LM</th>
<th>Q-LM</th>
<th>(\mu)-LM</th>
<th>(\mu)-LM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IavAD</td>
<td>IavAD</td>
<td>AMSE(_Q)</td>
<td>AMSE(_\mu)</td>
<td>OA(_2)-LHS</td>
</tr>
<tr>
<td>MLE</td>
<td>MinUP</td>
<td>MinUP</td>
<td>MC-BLUP</td>
<td>MC-BLUP</td>
<td>MC-BLUP</td>
</tr>
<tr>
<td>(x_{d_1}^*)</td>
<td>4.03 (4.00)</td>
<td>3.45 (3.50)</td>
<td>3.77 (3.75)</td>
<td>1.75 (1.00)</td>
<td>2.01 (2.00)</td>
</tr>
<tr>
<td>(x_{d_2}^*)</td>
<td>3.42 (3.50)</td>
<td>3.01 (3.00)</td>
<td>2.66 (2.50)</td>
<td>2.82 (2.50)</td>
<td>3.10 (3.00)</td>
</tr>
</tbody>
</table>

For each coordinate \(x_{d_i}^*, i \in \{1, 2\}\), we ranked the five methodologies by the values of the absolute distance, \(|\hat{x}_{d_i}^* - x_{d_i}^*|\) within each run. For each methodology, the mean and median of the 100 ranks for each coordinate are given in Table 12. We notice that at least 50% of the time the \(\mu\)-LM/AMSE\(_\mu\)/MC-BLUP methodology does the best job estimating \(x_{d_1}^*\), while all the methodologies struggle somewhat equally when estimating \(x_{d_2}^*\). This suggests that the \(\mu\)-LM model detects slope changes in the surface better than the Q-LM model. However, if the change is subtle, both models tend to oversmooth the surface, and poor estimates of \(x_d^*\) are obtained.

4.4 The \((k_d, k_e) = (4, 5)\) Problem

In this section we present a comparison of the five competing methodologies using 64 point experimental designs in a \((k_d, k_e) = (4, 5)\) parameter design problem. The
The response function for this simulation experiment is

\[ y(x_d, x_e | x_d^*) = 8.35(x_{d4} - a)(x_{d4} - x_{d4}^*)^2 + \left\{ \left(2.75(x_{d1} - x_{d1}^*)^2 + x_{d1}x_{e1} + \sqrt{27} \right) \left(4.5 + 2\pi(x_{d2} - x_{d2}^*)\sin\left(\frac{\pi}{2}(x_{d2} - x_{d2}^*)\right)\right) \right\} + 1.5e^{3.5|x_{d3} - x_{d3}^*|^{1.25}} - \\
3.2x_{e4}e^{-x_{d2}} - .72x_{e3} + 4x_{d3}x_{e5} + \pi x_{e2}\sin(x_{d4}). \]  

(4.4.1)

where \( x_d \in [-1,1]^4, x_e \in [-1,1]^5, x_d^* \in [-1,1]^4 \) and \(-2 \leq a \leq -1\). The expected integrated loss function equals

\[ L(x_d|x_d^*) = \left(2.75(x_{d1} - x_{d1}^*)^2 + \sqrt{27}\right) \left(2\pi(x_{d2} - x_{d2}^*)\sin\left(\frac{\pi}{2}(x_{d2} - x_{d2}^*)\right) + 4.5\right) + \\
8.35(x_{d4} - a)(x_{d4} - x_{d4}^*)^2 + 1.5e^{3.5|x_{d3} - x_{d3}^*|^{1.25}}. \]  

(4.4.2)

The loss function in (4.4.2) is not differentiable with respect to \( x_{d3} \) at \( x_{d3}^* \), however, it is still uniquely minimized at the point \( x_d^* \). Figures 7 and 8 show the \( \binom{k_d}{2} = \binom{4}{2} = 6 \) perspective plots of \( L(\cdot|x_d^*) \) as a function of each bivariate pair in \( x_d \) with the remaining variables in \( x_d \) fixed at their levels in \( x_d^* \). These pictures suggest that \( x_d^* \) is easiest to estimate when it is near the origin.

For the simulation experiment, we partition the margin \([-1,1]\) for each \( x_d \) variable into five equal subintervals. This partitions the whole space \( X_d \) into \( 5^4 = 625 \) equal cuboidal regions. Sampling once from each of these regions produces a computationally expensive simulation experiment; therefore, we use the software of Owen, based on the algorithm of Bose and Bush (1952), to generate a 125 point strength three orthogonal array with five levels for each variable in \( x_d \). This orthogonal array determines 125 cuboidal regions in such a way that the selected regions are uniformly distributed across each univariate, bivariate, and trivariate margin. We randomly
sample one point from each of the 125 selected cuboidal regions to be the true optimal levels \( x^*_d \) for one run of the computer code.

The proportion of the 125 simulation runs that achieved at least the given loss efficiency levels for each of the methodologies is given in Table 13. Table 13 clearly shows that the \( \mu\text{-LM}/OA_2\text{-LHS}/MC\text{-BLUP} \) methodology performs the best with the \( \mu\text{-LM}/AMSE_\mu/MC\text{-BLUP} \) methodology a close second. The methodology that was developed in this dissertation performs considerably worse, but again (4.4.1) may not be approximated very well by a second order model. Also note that the \( Q\text{-LM}/IavAD/\text{MinUP} \) and \( Q\text{-LM}/AMSE_Q/\text{MinUP} \) methodologies perform equally well again, so no conclusion can be drawn about whether design strategies that explicitly consider the estimation of \( x^*_d \) perform better in the parameter design problem.

Table 13: Proportion of the 125 simulation runs greater than or equal to the given loss efficiency levels for the \((k_d, k_e) = (4,5)\) parameter design problem. Highlighted values in the table designate the optimal methodology or methodologies.

<table>
<thead>
<tr>
<th>Loss Efficiency Level</th>
<th>Q-LM IavAD MLE</th>
<th>Q-LM IavAD MinUP</th>
<th>Q-LM AMSE ( \mu ) MinUP</th>
<th>( \mu \text{-LM AMSE}_\mu ) MC-BLUP</th>
<th>( \mu \text{-LM OA}_2\text{-LHS} ) MC-BLUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.950</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.925</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>0.900</td>
<td>0.01</td>
<td>\textbf{0.02}</td>
<td>\textbf{0.02}</td>
<td>0.01</td>
<td>\textbf{0.02}</td>
</tr>
<tr>
<td>0.850</td>
<td>0.02</td>
<td>0.02</td>
<td>0.04</td>
<td>0.09</td>
<td>0.15</td>
</tr>
<tr>
<td>0.800</td>
<td>0.04</td>
<td>0.06</td>
<td>0.11</td>
<td>0.22</td>
<td>\textbf{0.33}</td>
</tr>
<tr>
<td>0.750</td>
<td>0.11</td>
<td>0.18</td>
<td>0.22</td>
<td>0.40</td>
<td>0.54</td>
</tr>
<tr>
<td>0.600</td>
<td>0.46</td>
<td>0.82</td>
<td>0.81</td>
<td>0.82</td>
<td>\textbf{0.89}</td>
</tr>
<tr>
<td>0.500</td>
<td>0.58</td>
<td>\textbf{1.00}</td>
<td>0.99</td>
<td>0.99</td>
<td>0.98</td>
</tr>
</tbody>
</table>
Comparing Table 13 with Table 10, we notice that the performance of every methodology has significantly decreased as $k_d$ increased from two to four. The degradation in performance can be partially explained by examining the 125 true optimal operating conditions $x_d^*$. In Section 4.3, 64% of the true $x_d^*$'s where in the subspace $[-0.8, 0.8]^2$; however, in this simulation experiment only $51/125 = 40.8\%$ of the true $x_d^*$'s were in the subspace $[-0.8, 0.8]^4$. Table 14 presents the proportion of loss efficiencies that achieved the given levels for each of the competing methodologies using only the 51 runs of the simulation for which $x_d^* \in [-0.8, 0.8]^4$. We see slight improvement in performance for each of the methodologies on the interior of the space. The methodologies that use the $\mu$-LM model are still without question the best.

<table>
<thead>
<tr>
<th>Loss Efficiency Level</th>
<th>Q-LM</th>
<th>Q-LM</th>
<th>Q-LM</th>
<th>Q-LM</th>
<th>$\mu$-LM</th>
<th>$\mu$-LM</th>
<th>$\mu$-LM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IavAD</td>
<td>IavAD</td>
<td>AMSE$Q$</td>
<td>MinUP</td>
<td>AMSE$\mu$</td>
<td>MC-BLUP</td>
<td>OA$_2$-LHS</td>
</tr>
<tr>
<td>0.950</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.925</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>0.900</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>0.02</td>
<td>0.12</td>
<td>0.29</td>
<td>0.57</td>
</tr>
<tr>
<td>0.850</td>
<td>0.04</td>
<td>0.02</td>
<td>0.00</td>
<td>0.12</td>
<td>0.33</td>
<td>0.55</td>
<td>0.75</td>
</tr>
<tr>
<td>0.800</td>
<td>0.08</td>
<td>0.06</td>
<td>0.08</td>
<td>0.33</td>
<td>0.55</td>
<td>0.75</td>
<td>0.92</td>
</tr>
<tr>
<td>0.750</td>
<td>0.16</td>
<td>0.18</td>
<td>0.22</td>
<td>0.55</td>
<td>0.75</td>
<td>0.92</td>
<td>1.00</td>
</tr>
<tr>
<td>0.600</td>
<td>0.57</td>
<td>0.78</td>
<td>0.73</td>
<td>0.82</td>
<td>0.92</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>0.500</td>
<td>0.75</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 15: Average and (median) coordinate-wise ranks of the five competing methodologies in the 125 simulation runs for the \((k_d, k_r) = (4, 5)\) parameter design problem.

<table>
<thead>
<tr>
<th>Control Variable</th>
<th>Q-LM</th>
<th>Q-LM</th>
<th>Q-LM</th>
<th>(\mu)-LM</th>
<th>(\mu)-LM</th>
</tr>
</thead>
<tbody>
<tr>
<td>IavAD MLE</td>
<td>IavAD MinUP</td>
<td>AMSE(_Q)</td>
<td>AMSE(_\mu)</td>
<td>OA(_2)-LHS</td>
<td></td>
</tr>
<tr>
<td>(x_{d_1}^*)</td>
<td>2.00 (2.00)</td>
<td>2.24 (2.00)</td>
<td>2.84 (3.00)</td>
<td>4.18 (5.00)</td>
<td>3.74 (4.00)</td>
</tr>
<tr>
<td>(x_{d_2}^*)</td>
<td>4.13 (4.50)</td>
<td>3.30 (3.50)</td>
<td>3.35 (3.50)</td>
<td>2.16 (2.00)</td>
<td>2.05 (2.00)</td>
</tr>
<tr>
<td>(x_{d_3}^*)</td>
<td>3.20 (3.00)</td>
<td>3.30 (3.00)</td>
<td>3.32 (3.00)</td>
<td>2.73 (3.00)</td>
<td>2.45 (2.50)</td>
</tr>
<tr>
<td>(x_{d_4}^*)</td>
<td>3.24 (3.00)</td>
<td>2.83 (3.00)</td>
<td>2.95 (3.00)</td>
<td>2.78 (2.50)</td>
<td>3.20 (3.50)</td>
</tr>
</tbody>
</table>

For each coordinate \(x_{d_i}^*, i \in \{1, \ldots, 4\}\), we ranked the five methodologies by the values of the absolute distance \(|\hat{x}_{d_i}^* - x_{d_i}^*|\) within each run. The mean and median of the 125 ranks for each methodology and coordinate are given in Table 15. We notice the two methodologies that use the IavAD design criterion do the best job estimating the coordinate \(x_{d_i}^*\), but the \(\mu\)-LM model based methodologies/AMSE\(_\mu\)/MC-BLUP estimate the remaining coordinates better. Once again, we can attribute some of the success the \(\mu\)-LM methodologies to using a response function (4.4.1) that the \(\mu\)-LM model is known to predict well (Welch et al 1992), but the complete domination of these methodologies as seen in Tables 13 and 14 indicate that they are the methods of choice for the models of the form (4.4.1).
Figure 7: Perspective plots of the true integrated loss function $L(\cdot | x^*_d)$ for the $(k_d, k_e) = (4, 5)$ parameter design problem as a function of each pair in $x_d$ with the remaining variables in $x_d$ fixed at their levels in $x^*_d$. The true optimal operating conditions are $x^*_d = 0_4$. 
Figure 8: Perspective plots of the true integrated loss function $L(\cdot|x_d^*)$ for the $(k_d, k_e) = (4, 5)$ parameter design problem as a function of each pair in $x_d$ with the remaining variables in $x_d^*$ fixed at their levels in $x_d^*$. The true optimal operating conditions are $x_d^* = (-0.5, 0.75, 0.5, -0.75)'$. 
Table 16: The 25 point AMSE$_Q$ optimal design on $[-1, 1]^6$ that is used in the simulation experiment for the $(k_d, k_c) = (2, 4)$ parameter design problem.

<table>
<thead>
<tr>
<th>$x_{d_1}$</th>
<th>$x_{d_2}$</th>
<th>$x_{c_1}$</th>
<th>$x_{c_2}$</th>
<th>$x_{c_3}$</th>
<th>$x_{c_4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>-0.89</td>
<td>-0.97</td>
<td>-0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>0.40</td>
<td>0.97</td>
<td>0.97</td>
<td>0.64</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>-0.72</td>
<td>-0.71</td>
<td>-0.15</td>
<td>0.81</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>-0.89</td>
<td>-0.97</td>
<td>-0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>0.40</td>
<td>0.97</td>
<td>0.97</td>
<td>0.64</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>0.42</td>
<td>-0.75</td>
<td>-0.05</td>
<td>0.87</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>-0.84</td>
<td>0.08</td>
<td>0.61</td>
<td>-0.47</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>0.23</td>
<td>0.77</td>
<td>0.86</td>
<td>0.92</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>-0.84</td>
<td>-0.08</td>
<td>0.61</td>
<td>-0.47</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>0.19</td>
<td>0.89</td>
<td>-0.96</td>
<td>-0.88</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-0.36</td>
<td>-0.42</td>
<td>-0.67</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>-0.87</td>
<td>0.52</td>
<td>0.02</td>
<td>0.98</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-0.36</td>
<td>-0.42</td>
<td>-0.67</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>0.19</td>
<td>0.89</td>
<td>-0.96</td>
<td>-0.88</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>-0.87</td>
<td>0.52</td>
<td>0.02</td>
<td>0.98</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>0.29</td>
<td>0.68</td>
<td>-0.97</td>
<td>-0.03</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>-0.05</td>
<td>0.59</td>
<td>0.95</td>
<td>-0.67</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>0.49</td>
<td>-0.92</td>
<td>-0.64</td>
<td>0.51</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>-0.05</td>
<td>0.59</td>
<td>0.95</td>
<td>-0.67</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>0.49</td>
<td>-0.92</td>
<td>-0.64</td>
<td>0.51</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>0.39</td>
<td>-0.35</td>
<td>0.93</td>
<td>0.84</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>0.77</td>
<td>-0.51</td>
<td>0.72</td>
<td>0.15</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>-0.84</td>
<td>0.56</td>
<td>-0.99</td>
<td>-0.08</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>0.77</td>
<td>-0.51</td>
<td>0.72</td>
<td>0.15</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>-0.60</td>
<td>-0.68</td>
<td>-0.62</td>
<td>-0.83</td>
</tr>
</tbody>
</table>
Table 17: The 64 point AMSE\(_Q\) optimal design on \([-1, 1]^9\) that is used in the simulation experiment for the \((k_d, k_e) = (4, 5)\) parameter design problem.

<table>
<thead>
<tr>
<th>(x_{d1})</th>
<th>(x_{d2})</th>
<th>(x_{d3})</th>
<th>(x_{d4})</th>
<th>(x_{e1})</th>
<th>(x_{e2})</th>
<th>(x_{e3})</th>
<th>(x_{e4})</th>
<th>(x_{e5})</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.60</td>
<td>-0.08</td>
<td>-0.86</td>
<td>-0.70</td>
<td>-0.99</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.60</td>
<td>-0.08</td>
<td>-0.86</td>
<td>-0.70</td>
<td>-0.99</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>0.60</td>
<td>-0.08</td>
<td>-0.86</td>
<td>-0.70</td>
<td>-0.99</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>-0.49</td>
<td>-0.93</td>
<td>-0.15</td>
<td>0.82</td>
<td>-0.70</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>0.12</td>
<td>0.87</td>
<td>0.35</td>
<td>0.60</td>
<td>0.57</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>0.55</td>
<td>0.62</td>
<td>-0.55</td>
<td>-0.91</td>
<td>0.39</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.60</td>
<td>-0.08</td>
<td>-0.86</td>
<td>-0.70</td>
<td>-0.99</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.60</td>
<td>-0.08</td>
<td>-0.86</td>
<td>-0.70</td>
<td>-0.99</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.78</td>
<td>0.74</td>
<td>0.88</td>
<td>-0.69</td>
<td>0.19</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.41</td>
<td>0.98</td>
<td>-0.63</td>
<td>-0.13</td>
<td>0.89</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.63</td>
<td>-0.82</td>
<td>-0.94</td>
<td>-0.37</td>
<td>-0.60</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.78</td>
<td>0.74</td>
<td>0.88</td>
<td>-0.69</td>
<td>0.19</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>0.41</td>
<td>0.98</td>
<td>-0.63</td>
<td>-0.13</td>
<td>0.89</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>-0.91</td>
<td>0.31</td>
<td>0.09</td>
<td>0.56</td>
<td>-0.90</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>0.78</td>
<td>0.74</td>
<td>0.88</td>
<td>-0.69</td>
<td>0.19</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>-0.91</td>
<td>0.31</td>
<td>0.09</td>
<td>0.56</td>
<td>-0.90</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-0.36</td>
<td>-0.76</td>
<td>-0.27</td>
<td>-0.55</td>
<td>0.69</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.41</td>
<td>0.98</td>
<td>-0.63</td>
<td>-0.13</td>
<td>0.89</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>-0.91</td>
<td>0.31</td>
<td>0.09</td>
<td>0.56</td>
<td>-0.90</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.41</td>
<td>0.98</td>
<td>-0.63</td>
<td>-0.13</td>
<td>0.89</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>-0.36</td>
<td>-0.76</td>
<td>-0.27</td>
<td>-0.55</td>
<td>0.69</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-0.36</td>
<td>-0.76</td>
<td>-0.27</td>
<td>-0.55</td>
<td>0.69</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.78</td>
<td>0.74</td>
<td>0.88</td>
<td>-0.69</td>
<td>0.19</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-0.63</td>
<td>-0.82</td>
<td>-0.94</td>
<td>-0.37</td>
<td>-0.60</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>-0.63</td>
<td>-0.82</td>
<td>-0.94</td>
<td>-0.37</td>
<td>-0.60</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>0.41</td>
<td>0.98</td>
<td>-0.63</td>
<td>-0.13</td>
<td>0.89</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.94</td>
<td>-0.29</td>
<td>0.26</td>
<td>-0.97</td>
<td>0.34</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>-0.63</td>
<td>-0.82</td>
<td>-0.94</td>
<td>-0.37</td>
<td>-0.60</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.41</td>
<td>0.98</td>
<td>-0.63</td>
<td>-0.13</td>
<td>0.89</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>0.51</td>
<td>-0.64</td>
<td>-0.63</td>
<td>0.50</td>
<td>0.78</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-0.62</td>
<td>0.45</td>
<td>-0.14</td>
<td>-0.72</td>
<td>-0.29</td>
</tr>
</tbody>
</table>
Table 17 (continued)

<table>
<thead>
<tr>
<th>$x_{d_1}$</th>
<th>$x_{d_2}$</th>
<th>$x_{d_3}$</th>
<th>$x_{d_4}$</th>
<th>$x_{c_1}$</th>
<th>$x_{c_2}$</th>
<th>$x_{c_3}$</th>
<th>$x_{c_4}$</th>
<th>$x_{c_5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.17</td>
<td>0.02</td>
<td>0.29</td>
<td>-0.99</td>
<td>0.50</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-0.43</td>
<td>-0.50</td>
<td>-0.91</td>
<td>0.78</td>
<td>-0.10</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.85</td>
<td>-0.10</td>
<td>0.92</td>
<td>-0.58</td>
<td>-0.84</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.85</td>
<td>-0.10</td>
<td>0.92</td>
<td>-0.58</td>
<td>-0.84</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-0.83</td>
<td>0.82</td>
<td>0.49</td>
<td>0.17</td>
<td>-0.43</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-0.77</td>
<td>0.97</td>
<td>-0.83</td>
<td>0.97</td>
<td>-0.68</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-0.99</td>
<td>-0.24</td>
<td>0.90</td>
<td>0.13</td>
<td>0.43</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-0.77</td>
<td>0.97</td>
<td>-0.83</td>
<td>0.97</td>
<td>-0.68</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>-0.99</td>
<td>-0.24</td>
<td>0.90</td>
<td>0.13</td>
<td>0.43</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>-0.27</td>
<td>-0.63</td>
<td>-0.79</td>
<td>0.41</td>
<td>0.18</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-0.99</td>
<td>-0.24</td>
<td>0.90</td>
<td>0.13</td>
<td>0.43</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>-0.77</td>
<td>0.97</td>
<td>-0.83</td>
<td>0.97</td>
<td>-0.68</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.62</td>
<td>0.19</td>
<td>0.36</td>
<td>-0.44</td>
<td>-0.93</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-0.77</td>
<td>0.97</td>
<td>-0.83</td>
<td>0.97</td>
<td>-0.68</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.56</td>
<td>-0.97</td>
<td>0.67</td>
<td>-0.85</td>
<td>0.69</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>-0.99</td>
<td>-0.24</td>
<td>0.90</td>
<td>0.13</td>
<td>0.43</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-0.77</td>
<td>0.97</td>
<td>-0.83</td>
<td>0.97</td>
<td>-0.68</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-0.99</td>
<td>-0.24</td>
<td>0.90</td>
<td>0.13</td>
<td>0.43</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.60</td>
<td>-0.92</td>
<td>-0.11</td>
<td>0.96</td>
<td>-0.29</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>-0.97</td>
<td>-0.49</td>
<td>0.38</td>
<td>-0.82</td>
<td>0.51</td>
</tr>
<tr>
<td>-1.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.94</td>
<td>-0.14</td>
<td>0.86</td>
<td>0.70</td>
<td>0.85</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.78</td>
<td>0.59</td>
<td>0.55</td>
<td>0.57</td>
<td>-0.92</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.94</td>
<td>-0.14</td>
<td>0.86</td>
<td>0.70</td>
<td>0.85</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.60</td>
<td>-0.92</td>
<td>-0.11</td>
<td>0.96</td>
<td>-0.29</td>
</tr>
<tr>
<td>0.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.78</td>
<td>0.59</td>
<td>0.55</td>
<td>0.57</td>
<td>-0.92</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>0.60</td>
<td>-0.92</td>
<td>-0.11</td>
<td>0.96</td>
<td>-0.29</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>0.60</td>
<td>-0.92</td>
<td>-0.11</td>
<td>0.96</td>
<td>-0.29</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>-0.97</td>
<td>-0.49</td>
<td>0.38</td>
<td>-0.82</td>
<td>0.51</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>-1.00</td>
<td>1.00</td>
<td>0.94</td>
<td>-0.14</td>
<td>0.86</td>
<td>0.70</td>
<td>0.85</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.78</td>
<td>0.59</td>
<td>0.55</td>
<td>0.57</td>
<td>-0.92</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>0.94</td>
<td>-0.14</td>
<td>0.86</td>
<td>0.70</td>
<td>0.85</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>-1.00</td>
<td>-1.00</td>
<td>0.78</td>
<td>0.59</td>
<td>0.55</td>
<td>0.57</td>
<td>-0.92</td>
</tr>
<tr>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.60</td>
<td>-0.92</td>
<td>-0.11</td>
<td>0.96</td>
<td>-0.29</td>
</tr>
</tbody>
</table>
Table 18: The 25 point AMSE$_\mu$ optimal design on $[-1, 1]^6$ that is used in the simulation experiment for the $(k_d, k_c) = (2, 4)$ parameter design problem.

<table>
<thead>
<tr>
<th>$x_{d1}$</th>
<th>$x_{d2}$</th>
<th>$x_{e1}$</th>
<th>$x_{e2}$</th>
<th>$x_{e3}$</th>
<th>$x_{e4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.30</td>
<td>0.76</td>
<td>-0.46</td>
<td>-0.71</td>
<td>0.65</td>
<td>-0.85</td>
</tr>
<tr>
<td>-0.75</td>
<td>-0.41</td>
<td>-0.28</td>
<td>0.81</td>
<td>0.47</td>
<td>0.31</td>
</tr>
<tr>
<td>-0.47</td>
<td>0.89</td>
<td>0.31</td>
<td>-0.61</td>
<td>0.13</td>
<td>-0.58</td>
</tr>
<tr>
<td>0.15</td>
<td>-0.02</td>
<td>0.62</td>
<td>0.02</td>
<td>-0.90</td>
<td>0.64</td>
</tr>
<tr>
<td>-0.16</td>
<td>-0.39</td>
<td>-0.75</td>
<td>0.78</td>
<td>-0.73</td>
<td>-0.20</td>
</tr>
<tr>
<td>-0.96</td>
<td>-0.74</td>
<td>-0.45</td>
<td>-0.52</td>
<td>-0.55</td>
<td>0.81</td>
</tr>
<tr>
<td>0.90</td>
<td>0.90</td>
<td>-0.38</td>
<td>-0.69</td>
<td>-0.93</td>
<td>-0.78</td>
</tr>
<tr>
<td>0.86</td>
<td>0.87</td>
<td>0.60</td>
<td>0.55</td>
<td>0.64</td>
<td>-0.01</td>
</tr>
<tr>
<td>0.40</td>
<td>0.18</td>
<td>-0.48</td>
<td>-0.08</td>
<td>-0.78</td>
<td>-0.67</td>
</tr>
<tr>
<td>0.53</td>
<td>0.66</td>
<td>-0.10</td>
<td>-0.12</td>
<td>0.13</td>
<td>-0.98</td>
</tr>
<tr>
<td>0.38</td>
<td>-0.18</td>
<td>0.40</td>
<td>0.00</td>
<td>0.69</td>
<td>0.11</td>
</tr>
<tr>
<td>-0.74</td>
<td>0.00</td>
<td>-0.74</td>
<td>0.27</td>
<td>0.56</td>
<td>0.71</td>
</tr>
<tr>
<td>0.75</td>
<td>-0.84</td>
<td>0.43</td>
<td>0.19</td>
<td>0.37</td>
<td>-0.47</td>
</tr>
<tr>
<td>0.57</td>
<td>-0.34</td>
<td>0.99</td>
<td>0.50</td>
<td>-0.24</td>
<td>-0.32</td>
</tr>
<tr>
<td>0.18</td>
<td>0.07</td>
<td>-0.51</td>
<td>-0.49</td>
<td>0.24</td>
<td>-0.28</td>
</tr>
<tr>
<td>0.45</td>
<td>-0.75</td>
<td>0.54</td>
<td>0.36</td>
<td>-0.68</td>
<td>-0.04</td>
</tr>
<tr>
<td>-0.93</td>
<td>0.18</td>
<td>0.09</td>
<td>0.85</td>
<td>0.23</td>
<td>-0.31</td>
</tr>
<tr>
<td>0.93</td>
<td>-0.73</td>
<td>-0.61</td>
<td>-0.97</td>
<td>0.66</td>
<td>0.89</td>
</tr>
<tr>
<td>0.86</td>
<td>-0.68</td>
<td>0.87</td>
<td>0.42</td>
<td>0.41</td>
<td>0.82</td>
</tr>
<tr>
<td>-0.03</td>
<td>-0.20</td>
<td>-0.63</td>
<td>-0.09</td>
<td>-0.80</td>
<td>-0.78</td>
</tr>
<tr>
<td>0.60</td>
<td>-0.03</td>
<td>-0.92</td>
<td>0.47</td>
<td>0.15</td>
<td>0.98</td>
</tr>
<tr>
<td>0.38</td>
<td>0.09</td>
<td>-0.84</td>
<td>0.76</td>
<td>-0.78</td>
<td>0.09</td>
</tr>
<tr>
<td>-0.48</td>
<td>-0.86</td>
<td>0.51</td>
<td>-0.65</td>
<td>-0.09</td>
<td>-0.25</td>
</tr>
<tr>
<td>0.90</td>
<td>-0.77</td>
<td>0.98</td>
<td>-0.50</td>
<td>-0.29</td>
<td>0.82</td>
</tr>
<tr>
<td>-0.80</td>
<td>-0.61</td>
<td>0.76</td>
<td>-0.13</td>
<td>0.09</td>
<td>-0.09</td>
</tr>
</tbody>
</table>
Table 19: The 64 point AMSEμ optimal design on [-1,1]^9 that is used in the simulation experiment for the (kd, kr) = (4, 5) parameter design problem.

<table>
<thead>
<tr>
<th>x_{d1}</th>
<th>x_{d2}</th>
<th>x_{d3}</th>
<th>x_{d4}</th>
<th>x_{r1}</th>
<th>x_{r2}</th>
<th>x_{r3}</th>
<th>x_{r4}</th>
<th>x_{r5}</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.36</td>
<td>0.36</td>
<td>0.31</td>
<td>0.21</td>
<td>-0.84</td>
<td>0.35</td>
<td>-0.39</td>
<td>-0.98</td>
<td>0.39</td>
</tr>
<tr>
<td>0.44</td>
<td>-0.43</td>
<td>-0.37</td>
<td>0.89</td>
<td>0.53</td>
<td>-0.90</td>
<td>-0.96</td>
<td>0.81</td>
<td>0.01</td>
</tr>
<tr>
<td>-0.20</td>
<td>-0.22</td>
<td>-0.53</td>
<td>-0.72</td>
<td>-0.24</td>
<td>-0.96</td>
<td>-0.25</td>
<td>-0.82</td>
<td>-0.82</td>
</tr>
<tr>
<td>0.52</td>
<td>0.90</td>
<td>0.08</td>
<td>0.35</td>
<td>-0.34</td>
<td>-0.66</td>
<td>-0.57</td>
<td>-0.08</td>
<td>0.72</td>
</tr>
<tr>
<td>-0.04</td>
<td>0.92</td>
<td>0.12</td>
<td>-0.37</td>
<td>0.93</td>
<td>0.14</td>
<td>0.46</td>
<td>0.95</td>
<td>0.55</td>
</tr>
<tr>
<td>-0.64</td>
<td>0.28</td>
<td>0.28</td>
<td>-0.45</td>
<td>0.46</td>
<td>0.08</td>
<td>-0.82</td>
<td>0.71</td>
<td>-0.08</td>
</tr>
<tr>
<td>0.21</td>
<td>-0.37</td>
<td>-0.45</td>
<td>-0.52</td>
<td>0.32</td>
<td>0.86</td>
<td>0.75</td>
<td>0.23</td>
<td>0.14</td>
</tr>
<tr>
<td>-0.57</td>
<td>0.42</td>
<td>-0.01</td>
<td>-0.25</td>
<td>-0.17</td>
<td>0.72</td>
<td>-0.63</td>
<td>-0.73</td>
<td>0.68</td>
</tr>
<tr>
<td>0.51</td>
<td>0.25</td>
<td>0.63</td>
<td>-0.23</td>
<td>0.82</td>
<td>-0.45</td>
<td>-0.43</td>
<td>0.83</td>
<td>0.33</td>
</tr>
<tr>
<td>0.93</td>
<td>-0.38</td>
<td>0.27</td>
<td>0.47</td>
<td>0.13</td>
<td>0.28</td>
<td>-0.67</td>
<td>-0.02</td>
<td>-0.37</td>
</tr>
<tr>
<td>-0.94</td>
<td>-0.17</td>
<td>0.82</td>
<td>-0.80</td>
<td>-0.79</td>
<td>-0.32</td>
<td>0.94</td>
<td>0.29</td>
<td>-0.08</td>
</tr>
<tr>
<td>-0.59</td>
<td>-0.53</td>
<td>0.12</td>
<td>0.25</td>
<td>-0.60</td>
<td>0.10</td>
<td>0.45</td>
<td>-0.45</td>
<td>-0.61</td>
</tr>
<tr>
<td>-0.30</td>
<td>0.75</td>
<td>-0.47</td>
<td>-0.61</td>
<td>-0.02</td>
<td>-0.51</td>
<td>0.47</td>
<td>-0.74</td>
<td>0.22</td>
</tr>
<tr>
<td>-0.64</td>
<td>0.19</td>
<td>-0.67</td>
<td>-0.57</td>
<td>0.04</td>
<td>0.34</td>
<td>0.32</td>
<td>-0.70</td>
<td>0.82</td>
</tr>
<tr>
<td>0.08</td>
<td>0.83</td>
<td>0.12</td>
<td>0.99</td>
<td>0.25</td>
<td>0.80</td>
<td>0.79</td>
<td>-0.54</td>
<td>0.11</td>
</tr>
<tr>
<td>-0.57</td>
<td>-0.43</td>
<td>-0.27</td>
<td>-0.29</td>
<td>-0.71</td>
<td>-0.93</td>
<td>-0.91</td>
<td>-0.37</td>
<td>-0.39</td>
</tr>
<tr>
<td>-0.94</td>
<td>-0.34</td>
<td>0.21</td>
<td>-0.08</td>
<td>-0.14</td>
<td>0.19</td>
<td>0.84</td>
<td>0.28</td>
<td>0.01</td>
</tr>
<tr>
<td>0.25</td>
<td>0.60</td>
<td>0.68</td>
<td>0.25</td>
<td>-0.99</td>
<td>0.30</td>
<td>-0.24</td>
<td>-0.09</td>
<td>0.52</td>
</tr>
<tr>
<td>-0.80</td>
<td>0.86</td>
<td>-0.07</td>
<td>-0.26</td>
<td>-0.41</td>
<td>-0.84</td>
<td>-0.91</td>
<td>-0.99</td>
<td>0.22</td>
</tr>
<tr>
<td>0.44</td>
<td>-0.73</td>
<td>0.82</td>
<td>-0.98</td>
<td>-0.11</td>
<td>-0.74</td>
<td>0.49</td>
<td>-0.07</td>
<td>-0.68</td>
</tr>
<tr>
<td>0.36</td>
<td>0.43</td>
<td>0.51</td>
<td>-0.82</td>
<td>0.87</td>
<td>-0.28</td>
<td>0.56</td>
<td>0.23</td>
<td>-0.49</td>
</tr>
<tr>
<td>-0.04</td>
<td>0.62</td>
<td>0.76</td>
<td>-0.72</td>
<td>0.33</td>
<td>0.31</td>
<td>-0.39</td>
<td>-0.11</td>
<td>0.95</td>
</tr>
<tr>
<td>-0.24</td>
<td>0.11</td>
<td>-0.94</td>
<td>0.40</td>
<td>-0.31</td>
<td>-0.51</td>
<td>0.87</td>
<td>0.44</td>
<td>-0.13</td>
</tr>
<tr>
<td>0.01</td>
<td>-0.53</td>
<td>-0.36</td>
<td>0.44</td>
<td>-0.08</td>
<td>0.72</td>
<td>0.79</td>
<td>0.16</td>
<td>0.07</td>
</tr>
<tr>
<td>0.97</td>
<td>-0.10</td>
<td>-0.67</td>
<td>-0.86</td>
<td>0.70</td>
<td>0.78</td>
<td>-0.90</td>
<td>0.79</td>
<td>-0.27</td>
</tr>
<tr>
<td>-0.27</td>
<td>0.11</td>
<td>-0.86</td>
<td>0.74</td>
<td>-0.09</td>
<td>-0.76</td>
<td>0.24</td>
<td>0.70</td>
<td>0.02</td>
</tr>
<tr>
<td>0.45</td>
<td>0.06</td>
<td>0.53</td>
<td>0.56</td>
<td>-0.84</td>
<td>-0.55</td>
<td>-0.46</td>
<td>-0.75</td>
<td>0.54</td>
</tr>
<tr>
<td>0.53</td>
<td>-0.57</td>
<td>-0.28</td>
<td>0.96</td>
<td>-0.14</td>
<td>-0.10</td>
<td>0.73</td>
<td>-0.06</td>
<td>-0.01</td>
</tr>
<tr>
<td>0.96</td>
<td>0.67</td>
<td>-0.43</td>
<td>-0.45</td>
<td>0.08</td>
<td>0.15</td>
<td>0.53</td>
<td>-0.68</td>
<td>-0.29</td>
</tr>
<tr>
<td>-0.37</td>
<td>0.39</td>
<td>0.71</td>
<td>0.52</td>
<td>0.26</td>
<td>0.38</td>
<td>0.37</td>
<td>0.93</td>
<td>-0.95</td>
</tr>
<tr>
<td>0.71</td>
<td>-0.38</td>
<td>0.35</td>
<td>0.60</td>
<td>0.83</td>
<td>0.65</td>
<td>-0.48</td>
<td>0.13</td>
<td>-0.03</td>
</tr>
<tr>
<td>$x_{d_1}$</td>
<td>$x_{d_2}$</td>
<td>$x_{d_3}$</td>
<td>$x_{d_4}$</td>
<td>$x_{e_1}$</td>
<td>$x_{e_2}$</td>
<td>$x_{e_3}$</td>
<td>$x_{e_4}$</td>
<td>$x_{e_5}$</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>0.64</td>
<td>-0.19</td>
<td>-0.58</td>
<td>-0.69</td>
<td>-1.00</td>
<td>-0.91</td>
<td>0.75</td>
<td>0.00</td>
<td>0.38</td>
</tr>
<tr>
<td>-0.46</td>
<td>0.73</td>
<td>-0.69</td>
<td>0.43</td>
<td>0.96</td>
<td>0.42</td>
<td>0.14</td>
<td>0.57</td>
<td>-0.69</td>
</tr>
<tr>
<td>-0.23</td>
<td>-0.04</td>
<td>0.83</td>
<td>0.39</td>
<td>0.74</td>
<td>-0.27</td>
<td>0.88</td>
<td>0.65</td>
<td>0.97</td>
</tr>
<tr>
<td>0.55</td>
<td>0.16</td>
<td>0.21</td>
<td>0.05</td>
<td>-0.02</td>
<td>0.75</td>
<td>-0.82</td>
<td>-0.35</td>
<td>0.12</td>
</tr>
<tr>
<td>-0.01</td>
<td>-0.56</td>
<td>0.94</td>
<td>-0.25</td>
<td>-0.62</td>
<td>0.91</td>
<td>-0.26</td>
<td>0.46</td>
<td>-0.46</td>
</tr>
<tr>
<td>0.33</td>
<td>-0.49</td>
<td>0.44</td>
<td>-0.28</td>
<td>-0.26</td>
<td>-0.73</td>
<td>0.73</td>
<td>-0.61</td>
<td>-0.25</td>
</tr>
<tr>
<td>-0.89</td>
<td>-0.65</td>
<td>-0.77</td>
<td>-0.79</td>
<td>0.64</td>
<td>-0.39</td>
<td>0.45</td>
<td>-0.36</td>
<td>0.81</td>
</tr>
<tr>
<td>-0.97</td>
<td>-0.90</td>
<td>-0.89</td>
<td>-0.99</td>
<td>-0.20</td>
<td>-0.52</td>
<td>-0.77</td>
<td>-0.54</td>
<td>-0.11</td>
</tr>
<tr>
<td>0.78</td>
<td>0.03</td>
<td>-0.47</td>
<td>0.20</td>
<td>0.60</td>
<td>-0.98</td>
<td>-0.03</td>
<td>-0.26</td>
<td>0.69</td>
</tr>
<tr>
<td>-0.59</td>
<td>-0.37</td>
<td>-0.56</td>
<td>-0.45</td>
<td>0.46</td>
<td>-0.83</td>
<td>0.31</td>
<td>0.33</td>
<td>0.35</td>
</tr>
<tr>
<td>0.88</td>
<td>0.16</td>
<td>-0.90</td>
<td>0.73</td>
<td>0.20</td>
<td>-0.08</td>
<td>0.84</td>
<td>0.69</td>
<td>0.95</td>
</tr>
<tr>
<td>-0.44</td>
<td>0.15</td>
<td>-0.14</td>
<td>-0.21</td>
<td>0.50</td>
<td>0.99</td>
<td>-0.60</td>
<td>-0.54</td>
<td>0.59</td>
</tr>
<tr>
<td>-0.22</td>
<td>-0.99</td>
<td>-0.49</td>
<td>-0.48</td>
<td>-0.01</td>
<td>0.87</td>
<td>-0.96</td>
<td>0.04</td>
<td>-0.15</td>
</tr>
<tr>
<td>-0.74</td>
<td>0.94</td>
<td>0.10</td>
<td>-0.56</td>
<td>-0.48</td>
<td>-0.90</td>
<td>-0.44</td>
<td>-0.05</td>
<td>-0.07</td>
</tr>
<tr>
<td>0.15</td>
<td>-0.29</td>
<td>-0.69</td>
<td>0.64</td>
<td>-0.32</td>
<td>-0.22</td>
<td>-0.68</td>
<td>0.42</td>
<td>-0.39</td>
</tr>
<tr>
<td>0.34</td>
<td>0.28</td>
<td>-0.83</td>
<td>-0.68</td>
<td>-0.74</td>
<td>0.88</td>
<td>0.64</td>
<td>-0.36</td>
<td>0.86</td>
</tr>
<tr>
<td>-0.10</td>
<td>0.47</td>
<td>0.59</td>
<td>0.02</td>
<td>-0.42</td>
<td>-0.78</td>
<td>-0.71</td>
<td>0.48</td>
<td>-0.62</td>
</tr>
<tr>
<td>-0.69</td>
<td>0.06</td>
<td>0.94</td>
<td>-0.23</td>
<td>0.00</td>
<td>-1.00</td>
<td>0.87</td>
<td>0.70</td>
<td>-0.31</td>
</tr>
<tr>
<td>-0.23</td>
<td>-0.94</td>
<td>-0.15</td>
<td>0.48</td>
<td>0.45</td>
<td>-0.58</td>
<td>0.95</td>
<td>0.98</td>
<td>0.41</td>
</tr>
<tr>
<td>-0.96</td>
<td>-0.62</td>
<td>-0.78</td>
<td>-0.43</td>
<td>0.06</td>
<td>0.30</td>
<td>-0.83</td>
<td>0.64</td>
<td>0.73</td>
</tr>
<tr>
<td>-0.47</td>
<td>-0.89</td>
<td>-0.32</td>
<td>-0.70</td>
<td>0.63</td>
<td>-0.91</td>
<td>0.77</td>
<td>-0.92</td>
<td>0.89</td>
</tr>
<tr>
<td>-0.82</td>
<td>0.88</td>
<td>-0.46</td>
<td>0.01</td>
<td>-0.18</td>
<td>0.01</td>
<td>-0.95</td>
<td>-0.18</td>
<td>-0.36</td>
</tr>
<tr>
<td>-0.07</td>
<td>0.43</td>
<td>-0.28</td>
<td>0.27</td>
<td>0.50</td>
<td>-0.75</td>
<td>0.80</td>
<td>-0.58</td>
<td>-0.69</td>
</tr>
<tr>
<td>-0.16</td>
<td>-0.28</td>
<td>0.13</td>
<td>0.96</td>
<td>0.75</td>
<td>-0.46</td>
<td>-0.15</td>
<td>0.25</td>
<td>0.58</td>
</tr>
<tr>
<td>0.47</td>
<td>-0.04</td>
<td>0.42</td>
<td>-0.95</td>
<td>0.15</td>
<td>0.53</td>
<td>0.02</td>
<td>0.15</td>
<td>-0.18</td>
</tr>
<tr>
<td>0.98</td>
<td>0.58</td>
<td>-0.63</td>
<td>-0.30</td>
<td>-0.35</td>
<td>-0.33</td>
<td>-0.67</td>
<td>0.73</td>
<td>-0.41</td>
</tr>
<tr>
<td>-0.62</td>
<td>0.37</td>
<td>0.75</td>
<td>0.62</td>
<td>0.89</td>
<td>-0.19</td>
<td>-0.22</td>
<td>0.08</td>
<td>0.91</td>
</tr>
<tr>
<td>0.74</td>
<td>0.60</td>
<td>0.34</td>
<td>-0.15</td>
<td>0.92</td>
<td>0.72</td>
<td>0.38</td>
<td>-0.85</td>
<td>0.18</td>
</tr>
<tr>
<td>0.36</td>
<td>-0.33</td>
<td>0.64</td>
<td>0.19</td>
<td>-0.94</td>
<td>-0.66</td>
<td>-0.35</td>
<td>0.53</td>
<td>-0.79</td>
</tr>
<tr>
<td>-0.57</td>
<td>-0.80</td>
<td>0.98</td>
<td>-0.23</td>
<td>-0.20</td>
<td>0.13</td>
<td>-0.06</td>
<td>-0.62</td>
<td>0.30</td>
</tr>
<tr>
<td>0.57</td>
<td>0.97</td>
<td>-0.85</td>
<td>0.54</td>
<td>-0.44</td>
<td>0.41</td>
<td>0.55</td>
<td>0.50</td>
<td>0.21</td>
</tr>
<tr>
<td>0.18</td>
<td>0.05</td>
<td>-0.53</td>
<td>0.32</td>
<td>-0.03</td>
<td>-0.81</td>
<td>0.49</td>
<td>-0.01</td>
<td>0.09</td>
</tr>
<tr>
<td>0.67</td>
<td>0.25</td>
<td>0.24</td>
<td>-0.55</td>
<td>-0.54</td>
<td>0.35</td>
<td>0.18</td>
<td>-0.22</td>
<td>-0.94</td>
</tr>
</tbody>
</table>
Table 20: The 25 point strength two orthogonal array based Latin hypercube design on \([-1, 1]^6\) that is used in the simulation experiment for the \((k_d, k_e) = (2, 4)\) parameter design problem.

<table>
<thead>
<tr>
<th>(x_{d1})</th>
<th>(x_{d2})</th>
<th>(x_{e1})</th>
<th>(x_{e2})</th>
<th>(x_{e3})</th>
<th>(x_{e4})</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.13</td>
<td>0.98</td>
<td>-0.90</td>
<td>-0.42</td>
<td>-0.85</td>
<td>-0.94</td>
</tr>
<tr>
<td>-0.13</td>
<td>0.33</td>
<td>0.77</td>
<td>0.49</td>
<td>0.15</td>
<td>-0.43</td>
</tr>
<tr>
<td>-0.13</td>
<td>0.13</td>
<td>-0.58</td>
<td>-0.97</td>
<td>0.57</td>
<td>0.54</td>
</tr>
<tr>
<td>-0.13</td>
<td>-0.71</td>
<td>0.50</td>
<td>0.72</td>
<td>-0.31</td>
<td>0.68</td>
</tr>
<tr>
<td>-0.13</td>
<td>-0.43</td>
<td>-0.12</td>
<td>-0.01</td>
<td>0.96</td>
<td>0.19</td>
</tr>
<tr>
<td>-0.39</td>
<td>0.98</td>
<td>0.77</td>
<td>-0.97</td>
<td>-0.31</td>
<td>0.19</td>
</tr>
<tr>
<td>-0.39</td>
<td>0.33</td>
<td>-0.58</td>
<td>0.72</td>
<td>0.96</td>
<td>-0.94</td>
</tr>
<tr>
<td>-0.39</td>
<td>0.13</td>
<td>0.50</td>
<td>-0.01</td>
<td>-0.85</td>
<td>-0.43</td>
</tr>
<tr>
<td>-0.39</td>
<td>-0.71</td>
<td>-0.12</td>
<td>-0.42</td>
<td>0.15</td>
<td>0.54</td>
</tr>
<tr>
<td>-0.39</td>
<td>-0.43</td>
<td>-0.90</td>
<td>0.49</td>
<td>0.57</td>
<td>0.68</td>
</tr>
<tr>
<td>0.37</td>
<td>0.98</td>
<td>-0.58</td>
<td>-0.01</td>
<td>0.15</td>
<td>0.68</td>
</tr>
<tr>
<td>0.37</td>
<td>0.33</td>
<td>0.50</td>
<td>-0.42</td>
<td>0.57</td>
<td>0.19</td>
</tr>
<tr>
<td>0.37</td>
<td>0.13</td>
<td>-0.12</td>
<td>0.49</td>
<td>-0.31</td>
<td>-0.94</td>
</tr>
<tr>
<td>0.37</td>
<td>-0.71</td>
<td>-0.90</td>
<td>-0.97</td>
<td>0.96</td>
<td>-0.43</td>
</tr>
<tr>
<td>0.37</td>
<td>-0.43</td>
<td>0.77</td>
<td>0.72</td>
<td>-0.85</td>
<td>0.54</td>
</tr>
<tr>
<td>0.87</td>
<td>0.98</td>
<td>0.50</td>
<td>0.49</td>
<td>0.96</td>
<td>0.54</td>
</tr>
<tr>
<td>0.87</td>
<td>0.33</td>
<td>-0.12</td>
<td>-0.97</td>
<td>-0.85</td>
<td>0.68</td>
</tr>
<tr>
<td>0.87</td>
<td>0.13</td>
<td>-0.90</td>
<td>0.72</td>
<td>0.15</td>
<td>0.19</td>
</tr>
<tr>
<td>0.87</td>
<td>-0.71</td>
<td>0.77</td>
<td>-0.01</td>
<td>0.57</td>
<td>-0.94</td>
</tr>
<tr>
<td>0.87</td>
<td>-0.43</td>
<td>-0.58</td>
<td>-0.42</td>
<td>-0.31</td>
<td>-0.43</td>
</tr>
<tr>
<td>-0.86</td>
<td>0.98</td>
<td>-0.12</td>
<td>0.72</td>
<td>0.57</td>
<td>-0.43</td>
</tr>
<tr>
<td>-0.86</td>
<td>0.33</td>
<td>-0.90</td>
<td>-0.01</td>
<td>-0.31</td>
<td>0.54</td>
</tr>
<tr>
<td>-0.86</td>
<td>0.13</td>
<td>0.77</td>
<td>-0.42</td>
<td>0.96</td>
<td>0.68</td>
</tr>
<tr>
<td>-0.86</td>
<td>-0.71</td>
<td>-0.58</td>
<td>0.49</td>
<td>-0.85</td>
<td>0.19</td>
</tr>
<tr>
<td>-0.86</td>
<td>-0.43</td>
<td>0.50</td>
<td>-0.97</td>
<td>0.15</td>
<td>-0.94</td>
</tr>
</tbody>
</table>
Table 21: The 64 point OA₂-based Latin hypercube design on \([-1, 1]^9\) that is used in the simulation experiment for the \((k_d, k_e) = (4, 5)\) parameter design problem.

<table>
<thead>
<tr>
<th>(x_{d1})</th>
<th>(x_{d2})</th>
<th>(x_{d3})</th>
<th>(x_{d4})</th>
<th>(x_{e1})</th>
<th>(x_{e2})</th>
<th>(x_{e3})</th>
<th>(x_{e4})</th>
<th>(x_{e5})</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.31</td>
<td>0.65</td>
<td>0.11</td>
<td>0.34</td>
<td>-0.91</td>
<td>-0.85</td>
<td>-0.08</td>
<td>0.65</td>
<td>-0.09</td>
</tr>
<tr>
<td>-0.31</td>
<td>-0.45</td>
<td>0.88</td>
<td>0.65</td>
<td>-0.56</td>
<td>-0.08</td>
<td>0.15</td>
<td>0.76</td>
<td>0.20</td>
</tr>
<tr>
<td>-0.31</td>
<td>-0.85</td>
<td>-0.35</td>
<td>-0.36</td>
<td>0.57</td>
<td>0.76</td>
<td>-0.92</td>
<td>0.33</td>
<td>-0.53</td>
</tr>
<tr>
<td>-0.31</td>
<td>0.42</td>
<td>-0.13</td>
<td>0.18</td>
<td>0.13</td>
<td>0.39</td>
<td>0.71</td>
<td>-0.45</td>
<td>-0.94</td>
</tr>
<tr>
<td>-0.31</td>
<td>0.24</td>
<td>-0.91</td>
<td>0.87</td>
<td>-0.42</td>
<td>-0.47</td>
<td>-0.31</td>
<td>-0.81</td>
<td>0.40</td>
</tr>
<tr>
<td>-0.31</td>
<td>-0.23</td>
<td>-0.73</td>
<td>-0.16</td>
<td>-0.04</td>
<td>0.05</td>
<td>-0.52</td>
<td>-0.08</td>
<td>0.84</td>
</tr>
<tr>
<td>-0.31</td>
<td>-0.68</td>
<td>0.29</td>
<td>-0.79</td>
<td>0.82</td>
<td>0.70</td>
<td>0.41</td>
<td>0.25</td>
<td>-0.30</td>
</tr>
<tr>
<td>-0.31</td>
<td>0.84</td>
<td>0.70</td>
<td>-0.64</td>
<td>0.27</td>
<td>-0.61</td>
<td>0.97</td>
<td>-0.55</td>
<td>0.75</td>
</tr>
<tr>
<td>0.42</td>
<td>0.65</td>
<td>0.88</td>
<td>-0.36</td>
<td>0.13</td>
<td>-0.47</td>
<td>-0.52</td>
<td>0.25</td>
<td>0.75</td>
</tr>
<tr>
<td>0.42</td>
<td>-0.45</td>
<td>0.11</td>
<td>0.18</td>
<td>0.57</td>
<td>0.05</td>
<td>-0.31</td>
<td>-0.55</td>
<td>-0.30</td>
</tr>
<tr>
<td>0.42</td>
<td>-0.85</td>
<td>-0.35</td>
<td>0.65</td>
<td>-0.56</td>
<td>0.70</td>
<td>0.97</td>
<td>-0.81</td>
<td>0.84</td>
</tr>
<tr>
<td>0.42</td>
<td>0.42</td>
<td>-0.35</td>
<td>0.65</td>
<td>0.27</td>
<td>-0.85</td>
<td>0.15</td>
<td>0.33</td>
<td>-0.94</td>
</tr>
<tr>
<td>0.42</td>
<td>-0.23</td>
<td>-0.91</td>
<td>0.87</td>
<td>-0.42</td>
<td>-0.85</td>
<td>0.15</td>
<td>0.33</td>
<td>-0.94</td>
</tr>
<tr>
<td>0.42</td>
<td>-0.68</td>
<td>0.70</td>
<td>-0.87</td>
<td>-0.04</td>
<td>0.76</td>
<td>0.71</td>
<td>0.65</td>
<td>0.20</td>
</tr>
<tr>
<td>0.42</td>
<td>-0.84</td>
<td>-0.79</td>
<td>0.16</td>
<td>0.42</td>
<td>0.39</td>
<td>-0.92</td>
<td>0.76</td>
<td>-0.09</td>
</tr>
<tr>
<td>-0.70</td>
<td>-0.65</td>
<td>-0.35</td>
<td>0.87</td>
<td>0.82</td>
<td>0.05</td>
<td>0.97</td>
<td>0.76</td>
<td>-0.94</td>
</tr>
<tr>
<td>-0.70</td>
<td>-0.45</td>
<td>-0.13</td>
<td>0.16</td>
<td>0.27</td>
<td>-0.47</td>
<td>0.41</td>
<td>0.65</td>
<td>-0.53</td>
</tr>
<tr>
<td>-0.70</td>
<td>-0.85</td>
<td>0.11</td>
<td>-0.79</td>
<td>-0.42</td>
<td>-0.61</td>
<td>-0.52</td>
<td>-0.45</td>
<td>0.20</td>
</tr>
<tr>
<td>-0.70</td>
<td>-0.23</td>
<td>-0.03</td>
<td>0.34</td>
<td>0.57</td>
<td>-0.08</td>
<td>0.71</td>
<td>-0.08</td>
<td>0.75</td>
</tr>
<tr>
<td>-0.70</td>
<td>-0.23</td>
<td>0.29</td>
<td>0.34</td>
<td>0.13</td>
<td>-0.85</td>
<td>-0.92</td>
<td>-0.81</td>
<td>-0.30</td>
</tr>
<tr>
<td>-0.70</td>
<td>-0.68</td>
<td>-0.91</td>
<td>0.36</td>
<td>-0.91</td>
<td>0.39</td>
<td>0.15</td>
<td>-0.55</td>
<td>0.84</td>
</tr>
<tr>
<td>-0.70</td>
<td>0.84</td>
<td>-0.73</td>
<td>0.18</td>
<td>-0.56</td>
<td>0.76</td>
<td>-0.08</td>
<td>0.25</td>
<td>0.40</td>
</tr>
<tr>
<td>-0.13</td>
<td>0.65</td>
<td>-0.13</td>
<td>-0.79</td>
<td>-0.04</td>
<td>-0.08</td>
<td>-0.92</td>
<td>-0.81</td>
<td>0.75</td>
</tr>
<tr>
<td>-0.13</td>
<td>0.45</td>
<td>-0.35</td>
<td>-0.64</td>
<td>-0.42</td>
<td>-0.85</td>
<td>0.71</td>
<td>-0.25</td>
<td>0.84</td>
</tr>
<tr>
<td>-0.13</td>
<td>-0.85</td>
<td>0.88</td>
<td>0.87</td>
<td>0.27</td>
<td>0.39</td>
<td>-0.08</td>
<td>-0.08</td>
<td>-0.30</td>
</tr>
<tr>
<td>-0.13</td>
<td>-0.42</td>
<td>0.11</td>
<td>-0.16</td>
<td>0.82</td>
<td>0.76</td>
<td>0.15</td>
<td>-0.81</td>
<td>0.75</td>
</tr>
<tr>
<td>-0.13</td>
<td>0.24</td>
<td>0.70</td>
<td>-0.36</td>
<td>0.56</td>
<td>0.05</td>
<td>0.41</td>
<td>-0.45</td>
<td>-0.09</td>
</tr>
<tr>
<td>-0.13</td>
<td>-0.23</td>
<td>0.29</td>
<td>-0.18</td>
<td>-0.91</td>
<td>-0.47</td>
<td>0.97</td>
<td>0.33</td>
<td>0.20</td>
</tr>
<tr>
<td>-0.13</td>
<td>-0.68</td>
<td>0.73</td>
<td>-0.34</td>
<td>0.13</td>
<td>-0.61</td>
<td>-0.31</td>
<td>0.76</td>
<td>-0.53</td>
</tr>
<tr>
<td>$x_{d1}$</td>
<td>$x_{d2}$</td>
<td>$x_{d3}$</td>
<td>$x_{d4}$</td>
<td></td>
<td>$x_{e1}$</td>
<td>$x_{e2}$</td>
<td>$x_{e3}$</td>
<td>$x_{e4}$</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td></td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>-0.13</td>
<td>0.84</td>
<td>-0.91</td>
<td>0.65</td>
<td>0.57</td>
<td>0.70</td>
<td>-0.52</td>
<td>0.65</td>
<td>-0.94</td>
</tr>
<tr>
<td>0.23</td>
<td>-0.65</td>
<td>-0.91</td>
<td>-0.16</td>
<td>-0.56</td>
<td>-0.61</td>
<td>0.71</td>
<td>0.33</td>
<td>-0.30</td>
</tr>
<tr>
<td>0.23</td>
<td>-0.45</td>
<td>-0.73</td>
<td>0.87</td>
<td>-0.91</td>
<td>0.70</td>
<td>-0.92</td>
<td>-0.45</td>
<td>-0.75</td>
</tr>
<tr>
<td>0.23</td>
<td>-0.85</td>
<td>0.29</td>
<td>-0.64</td>
<td>0.13</td>
<td>0.05</td>
<td>0.15</td>
<td>0.65</td>
<td>0.40</td>
</tr>
<tr>
<td>0.23</td>
<td>0.42</td>
<td>0.70</td>
<td>-0.79</td>
<td>0.57</td>
<td>-0.47</td>
<td>-0.08</td>
<td>0.76</td>
<td>0.84</td>
</tr>
<tr>
<td>0.23</td>
<td>0.24</td>
<td>0.11</td>
<td>0.65</td>
<td>-0.04</td>
<td>0.39</td>
<td>0.97</td>
<td>0.25</td>
<td>-0.53</td>
</tr>
<tr>
<td>0.23</td>
<td>-0.23</td>
<td>0.88</td>
<td>0.34</td>
<td>-0.42</td>
<td>0.76</td>
<td>0.41</td>
<td>-0.55</td>
<td>-0.94</td>
</tr>
<tr>
<td>0.23</td>
<td>-0.68</td>
<td>-0.35</td>
<td>0.18</td>
<td>0.27</td>
<td>-0.08</td>
<td>-0.52</td>
<td>-0.81</td>
<td>-0.09</td>
</tr>
<tr>
<td>0.23</td>
<td>0.84</td>
<td>-0.13</td>
<td>-0.36</td>
<td>0.82</td>
<td>-0.85</td>
<td>-0.31</td>
<td>-0.08</td>
<td>0.20</td>
</tr>
<tr>
<td>0.94</td>
<td>0.65</td>
<td>-0.73</td>
<td>-0.64</td>
<td>0.57</td>
<td>0.39</td>
<td>0.41</td>
<td>-0.81</td>
<td>0.20</td>
</tr>
<tr>
<td>0.94</td>
<td>-0.45</td>
<td>-0.91</td>
<td>-0.79</td>
<td>0.13</td>
<td>0.76</td>
<td>0.97</td>
<td>-0.08</td>
<td>-0.09</td>
</tr>
<tr>
<td>0.94</td>
<td>-0.85</td>
<td>0.70</td>
<td>-0.16</td>
<td>-0.91</td>
<td>-0.08</td>
<td>0.31</td>
<td>0.25</td>
<td>-0.94</td>
</tr>
<tr>
<td>0.94</td>
<td>0.42</td>
<td>0.29</td>
<td>0.87</td>
<td>-0.56</td>
<td>-0.85</td>
<td>-0.52</td>
<td>-0.55</td>
<td>-0.53</td>
</tr>
<tr>
<td>0.94</td>
<td>0.24</td>
<td>0.88</td>
<td>0.18</td>
<td>0.82</td>
<td>-0.61</td>
<td>-0.92</td>
<td>0.65</td>
<td>0.84</td>
</tr>
<tr>
<td>0.94</td>
<td>-0.23</td>
<td>0.11</td>
<td>-0.36</td>
<td>0.27</td>
<td>0.70</td>
<td>0.71</td>
<td>0.76</td>
<td>0.40</td>
</tr>
<tr>
<td>0.94</td>
<td>-0.68</td>
<td>-0.13</td>
<td>0.65</td>
<td>-0.42</td>
<td>0.05</td>
<td>-0.08</td>
<td>0.33</td>
<td>0.75</td>
</tr>
<tr>
<td>0.94</td>
<td>0.84</td>
<td>-0.35</td>
<td>0.34</td>
<td>-0.04</td>
<td>-0.47</td>
<td>0.15</td>
<td>-0.45</td>
<td>-0.30</td>
</tr>
<tr>
<td>-0.92</td>
<td>0.65</td>
<td>0.29</td>
<td>0.65</td>
<td>0.27</td>
<td>0.76</td>
<td>-0.31</td>
<td>-0.45</td>
<td>0.84</td>
</tr>
<tr>
<td>-0.92</td>
<td>-0.45</td>
<td>0.70</td>
<td>0.34</td>
<td>0.82</td>
<td>0.39</td>
<td>-0.52</td>
<td>0.33</td>
<td>0.40</td>
</tr>
<tr>
<td>-0.92</td>
<td>-0.85</td>
<td>-0.91</td>
<td>0.18</td>
<td>-0.04</td>
<td>-0.85</td>
<td>0.41</td>
<td>0.76</td>
<td>0.75</td>
</tr>
<tr>
<td>-0.92</td>
<td>0.42</td>
<td>-0.73</td>
<td>-0.36</td>
<td>-0.42</td>
<td>-0.08</td>
<td>0.97</td>
<td>0.65</td>
<td>-0.30</td>
</tr>
<tr>
<td>-0.92</td>
<td>0.24</td>
<td>-0.35</td>
<td>-0.16</td>
<td>0.13</td>
<td>0.70</td>
<td>-0.08</td>
<td>-0.55</td>
<td>0.20</td>
</tr>
<tr>
<td>-0.92</td>
<td>-0.23</td>
<td>-0.13</td>
<td>0.87</td>
<td>0.57</td>
<td>-0.61</td>
<td>0.15</td>
<td>0.25</td>
<td>-0.09</td>
</tr>
<tr>
<td>-0.92</td>
<td>-0.68</td>
<td>0.11</td>
<td>-0.64</td>
<td>-0.56</td>
<td>-0.47</td>
<td>-0.92</td>
<td>-0.08</td>
<td>-0.94</td>
</tr>
<tr>
<td>-0.92</td>
<td>0.84</td>
<td>0.88</td>
<td>-0.79</td>
<td>-0.91</td>
<td>0.05</td>
<td>0.71</td>
<td>-0.81</td>
<td>-0.53</td>
</tr>
<tr>
<td>0.66</td>
<td>0.65</td>
<td>0.70</td>
<td>0.18</td>
<td>-0.42</td>
<td>0.70</td>
<td>0.15</td>
<td>-0.08</td>
<td>-0.53</td>
</tr>
<tr>
<td>0.66</td>
<td>-0.45</td>
<td>0.29</td>
<td>-0.36</td>
<td>-0.04</td>
<td>-0.61</td>
<td>-0.08</td>
<td>-0.81</td>
<td>-0.94</td>
</tr>
<tr>
<td>0.66</td>
<td>-0.85</td>
<td>-0.73</td>
<td>0.65</td>
<td>0.82</td>
<td>-0.47</td>
<td>0.71</td>
<td>-0.55</td>
<td>-0.09</td>
</tr>
<tr>
<td>0.66</td>
<td>0.42</td>
<td>-0.91</td>
<td>0.34</td>
<td>0.27</td>
<td>0.05</td>
<td>-0.92</td>
<td>0.25</td>
<td>0.20</td>
</tr>
<tr>
<td>0.66</td>
<td>0.24</td>
<td>-0.13</td>
<td>-0.64</td>
<td>-0.91</td>
<td>0.76</td>
<td>-0.52</td>
<td>0.76</td>
<td>-0.30</td>
</tr>
<tr>
<td>0.66</td>
<td>-0.23</td>
<td>-0.35</td>
<td>-0.79</td>
<td>-0.56</td>
<td>0.39</td>
<td>-0.31</td>
<td>0.65</td>
<td>0.75</td>
</tr>
<tr>
<td>0.66</td>
<td>-0.68</td>
<td>0.88</td>
<td>-0.16</td>
<td>0.57</td>
<td>-0.85</td>
<td>0.97</td>
<td>-0.45</td>
<td>0.40</td>
</tr>
<tr>
<td>0.66</td>
<td>0.84</td>
<td>0.11</td>
<td>0.87</td>
<td>0.13</td>
<td>-0.08</td>
<td>0.41</td>
<td>0.33</td>
<td>0.84</td>
</tr>
</tbody>
</table>
CHAPTER V

Stochastic Domination Designs for Linear Estimators

5.1 Introduction

In this chapter we consider a "new" stochastic design criterion for making inferences on $K\beta$ where $\beta \in \mathbb{R}^p$ is an unknown parameter vector and $K$ is a known $k \times p$ matrix of rank $k$. We assume the classical linear model in (1.3.2) with a spherically contoured distribution assumption for the error term. In this set-up, the least-squares estimator $K\hat{\beta}$ stochastically dominates any other linear unbiased estimator for $K\beta$ under the generalized Euclidean error with respect to a given $k \times k$ positive definite matrix $D$ (see Corollary 1.3.14). Therefore, the goal of this chapter is to determine the stochastically smallest distribution of $|K\hat{\beta} - K\beta|_D$ by choice of design.

Design Criterion 5.1.1 (SD) A design measure $\xi^* \in \Xi$ is called a stochastic domination design for estimating $K\beta$ if and only if for all $\beta \in \mathbb{R}^p$ and all $\epsilon > 0$

$$P_{\xi^*} \{ |K\hat{\beta} - K\beta|_D > \epsilon \} \leq P_\xi \{ |K\hat{\beta} - K\beta|_D > \epsilon \}. \quad (5.1.1)$$
for every $\xi \in \Xi$ where $|K\hat{\beta} - K\beta|_D$ is the generalized Euclidean error of $K\hat{\beta}$ with respect to $D$.

For $D = I_k$, the Design Criterion 5.1.1 reduces to the criterion considered by Sinha (1970). Under the assumption of normality, he proved that equal allocation of the experimental units in a completely randomized design minimizes the probability in (5.1.1) for estimating the treatment effects in a one-way set-up. Further, he showed that the balanced block design is optimal for estimating the treatment contrasts for two-way classified data.

Shah and Sinha (1989) claim the SD criterion satisfies the conditions for their definition of universal optimality. As a member of this class of criteria, the results cited in the previous paragraph become immediate consequences. In the next section, we shall verify a slightly stronger property for the SD criterion. The consequence of this result is a more appropriate definition of universal optimality.

5.2 SD-Optimal Designs

Let $\mathcal{X}$ be the design space, and let $\Xi$ be the set of design measures defined on the Borel sets of $\mathcal{X}$. In the following theorem, we assume that the information matrix $M(\xi)$ for the vector $\beta$ is invertible for each $\xi \in \Xi$. In addition, for a given $k \times k$ positive definite matrix $D$, let $D^{1/2}$ be the positive definite matrix such that $D = D^{1/2}D^{1/2}$.

**Theorem 5.2.1** Assume the model in (1.3.2) with a spherically distributed error term for which the density exists. For given $k \times k$ positive definite matrix $D$ and real
$k \times p$ matrix $K$ of rank $k$, the design $\xi^* \in \Xi$ is a stochastic domination design for estimating $K\beta$ if the eigenvalues of $(D^{1/2}KM(\xi^*)^{-1}K'D^{1/2})^{-1}$ are majorized by the eigenvalues of $(D^{1/2}KM(\xi)^{-1}K'D^{1/2})^{-1}$ for every $\xi \in \Xi$.

**Proof of Theorem 5.2.1:** Let $|K\hat{\beta}_\xi - K\beta|_D$ be the generalized Euclidean error of $K\hat{\beta}$ with respect to $D$ for the design $\xi \in \Xi$. This random variable can be equivalently written as

$$
|K\hat{\beta}_\xi - K\beta|_D = |D^{1/2}(K\hat{\beta}_\xi - K\beta)|
$$

(5.2.1)

where $|\cdot|$ is the usual Euclidean norm. The vector $K\hat{\beta}_\xi - K\beta$ is an elliptically distributed random variable with $\eta = 0_k$ and positive definite $\Sigma = \sigma^2 KM(\xi)^{-1}K'$ (Kelker, 1970). This implies the vector $D^{1/2}(K\hat{\beta}_\xi - K\beta)$ is elliptically distributed with $\eta = 0_k$ and $\Sigma = \sigma^2 D^{1/2}KM(\xi)^{-1}K'D^{1/2}$. Using (1.3.17), the characteristic function of $D^{1/2}(K\hat{\beta}_\xi - K\beta)$ is of the form

$$
\Psi(t) = \psi(\sigma^2 t'D^{1/2}KM(\xi)^{-1}K'D^{1/2}t).
$$

(5.2.2)

Lord (1954) showed that (5.2.2) is also the characteristic function for the random vector $(D^{1/2}KM(\xi)^{-1}K'D^{1/2})^{1/2} Z$ where $Z$ is a spherically distributed random vector with density $f_Z$ and $\Sigma = \sigma^2 I_k$. By the uniqueness of characteristic functions, we have that

$$
|K\hat{\beta}_\xi - K\beta|_D = \left| (D^{1/2}KM(\xi)^{-1}K'D^{1/2})^{1/2} Z \right|.
$$

(5.2.3)

Let $T$ be the orthogonal matrix that diagonalizes the $k \times k$ positive definite matrix $(D^{1/2}KM(\xi^*)^{-1}K'D^{1/2})^{-1}$, and let $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k)$ be a diagonal matrix with
the eigenvalues of \((D^{1/2}K\lambda(M(\xi^*)^{-1}^{-1})D^{1/2})^{-1}\) along the diagonal. Since \(Z\) is spherically distributed, \(TZ = Z\) in distribution. It follows that

\[
|K\hat{\beta}_\xi - K\beta|_D = \left(\sum_{i=1}^{k} \frac{Z_i^2}{\lambda_i}\right)^{1/2}
\]

For given \(\epsilon > 0\), let \(g(z) = \epsilon^2z\) for \(z > 0\); \(g(\cdot)\) is positive, increasing and convex. Since the density \(f_Z\) is Schur concave, Theorem 1.3.9 says that \(g(\cdot)\) is a Schur concave function of \((\lambda_1, \lambda_2, \ldots, \lambda_k)\). This implies

\[
P\left\{\left(\sum_{i=1}^{k} \frac{Z_i^2}{\lambda_i}\right)^{1/2} < \epsilon\right\}
\]

is a Schur convex function of \((\lambda_1, \lambda_2, \ldots, \lambda_k)\) for every \(\epsilon > 0\). Hence, the theorem is proved. □

Let \(D = I_k\), then the probability in Design Criterion 5.1.1 is a Schur convex function on the eigenvalues of the information matrix for \(K\beta\). This counters the assertion on page 14 of Shah and Sinha (1989) that the criterion of Sinha (1970) is not weakly convex. In fact, from Theorem 5.2.1 we have that the criterion is invariant under orthogonal transformations, i.e., the design problem is the same if we estimate \(TK\beta\) for \(T\) orthogonal. This implies Design Criterion 5.1.1 is a function of the eigenvalues of the information matrix for \(K\beta\). In addition, using (5.2.4), we see that a design \(\xi\) is preferred to \(\zeta\) whenever the two sets of eigenvalues of \((K\lambda(M(\cdot)^{-1})^{-1})^{-1}\) satisfy \(\lambda^\xi_i \leq \lambda^\zeta_i\) for \(i = 1, 2, \ldots, k\) with strict inequality for some \(i\).
The three properties mentioned in the previous paragraph are the conditions on design criteria for Bondar's (1983) definition of universally optimal designs. They imply the conditions for extended universal optimality (Shah and Sinha, 1989). Nearly all criteria that satisfy the conditions for extended universal optimality also satisfy the conditions of Bondar, so we might as well require the stronger properties for the definition of universal optimality. It is not known if Design Criterion 5.1.1 satisfies the conditions for Kiefer's (1975) universal optimality, since it is unclear how one would interpret the usual notion of convexity for the probability in (5.1.1).
CHAPTER VI

Future Work

My future research will follow two somewhat divergent paths. The first path is directed toward improving the exact designs constructed in Chapter III, while the second involves further investigation of the \( \mu \)-LM model (4.2.2) in the response model approach to parameter design by means of computer experimentation. Although some may dismiss the first research path based on the results of Chapter IV, it is of interest to know how well the competing methodologies fare in situations where the \( \mu \)-LM model struggles. Further, we note that the design columns for the controllable variables \( x_d \) in the experimental designs of Theorem 3.5.1 and Theorem 3.6.1 are efficient designs for the IavA and IavD criteria, respectively, for estimating the optimal point of a quadratic response surface in \( x_d \). Hence, these designs have applications beyond the parameter design problem.

To construct the optimal designs of Chapter III, we considered for the support of the control variables only the triples of isozero sets \( \{\mathcal{H}^i, \mathcal{H}^j, \mathcal{H}^k\} \) with either \( 0 = i < j < k < k_d \) or \( 0 < i < j < k = k_d \). The exclusion of the central composite designs (c.c.d) \( \mathcal{D}(0, k_d - 1, k_d) \) stemmed from Conjecture 3.7.1 and the fact that c.c.d’s provide only one isozero set that can be split for the IavA criterion. However,
during the construction of the sample designs in Section 3.8, we found that efficient
designs for the IavA criterion, with $\theta \neq 1$, could be obtained by splitting only one
of the isozero sets. Since the c.c.d's are the smallest designs supported on a triple of
isozero sets this suggests that we should reexamine the central composite designs for
the parameter design problem.

In Table 22 we present the control variable levels of a 41 point central composite
design for the $(k_d, k_e) = (4, 5)$ parameter design problem. The model has 40 parameters,
so this design is nearly saturated with IavA and IavD criteria efficiencies of .821
and .454, respectively.

Table 22: The control variable portion of a 41 point central composite design for the
$(k_d, k_e) = (4, 5)$ parameter design problem with $s_d = 1$ and $\theta = 1$ (IavA criterion
efficiency = .821; IavD criterion efficiency = .454).

<table>
<thead>
<tr>
<th>Isozero Set</th>
<th>Cardinality</th>
<th>Replications</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{H}^0$</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>$\mathcal{H}^{(0)}$</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>$\mathcal{H}^{(1)}$</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>$\mathcal{H}^4$</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 23 presents the control variable levels for a 163 point central composite
design for the $(k_d, k_e) = (7, 5)$ parameter design problem in Example 3.8.3. This
design has an IavA criterion efficiency of .570. Further, if we reduce the number of
replicates for each point in $\mathcal{H}^3$ to four, then there are 42 fewer design points, but the
efficiency only drops to .510.
Table 23: The control variable portion of a 163 point central composite design for the 
\((k_d, k_r) = (7, 5)\) parameter design problem with \(s_d = 1.5\) and \(\theta = 0.5\) (IavA criterion 
efficiency = .570).

<table>
<thead>
<tr>
<th>Isozero Set</th>
<th>Cardinality</th>
<th>Replications</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2 fraction of (\mathcal{H}^0)</td>
<td>64</td>
<td>1</td>
</tr>
<tr>
<td>(\mathcal{H}^{6(0)})</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>(\mathcal{H}^{6(1)})</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>(\mathcal{H}^7)</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

The design in Table 23 uses only a 1/2 fraction of the corner points \(\mathcal{H}^0\). It 
was noted in Chapter III that the isozero sets could be fractionated as long the 
design was of Resolution V. However, Hartley (1959) found that designs as low 
as Resolution III could be used to estimate the \(\binom{k_d+2}{2}\) regression parameters for 
the controllable variables provided two-way interactions were not aliased together. 
Draper (1985) and Draper and Lin (1990) construct small central composite designs of 
Resolution III for fitting a second-order response surface using columns of a Plackett 
and Burman design. These designs provide a substantial savings in experimental 
units, and they are easy to construct. Even though these designs will violate the 
structure of the information matrix in Theorem 3.5.1 and Theorem 3.6.1, they merit 
further consideration for the parameter design problem.

I also intend to develop better methods for selecting the environmental variable 
levels for the optimal exact design. The ad hoc procedure that we described in 
Section 2.8 does not perform adequately when the number of design points \(N\) is
large. This problem is partially mitigated by reducing the number of design points using the procedures described in the previous paragraphs. However, it would be better to have a sampling procedure, similar to Latin hypercube sampling, that could be easily implemented for any experimental situation.

The $\mu$-LM model in (4.2.2) has gained popularity with the statistical community because of its flexibility. Several articles have discussed the appropriate choice of correlation function and have detailed procedures for estimating the model parameters. The role of experimental design in the methodology is less well-developed, so this will continue to be an active area of research.

My majorization result in Chapter V is only a sufficient condition for the existence of a stochastic domination design. To appreciate the significance of the SD criterion, necessary conditions must also be determined. However, the determination of necessary conditions is hampered because the criterion is not real-valued. I propose studying the modified criterion

$$\int_0^\infty P_{\xi} \cdot \{ |K\beta - K\beta|_D > \epsilon \} \, d\pi(\epsilon)$$

(6.0.1)

where $\pi(\cdot)$ is a positive (density) function. When an SD-optimal design exists, it will minimize (6.0.1). Thus, minimization of (6.0.1) may lead us toward necessary conditions for the SD optimality.
Appendix A

Calculations for the IavA-optimal design

A.1 Proof of Lemma 3.5.2

Assume the prior \( \pi^* = \pi_B^* \pi_d^* \) for which the hypotheses in Assumptions 3.2.1 are satisfied. The matrix \( H^t A_d^{-2} H \) equals

\[
\begin{pmatrix}
0 & 0_{k_d}^t & 0_{k_d}^t & 0_{(k_d+1)k_r}^t & 0_{(k_d+1)k_r}^t \\
0_{k_d} & \frac{1}{4} A_d^{-2} & \frac{1}{2} A_d^{-2} A_u & \frac{1}{4} A_d^{-2} \Gamma_u & 0_{(k_d+1)k_r} \\
0_{k_d} & \frac{1}{2} A_u A_d^{-2} & A_u A_d^{-2} A_u & \frac{1}{2} A_u A_d^{-2} \Gamma_u & 0_{(k_d+1)k_r} \\
0_{(k_d+1)k_r} & \frac{1}{4} \Gamma_u A_d^{-2} & \frac{1}{2} \Gamma_u A_d^{-2} A_u & \frac{1}{4} \Gamma_u A_d^{-2} \Gamma_u & 0_{(k_d+1)k_r} \\
0_{(k_d+1)k_r} & 0_{(k_d+1)k_r} & 0_{(k_d+1)k_r} & 0_{(k_d+1)k_r} & 0_{(k_d+1)k_r} \\
\end{pmatrix}
\]

(A.1.1)

where the matrices \( \Gamma_u^* \) and \( A_u^* \), defined in (3.4.12) and (3.4.14), are functions of \( u^* \) only, and \( A_d^{-2} \) is free of \( u^* \). To calculate \( E_{\pi^*}[H^t A_d^{-2} H] \), we proceed by first integrating \( H^t A_d^{-2} H \) with respect to \( \pi_d^* \) then integrating with respect to \( \pi_B^* \).

171
Recall the first hypothesis, i.e., $E^\pi_d[u^*|\alpha_0, A_d, \alpha_e, A_I] = 0_{k_d}$. Under this assumption $E^\pi_u[\Lambda^*_u] = 0_{k_d \times k_d}$ and $E^\pi_u[\Gamma_u^*] = 0_{k_d \times (k_d^*)}$, so $E^\pi_u[H^tA_d^{-2}H]$ equals

$$
\begin{pmatrix}
0 & 0_{k_d}^t & 0_{k_d}^t & 0_{k_d}^t(k_d^*) & 0_{(k_d+1)k_r}^t \\
0_{k_d} & \frac{1}{4}E^\pi_n[A_{d}^{-2}] & 0_{k_d \times k_d} & 0_{k_d \times (k_d^*)} & 0_{k_d \times (k_d+1)k_r} \\
0_{k_d} & 0_{k_d \times k_d} & E^\pi_u[\Lambda^*_uA_{d}^{-2}\Lambda^*_u] & \frac{1}{2}E^\pi_u[\Lambda^*_uA_{d}^{-2}\Gamma_u^*] & 0_{k_d \times (k_d+1)k_r} \\
0_{(k_d^*)} & 0_{(k_d^*) \times k_d} & \frac{1}{4}E^\pi_u[\Gamma^t_uA_{d}^{-2}\Lambda^*_u] & \frac{1}{4}E^\pi_u[\Gamma^t_uA_{d}^{-2}\Gamma_u^*] & 0_{(k_d^*) \times (k_d+1)k_r} \\
0_{(k_d+1)k_r} & 0_{(k_d+1)k_r \times k_d} & 0_{(k_d+1)k_r \times k_d} & 0_{(k_d+1)k_r \times (k_d^*)} & 0_{(k_d+1)k_r \times (k_d+1)k_r}
\end{pmatrix}
$$

(A.1.2)

The nonzero elements of both $\Gamma_u^*$ and $\Lambda_u^*$ are first-order functions of the coordinates of $u^*$. Therefore, the nonzero elements in the three matrix integrands in (A.1.2), i.e. $\Lambda_u^tA_{d}^{-2}\Lambda_u^*$, $\Gamma_u^tA_{d}^{-2}\Lambda_u^*$, and $\Gamma_u^tA_{d}^{-2}\Gamma_u^*$, are quadratic functions of the coordinates of $u^*$. To find the expected values of these integrands with respect to $\pi_d^*$ requires the second hypothesis, i.e. $E^\pi_d[u^*u^{*\dagger}|\alpha_0, A_d, \alpha_e, A_I] = s_d^{-2}I_{k_d}$. Since the values of these expectations, as well as $A_{d}^{-2}$, are functions of the random variables in $\alpha_0, A_d, \alpha_e$, and $A_I$, we also need the final hypothesis, i.e. $E^\pi_n[A_{d}^{-2}] = s_d^{-4}\Lambda_Q$, to complete the calculation of $E^\pi_u[H^tA_d^{-2}H]$.

Let us consider the expectation $E^\pi_u[\Lambda_u^tA_{d}^{-2}\Lambda_u^*]$. Given $1 \leq i, j \leq k_d$, the expected value of the $(i, j)$ element of $\Lambda_u^tA_{d}^{-2}\Lambda_u^*$ is

$$
E^\pi_u[(\Lambda_u^tA_{d}^{-2}\Lambda_u^*)_{i,j}] = E^\pi_n[(A_{d}^{-2})_{i,j}]E^\pi_d[u_i^*u_j^*|\alpha_0, A_d, \alpha_e, A_I]]
= s_d^{-2}(E^\pi_n[A_{d}^{-2}])_{i,j}\delta_{ij}
= s_d^{-6}(\Lambda_Q)_{i,j}
$$

(A.1.3)
This implies

\[ E^n \{ \Lambda^t_{u*} A_d^{-2} \Lambda_{u*} \} = s_d^{-6} \Lambda_Q. \]  \hspace{1cm} (A.1.4)

Next, we consider the expectation \( E^n \{ \Gamma^t_{u*} A_d^{-2} \Lambda_{u*} \} \). To determine this expectation, we first define the bijective relationship

\[ h_{ij} = \frac{(2k_d - i)(i - 1)}{2} + j - i \quad \text{for all } \{(i, j) : 1 \leq i < j \leq k_d\} \]  \hspace{1cm} (A.1.5)

between the set \( \{(i, j) : 1 \leq i < j \leq k_d\} \) and the integers \( \{1, 2, \ldots, \binom{k_d}{2}\} \). We denote this relationship by \( h_{ij} \downarrow (i, j) \).

Fix \( k \in \{1, \ldots, k_d\} \) and \( h_{ij} \in \{1, \ldots, \binom{k_d^i}{2}\} \). The expected value of the \( (h_{ij}, k) \) element of \( \Gamma^t_{u*} A_d^{-2} \Lambda_{u*} \) is

\[
E^n \{ (\Gamma^t_{u*} A_d^{-2} \Lambda_{u*})_{h_{ij}, k} \} = E^n \{ \sum_{l=1}^{k_d} ((A_d^{-2})_{i,l} u_j^* + (A_d^{-2})_{j,l} u_i^*) u_k^* \delta_{kl} \}
\]

\[
= E^n [ (A_d^{-2})_{i,k} E^n [u_j^* u_k^* | \alpha_0, A_d, \alpha_c, A_I] ]
\]

\[
+ E^n [ (A_d^{-2})_{j,k} E^n [u_i^* u_k^* | \alpha_0, A_d, \alpha_c, A_I] ]
\]

\[
= s_d^{-2} E^n [ (A_d^{-2})_{i,k} \delta_{jk} + (A_d^{-2})_{j,k} \delta_{ik} ]
\]

\[
= s_d^{-6} (\Lambda_Q)_{i,j}. \]  \hspace{1cm} (A.1.6)

Since \( \Lambda_Q \) is diagonal and \( i < j \),

\[
E^n \{ \Gamma^t_{u*} A_d^{-2} \Lambda_{u*} \} = 0_{(\binom{k_d}{2}) \times k_d}. \]  \hspace{1cm} (A.1.7)

Finally, we determine the expected value of the matrix \( \Gamma^t_{u*} A_d^{-2} \Gamma_{u*} \). By symmetry, it suffices to consider the expected value for fixed \( h_{ij}^{(1)} \leq h_{ij}^{(2)} \) where \( h_{ij}^{(1)}, h_{ij}^{(2)} \in \{1, \ldots, \binom{k_d}{2}\} \) and \( h_{ij}^{(m)} \downarrow (i_m, j_m) \), with \( m = 1, 2 \). When \( h_{ij}^{(1)} \leq h_{ij}^{(2)} \) the following two relations are true: \( i_1 < j_1 \) and \( i_1 \leq i_2 < j_2 \).
Fix $h_{ij}^{(1)} \leq h_{ij}^{(2)}$. The $(h_{ij}^{(1)}, h_{ij}^{(2)})$ element of $\Gamma_u^t \cdot A_d^{-2} \Gamma_u^* \cdot$ is

$$(\Gamma_u^t \cdot A_d^{-2} \Gamma_u^*)_{h_{ij}^{(1)}, h_{ij}^{(2)}} = \sum_{k=1}^{k_d} [(A_d^{-2})_{1i,k} u_{j1}^* + (A_d^{-2})_{j1,k} u_{i1}^*] [u_{i2}^* \delta_{kj2} + u_{j2}^* \delta_{k12}]$$

$$= (A_d^{-2})_{i1,j2} u_{i1}^* u_{j2}^* + (A_d^{-2})_{i2,j1} u_{i2}^* u_{j1}^* + (A_d^{-2})_{i1,j2} u_{i1}^* u_{j2}^*.$$ (A.1.8)

Hence, the expected value of the $(h_{ij}^{(1)}, h_{ij}^{(2)})$ element of $\Gamma_u^t \cdot A_d^{-2} \Gamma_u^*$ equals

$$E^*[(\Gamma_u^t \cdot A_d^{-2} \Gamma_u^*)_{h_{ij}^{(1)}, h_{ij}^{(2)}}] = s_d^{-6} E^* h[(A_d^{-2})_{i1,j2} \delta_{j1,j2} + (A_d^{-2})_{i1,i2} \delta_{j1,j2} + (A_d^{-2})_{j1,j2} \delta_{i1,i2}].$$ (A.1.9)

The fourth component of the sum in (A.1.8) is not in (A.1.9) since $\delta_{i1,j2} = 0$. Recall $i_1 < j_2$ whenever $h_{ij}^{(1)} \leq h_{ij}^{(2)}$. To find the value of the expectation in (A.1.9), we consider separately each case of $i_1 < j_1$ and $i_1 < i_2 < j_2$ for which $h_{ij}^{(1)} \leq h_{ij}^{(2)}$. The relation $h_{ij}^{(1)} = h_{ij}^{(2)}$ is true if and only if $i_1 = i_2 < j_1 = j_2$. In this case,

$$E^*[(\Gamma_u^t \cdot A_d^{-2} \Gamma_u^*)_{h_{ij}^{(1)}, h_{ij}^{(2)}}] = s_d^{-6} (\lambda_{i1} + \lambda_{j1}).$$ (A.1.10)

where $\lambda_k$ is the eigenvalue of $Q$ (3.2.1) in the $k^{th}$ position on diagonal $\Lambda_Q$. When $h_{ij}^{(1)} < h_{ij}^{(2)}$, there are four possible situations. First, if $i_1 = i_2 < j_1 < j_2$, then

$$E^*[(\Gamma_u^t \cdot A_d^{-2} \Gamma_u^*)_{h_{ij}^{(1)}, h_{ij}^{(2)}}] = s_d^{-6} (\Lambda_Q)_{j1,j2} = 0.$$ (A.1.11)

Second, if $i_1 < i_2 < j_1 = j_2$, then

$$E^*[(\Gamma_u^t \cdot A_d^{-2} \Gamma_u^*)_{h_{ij}^{(1)}, h_{ij}^{(2)}}] = s_d^{-6} (\Lambda_Q)_{i1,i2} = 0.$$ (A.1.12)

Third, if $i_1 < j_1 = i_2 = j_2$, then

$$E^*[(\Gamma_u^t \cdot A_d^{-2} \Gamma_u^*)_{h_{ij}^{(1)}, h_{ij}^{(2)}}] = s_d^{-6} (\Lambda_Q)_{i1,j2} = 0.$$ (A.1.13)
Finally, if \( i_2 \neq j_1 \neq j_2 \), then
\[
E^{\pi^*}[\Gamma'_u \cdot A_d^{-2} \Gamma_{u^*} h_{i_j}^{(1)} h_{i_j}^{(2)}] = 0. \tag{A.1.14}
\]
Together these imply that
\[
E^{\pi^*}[\Gamma'_u \cdot A_d^{-2} \Gamma_{u^*}] = 0 \text{ if } h_{i_j}^{(1)} < h_{i_j}^{(2)} \tag{A.1.15}
\]
By symmetry of the matrix \( \Gamma'_u \cdot A_d^{-2} \Gamma_{u^*} \), the expected value of \( \Gamma'_u \cdot A_d^{-2} \Gamma_{u^*} \), with respect to \( \pi^* \), is the diagonal matrix
\[
E_{\pi^*}[\Gamma'_u \cdot A_d^{-2} \Gamma_{u^*}] = s_d^{-6} \text{diag}(\lambda_1 + \lambda_2, \lambda_1 + \lambda_3, \ldots, \lambda_{k_d-1} + \lambda_{k_d}). \tag{A.1.16}
\]
Let \( \Lambda_{QI} \) denote this diagonal matrix. Recall that either \( \Lambda_Q = \lambda I_{k_d} \) for some \( \lambda > 0 \) or \( \Lambda_Q = \text{diag}(\gamma, \lambda, \ldots, \lambda) \) where \( \gamma > 0, \gamma \neq \lambda > 0 \). This means
\[
\Lambda_{QI} = \text{diag} \left( (\gamma + \lambda) I_{k_d-1}, 2\lambda I_{(k_d-1)^2} \right) \tag{A.1.17}
\]
where \( \gamma \) and \( \lambda \) are possibly equal.

From (A.1.4), (A.1.7) and (A.1.16), \( E^{\pi^*}[\mathbf{H}' A_d^{-1} \mathbf{H}] \) is block diagonal with diagonal matrices in the blocks. Therefore, \( E^{\pi^*}[\mathbf{H}' A_d^{-1} \mathbf{H}] \) is a diagonal matrix of the form
\[
\begin{pmatrix}
0 & 0_{k_d} & 0_{k_d} & 0_{(k_d+1)k_r} \\
0_{k_d} & \frac{s_d^{-4}}{4} \Lambda_Q & 0_{k_d \times k_d} & 0_{k_d \times (k_d+1)k_r} \\
0_{k_d} & 0_{k_d \times k_d} & s_d^{-6} \Lambda_Q & 0_{k_d \times (k_d+1)k_r} \\
0_{(k_d+1)k_r} & 0_{(k_d+1)k_r \times k_d} & 0_{(k_d+1)k_r \times k_d} & 0_{(k_d+1)k_r \times (k_d+1)k_r} \\
0_{(k_d+1)k_r} & 0_{(k_d+1)k_r \times k_d} & 0_{(k_d+1)k_r \times k_d} & 0_{(k_d+1)k_r \times (k_d+1)k_r}
\end{pmatrix}. \tag{A.1.18}
\]
A.2 Calculating $\mathbf{M}(\xi^*)$ in Lemma 3.5.3

In this section we derive the information matrix $\mathbf{M}(\xi^*)$ for the parameter vector $\alpha$ given a design $\xi^*$ that is characterized by the conditions in Theorem 3.5.1. Recall the partition the vector $\alpha$ in (3.5.4). Applying a similar partition to the regression vector $\mathbf{f}$, i.e.

$$\mathbf{f}(\mathbf{u}, \mathbf{v}) = (1, u_1, u_1^2, \ldots, u^2_{k_d}, u_1, u_3, \ldots, u_{k_d-1} u_{k_d} | v', u_1 v_1, u_1 v_2, \ldots, u_{k_d} v_r)$$

$$= (f_1, f_2, f_3), \quad (A.2.1)$$

the information matrix based on the design $\xi^*$ is of the form

$$\begin{bmatrix}
\int f_1 f_1^* d\xi^* & \int f_1 f_2^* d\xi^* & \int f_1 f_3^* d\xi^* \\
\int f_2 f_1^* d\xi^* & \int f_2 f_2^* d\xi^* & \int f_2 f_3^* d\xi^* \\
\int f_3 f_1^* d\xi^* & \int f_3 f_2^* d\xi^* & \int f_3 f_3^* d\xi^*
\end{bmatrix} \quad (A.2.2)$$

Every element in the submatrices $\int f_3 f_1^* d\xi^*$ and $\int f_3 f_2^* d\xi^*$ is of the form

$$E^{\xi^*} u_i u_j u_k u_l u_m u_n \quad (A.2.3)$$

where each $i_j \in \{1, 2, \ldots, k_d\}$, each $q_i \in \{0, 1, 2, 3\}$ with $0 \leq \sum_{j=1}^4 q_{i_j} \leq 3$, and $k \in \{1, 2, \ldots, k_r\}$. Condition vi.) of Theorem 3.5.1 says each of these elements equals zero. This means the information for the parameters in $\alpha_3$ is independent of the information for the parameters in $\alpha_1$ and $\alpha_2$. Similarly, every element of the submatrix $\int f_2 f_1^* d\xi^*$ is equal to

$$E^{\xi^*} u_i^4 u_j u_k u_l u_m u_n \quad (A.2.4)$$

where each $i_j \in \{1, 2, \ldots, k_d\}$, each $q_i \in \{0, 1, 2, 3\}$ with $0 \leq \sum_{j=1}^4 q_{i_j} \leq 4$, and at least one $q_{i_j}$ is odd. From Condition i.) of Theorem 3.5.1, we know each of these
elements is also equal to zero, so the information for the parameters in $\alpha_2$ is also independent of the information for the parameters in $\alpha_1$.

We have established that $M(\xi^*)$ is a block diagonal matrix. Now we determine the values of the submatrices along the diagonal. First, we find the value of the submatrix $\int f_i f_i^t d\xi^*$. Consider the partition of the regression vector $f_1$ such that

$$
\int f_i f_i^t d\xi^* = \begin{pmatrix}
\int (1, u)^t (1, u) d\xi^* & \int (1, u)^t (u_1^2, \ldots, u_{k_d}^2) d\xi^* \\
\int (u_1^2, \ldots, u_{k_d}^2)^t (1, u) d\xi^* & \int (u_1^2, \ldots, u_{k_d}^2)^t (u_1^2, \ldots, u_{k_d}^2) d\xi^*
\end{pmatrix} \quad (A.2.5)
$$

For notational reasons, let $M_d$ denote the the matrix $\int f_i f_i^t d\xi^*$. Then, using (A.2.5) we can express this submatrix as

$$
M_d = \int f_i f_i^t d\xi^* = \begin{pmatrix}
M_{d11} & M_{d12} \\
M_{d21} & M_{d22}
\end{pmatrix} \quad (A.2.6)
$$

The elements of the matrix $M_{d11}$ are at most second-order functions of the control variables, and the only elements that are even functions appear on the diagonal. From Conditions $i., ii.$ and $iii.$ of Theorem 3.5.1, we get

$$
M_{d11} = \begin{pmatrix}
1 & 0_{k_d}^t \\
0_{k_d} & \Lambda_R
\end{pmatrix} \quad (A.2.7)
$$

where $\Lambda_R = \text{diag}(r_1, r_2 I_{k_d-1})$. Each element of the matrix $M_{d22}$ is the product of two quadratic terms for the control variables. Considering Conditions $ii.$ through $v.$ of the theorem, this means the matrix $M_{d22}$ equals

$$
M_{d22} = \begin{pmatrix}
r_1 & t_1 1_{k_d-1}^t \\
t_1 1_{k_d-1} & (r_2 - t_2)I_{k_d-1} + t_2 J_{k_d-1}
\end{pmatrix} \quad (A.2.8)
$$
The only elements of the submatrix \( M_{d12} \) that are even functions of the control variables appear in the top row; hence, \( M_{d12} \) equals

\[
M_{d12} = \begin{pmatrix}
r_1 & r_2 1_{k_d-1} \\
0_{k_d} & 0_{k_d \times (k_d-1)}
\end{pmatrix}.
\] (A.2.9)

Next, we determine the value of the submatrix \( \int f_2 g_2 d\xi^* \). Each element of this matrix has the form

\[
E^{\xi^*} u_{i_1} u_{j_2} u_{i_2} u_{j_2} \tag{A.2.10}
\]

where \( i_1 < j_1 \in \{1, \ldots, k_d\} \) and \( i_2 < j_2 \in \{1, \ldots, k_d\} \). Recalling the bijection \( h_{ij}^{(m)} \leftrightarrow (i_m, j_m) \) for \( m = \{1, 2\} \), we find, using Conditions i., iv., and v.) of Theorem 3.5.1, that

\[
(\int f_2 g_2 d\xi^*)_{h_{ij}^{(1)}, h_{ij}^{(2)}} = \begin{cases}
t_1 \delta_{h_{ij}^{(1)}, h_{ij}^{(2)}} & \text{if } h_{ij}^{(1)} < k_d \\
t_2 \delta_{h_{ij}^{(1)}, h_{ij}^{(2)}} & \text{if } h_{ij}^{(1)} \geq k_d
\end{cases}.
\] (A.2.11)

Hence, \( \int f_2 g_2 d\xi^* \) is a diagonal matrix, which we denote by \( \Lambda_T \), i.e.

\[
\Lambda_T = \int f_2 g_2 d\xi^* = \text{diag}(t_1 I_{k_d-1}, t_2 I_{(k_d-1)}) \tag{A.2.12}
\]

Finally, we consider the matrix \( \int f_3 g_2 d\xi^* \). This matrix contains the information for the noise variables and control-noise interactions. As such, it does not factor into the determination of the optimal operating conditions and so is ignored.

Putting together all of the pieces, we have that the information matrix \( M(\xi^*) \) for the candidate design \( \xi^* \) equals

\[
M_d \quad 0_{(2k_d+1) \times (k_d/2)} \quad 0_{(2k_d+1) \times (k_d+1)k_r} \\
0_{(k_d/2) \times (2k_d+1)} \quad \Lambda_T \quad 0_{(k_d/2) \times (k_d+1)k_r} \\
0_{(k_d+1)k_r \times (2k_d+1)} \quad 0_{(k_d+1)k_r \times (k_d/2)} \quad \int f_3 g_2 d\xi^* 
\] (A.2.13)
A.3 Calculating $M^{-}(\xi^{*})$ in Lemma 3.5.4

In this section we derive $M^{-}(\xi^{*})$ the Moore-Penrose generalized inverse for the information matrix $M(\xi^{*})$ given in Lemma 3.5.3. The block diagonal structure of $M(\xi^{*})$ simplifies the inversion, because we need only calculate the generalized inverse of each block on the diagonal. That means

$$M^{-}(\xi^{*}) = \text{diag}(M_{d}^{-1}, \Lambda_{T}^{-1}, (\int f_{3}f_{3}^{d}d\xi^{*})^{-})$$

(A.3.1)

where $M_{d}^{-1}$ is the unique inverse of $M_{d}$, $\Lambda_{T}^{-1} = \text{diag}\left(t_{1}^{-1}1_{k_{d}-1}, t_{2}^{-1}1_{d}^{*}1_{k_{d}-1}\right)$, and (where $M_{d}^{-1}$ is the unique inverse of $M_{d}$, $\Lambda_{T}^{-1} = \text{diag}\left(t_{1}^{-1}1_{k_{d}-1}, t_{2}^{-1}1_{d}^{*}1_{k_{d}-1}\right)$, and

$(\int f_{3}f_{3}^{d}d\xi^{*})^{-}$ is the Moore-Penrose generalized inverse of $\int f_{3}f_{3}^{d}d\xi^{*}$.

Considering the form of $M^{-}(\xi^{*})$ in (A.3.1), we only need to determine the value of the unique inverse $M_{d}^{-1}$. The existence of a unique inverse means that we are assuming $M_{d}$ is a nonsingular matrix. To make this supposition valid, we impose certain restrictions on the values of the design parameters $r_{1}, r_{2}, t_{1},$ and $t_{2}$. These restrictions are in addition to current relationships that are a consequence of the fact that the random variables $u_{i}, i \in \{1, 2, \ldots, k_{d}\}$, belong to the space $[-1, 1]^{k_{d}}$. The obligatory restrictions are

$$0 \leq t_{1} \leq r_{1} \leq 1 \quad 0 \leq t_{2} \leq r_{2} \leq 1 \quad \text{and} \quad 0 \leq t_{1} \leq r_{2},$$

(A.3.2)

and our additional restrictions are

$$r_{2}^{2} \leq t_{2}$$

(A.3.3)

and

$$r_{1}r_{2} \leq t_{1} \leq r_{1}r_{2} + \sqrt{\frac{(r_{1} - r_{2}^{2})(r_{2} - r_{2}^{2}) + (k_{d} - 2)(t_{2} - r_{2}^{2})}{k_{d} - 1}}.$$  (A.3.4)
Define the constant $w$ as

$$w = (r_1 - r_1^2)(r_2 - r_2^2 + (k_d - 2)(t_2 - r_2^2)) - (k_d - 1)(t_1 - r_1 r_2)^2. \quad (A.3.5)$$

The conditions in (A.3.2), (A.3.3), and (A.3.4) imply that $w > 0$. This constant $w$, which appears numerous times in the generalized inverse matrix $M^{-1}(\xi^*)$, is a divisor of the determinant of $M_d$.

To express $M_d^{-1}$, consider the partition of the matrix $M_d$ given in (A.2.6). The matrix $M_d^{-1}$ equals

$$\begin{pmatrix}
M_{d11}^{-1} & M_{d12}^{-1} \\
M_{d21}^{-1} & M_{d22}^{-1}
\end{pmatrix} \quad (A.3.6)$$

where

$$M_{d11}^{-1} = \begin{pmatrix}
\frac{r_1 r_2 + (k_d - 2)r_1 t_2 - (k_d - 1)t_2^2}{w} & 0'_{k_d} \\
0_{k_d} & \Lambda_R^{-1}
\end{pmatrix}, \quad (A.3.7)$$

with $\Lambda_R^{-1} = \text{diag}(r_1^{-1}, r_2^{-1} I_{k_d-1})$,

$$M_{d12}^{-1} = \begin{pmatrix}
\frac{-r_1 r_2 + (k_d - 2)r_1 t_2 - (k_d - 1)t_2^2}{w} & \frac{-r_1 r_2 - r_1 t_1}{w} 1'_{k_d-1} \\
0_{k_d} & 0_{k_d \times (k_d - 1)}
\end{pmatrix}, \quad (A.3.8)$$

and

$$M_{d22}^{-1} = \begin{pmatrix}
\frac{r_2 - r_1^2 + (k_d - 2)(t_2 - r_1^2)}{w} & \frac{-t_1 r_1 t_2}{w} 1'_{k_d-1} \\
\frac{-t_1 r_1 t_2}{w} & \frac{1}{r_2 - t_2} (I_{k_d-1} - \frac{(r_1 - r_1^2)(t_2 - r_1^2) - (t_1 - r_1 r_2)^2}{w} J_{k_d-1})
\end{pmatrix}. \quad (A.3.9)$$
Appendix B

Calculations for the IavD-optimal design

B.1 Proof of Lemma 3.6.4

Assume a prior distribution \( \pi_d^* \) for which the hypotheses in Assumptions 3.2.2 are satisfied. The matrix \( \mathbf{H}'\mathbf{H} \) is equal to

\[
\left( \begin{array}{cccccc}
0 & 0_{k_d} & 0_{k_d} & 0_{(k_d+1)k_e} \\
0_{k_d} & \frac{1}{4} \mathbf{I}_{k_d} & \frac{1}{2} \mathbf{A}_{u^*} & \frac{1}{k_e} \mathbf{\Gamma}_{u^*} & 0_{k_d \times (k_d+1)k_e} \\
0_{k_d} & \frac{1}{2} \mathbf{A}_{u^*} & \mathbf{A}_{u^*}^2 & \frac{1}{2} \mathbf{A}_{u^*} \mathbf{\Gamma}_{u^*} & 0_{k_d \times (k_d+1)k_e} \\
0_{(k_d+1)k_e} & \frac{1}{4} \mathbf{\Gamma}_{u^*} & \frac{1}{2} \mathbf{\Gamma}_{u^*} \mathbf{A}_{u^*} & \frac{1}{4} \mathbf{\Gamma}_{u^*} \mathbf{\Gamma}_{u^*} & 0_{(k_d+1)^2 \times (k_d+1)k_e} \\
0_{(k_d+1)k_e \times k_d} & 0_{(k_d+1)k_e \times k_d} & 0_{(k_d+1)k_e \times k_d} & 0_{(k_d+1)k_e \times (k_d+1)k_e} & 0_{(k_d+1)k_e \times (k_d+1)k_e}
\end{array} \right)
\]

(B.1.1)

where the matrices \( \mathbf{\Gamma}_{u^*} \) and \( \mathbf{A}_{u^*} \), defined in (3.4.12) and (3.4.13), are functions of \( u^* \) only. Using the first hypothesis, i.e., \( E_{\pi_d^*}[u^*] = 0_{k_d} \), we get \( E_{\pi_d^*}[\mathbf{A}_{u^*}] = 0_{k_d \times k_d} \) and \( E_{\pi_d^*}[\mathbf{\Gamma}_{u^*}] = 0_{k_d \times (k_d+1)} \). The remaining nonconstant elements of \( \mathbf{H}'\mathbf{H} \) are second-order functions of the \( u_i^* \)'s, so we need the second hypothesis, i.e., \( E_{\pi_d^*}[^{u_i^*}u_j^*] = s_d^{-2} \delta_{ij} \), to complete the calculation of the expectation. We shall consider, in turn, the expected values of the matrices \( \mathbf{A}_{u^*}^2 \), \( \mathbf{A}_{u^*}' \mathbf{\Gamma}_{u^*} \), and \( \mathbf{\Gamma}_{u^*}' \mathbf{\Gamma}_{u^*} \).
First, notice that $\Lambda^2_{u^*}$ is diagonal with entries $u_i^2$ on the diagonal. Thus,

$$E^d[(\Lambda^2_{u^*})_{ij}] = s_d^{-2} I_{k_d}. \quad (B.1.2)$$

Next, fix $k \in \{1, \ldots, k_d\}$ and $h_{ij} \in \{1, \ldots, \binom{k_d}{2}\}$ where $h_{ij} \overset{i.i.d.}{\sim} (i,j)$ with $i < j \in \{1, \ldots, k_d\}$ (see Appendix A.1). The expected value of the $(k, h_{ij})$ element of $\Lambda_{u^*} \Gamma_{u^*}$ is equal to

$$E^d[(\Lambda_{u^*} \Gamma_{u^*})_{k,h_{ij}}] = E^d[u_i^* u_j^* \delta_{k_i} + u_k^* u_i^* \delta_{k_j}]$$

$$= E^d[u_i^* u_j^* (\delta_{ki} + \delta_{kj})]. \quad (B.1.3)$$

Since $i < j$, we get

$$E^d[(\Lambda_{u^*} \Gamma_{u^*})] = 0_{k_d \times \binom{k_d}{2}}. \quad (B.1.4)$$

Finally, fix $h^{(1)}_{ij} \leq h^{(2)}_{ij}$ where $h^{(1)}_{ij}, h^{(2)}_{ij} \in \{1, \ldots, \binom{k_d}{2}\}$ and $h^{(m)}_{ij} \overset{i.i.d.}{\sim} (i_m, j_m)$, for $m = 1, 2$. The expected value of the $(h^{(1)}_{ij}, h^{(2)}_{ij})$ element of $\Gamma_{u^*} \Gamma_{u^*}$ equals

$$E^d[(\Gamma_{u^*}^t \Gamma_{u^*})_{h^{(1)}_{ij}, h^{(2)}_{ij}}] = \sum_{k=1}^{k_d} E^d[(u_i^* \delta_{kj_1} + u_j^* \delta_{kj_2})(u_i^* \delta_{kj_2} + u_j^* \delta_{kj_1})]$$

$$= E^d[u_i^* u_j^* \delta_{j_1j_2} + u_i^* u_j^* \delta_{j_2j_1} + u_i^* u_j^* \delta_{j_1j_2}]$$

$$= 2s_d^{-2} \delta_{h^{(1)}_{ij} h^{(2)}_{ij}}. \quad (B.1.5)$$

Hence,

$$E^d[\Gamma_{u^*}^t \Gamma_{u^*}] = 2s_d^{-2} I_{\binom{k_d}{2}}. \quad (B.1.6)$$
From (B.1.2), (B.1.4) and (B.1.6), we see the matrix $\mathbf{E}^r_{\mathbf{d}}[\mathbf{H}'\mathbf{H}]$ is block diagonal with diagonal matrices in the blocks. The value of $\mathbf{E}^r_{\mathbf{d}}[\mathbf{H}'\mathbf{H}]$ is

$$
\begin{pmatrix}
0 & 0_{k_d}^t & 0_{k_d} & 0_{(k_d+1)k_r}^t & 0_{(k_d+1)k_r} \\
0_{k_d} & \frac{1}{4}\mathbf{I}_{k_d} & 0_{k_d\times k_d} & 0_{k_d\times \left(\begin{smallmatrix}k_d \\ 2 \end{smallmatrix}\right)} & 0_{k_d\times \left(\begin{smallmatrix}k_d \\ 2 \end{smallmatrix}\right)} \\
0_{k_d} & 0_{k_d\times k_d} & \frac{s_d^{-2}}{2}\mathbf{I}_{k_d} & 0_{k_d\times \left(\begin{smallmatrix}k_d \\ 2 \end{smallmatrix}\right)} & 0_{k_d\times \left(\begin{smallmatrix}k_d \\ 2 \end{smallmatrix}\right)} \\
0_{\left(\begin{smallmatrix}k_d \\ 2 \end{smallmatrix}\right)} & 0_{\left(\begin{smallmatrix}k_d \\ 2 \end{smallmatrix}\right)\times k_d} & 0_{\left(\begin{smallmatrix}k_d \\ 2 \end{smallmatrix}\right)\times k_d} & \frac{s_d^{-2}}{2}\mathbf{I}_{\left(\begin{smallmatrix}k_d \\ 2 \end{smallmatrix}\right)} & 0_{\left(\begin{smallmatrix}k_d \\ 2 \end{smallmatrix}\right)\times \left(\begin{smallmatrix}k_d+1 \\ k_r \end{smallmatrix}\right)} \\
0_{\left(\begin{smallmatrix}k_d+1 \\ k_r \end{smallmatrix}\right)} & 0_{\left(\begin{smallmatrix}k_d+1 \\ k_r \end{smallmatrix}\right)\times k_d} & 0_{\left(\begin{smallmatrix}k_d+1 \\ k_r \end{smallmatrix}\right)\times k_d} & 0_{\left(\begin{smallmatrix}k_d+1 \\ k_r \end{smallmatrix}\right)\times \left(\begin{smallmatrix}k_d+1 \\ k_r \end{smallmatrix}\right)} & 0_{\left(\begin{smallmatrix}k_d+1 \\ k_r \end{smallmatrix}\right)\times \left(\begin{smallmatrix}k_d+1 \\ k_r \end{smallmatrix}\right)}
\end{pmatrix}
\quad (B.1.7)
$$

**B.2 Proof of Lemma 3.6.5**

Assume the prior distribution $\pi^*_{\mathbf{d}}$ for which the hypotheses in Assumptions 3.2.2 are satisfied. It is easy to show that the integrand $\mathbf{H}'\mathbf{M}^{-}(\xi^*)\mathbf{H}$ is equal to

$$
\mathbf{H}'\mathbf{M}^{-}(\xi^*)\mathbf{H} = \Lambda_{\mathbf{u}^*}
\left[
\frac{1}{r-t} \left(\mathbf{I}_{k_d} - \frac{t-r^2}{r-r^2 + (k_d-1)(t-r^2)} \mathbf{J}_{k_d}\right)
\right] \Lambda_{\mathbf{u}^*} + \frac{1}{4r} \mathbf{I}_{k_d} + \frac{1}{4t} \Gamma_{\mathbf{u}^*} \Gamma'_{\mathbf{u}^*}
\quad (B.2.1)
$$

Hence, we can determine the expected value by finding, in turn, the expected values of the three matrices on the right hand side of (B.2.1).

First, we find the expected value of the big matrix on the right hand side of (B.2.1), i.e.,

$$
\mathbf{E}^r_{\mathbf{d}} \left[ \Lambda_{\mathbf{u}^*} \left[ \frac{1}{r-t} \left( \mathbf{I}_{k_d} - \frac{t-r^2}{r-r^2 + (k_d-1)(t-r^2)} \mathbf{J}_{k_d}\right) \right] \Lambda_{\mathbf{u}^*} \right]
= \frac{1}{r-t} \mathbf{E}^r_{\mathbf{d}}[\Lambda_{\mathbf{u}^*}] - \frac{t-r^2}{(r-t)[r-r^2 + (k_d-1)(t-r^2)]} \mathbf{E}^r_{\mathbf{d}}[\Lambda_{\mathbf{u}^*} \mathbf{1}_{k_d} \mathbf{1}_d' \Lambda_{\mathbf{u}^*}']
$$
Next, we find $E^\pi_d[\Gamma_u \cdot \Gamma_u']$. Considering the matrix in (3.4.14), we see

$$
(\Gamma_u \cdot \Gamma_u')_{ij} = \left\{ \begin{array}{ll}
\frac{k_d}{u_i^* u_j^*} & \text{if } i = j \\
\sum_{m \neq i} u_m^2 & \text{if } i \neq j
\end{array} \right. \text{ for } i, j \in \{1, 2, \ldots, k_d\}.
$$

(B.2.3)

This implies

$$
E^\pi_d[\Gamma_u \cdot \Gamma_u'] = s_d^{-2}(k_d - 1)I_{k_d}.
$$

Hence,

$$
E^\pi_d[HM^{-1}(\xi^*)H^t] = \left\{ \begin{array}{c}
\frac{1}{4r} + \frac{s_d^{-2}}{r - t} \left( \frac{r - r^2 + (k_d - 2)(t - r^2)}{r - r^2 + (k_d - 1)(t - r^2)} \right) + \frac{s_d^{-2}(k_d - 1)}{4t}
\end{array} \right\} I_{k_d}.
$$

(B.2.5)

### B.3 The IavD criterion with respect to the prior $\pi^*$

In this section we show that the IavD-optimal design is unchanged if instead of a prior $\pi_d^*$ for which the hypotheses in Assumptions 3.2.2 are satisfied, we consider a prior $\pi^*$ for which the hypotheses in Assumptions 3.2.1 are satisfied. This result further demonstrates the similarities between the IavA and IavD design criteria. However, for the IavD criterion, we are not required to quantify the ratio of the two eigenvalues for the positive definite, completely symmetric matrix $Q$ in (3.2.1). We state the result of this section in the form of a lemma.
Lemma B.3.1 Assume the model in (3.3.2), and let \( \pi_d^* \) be a prior distribution for which the hypotheses in Assumptions 3.2.1 are satisfied. A design \( \xi^* \in \Xi_1 \) is IavD-optimal for the prior \( \pi^* \) if the conditions in Theorem 3.6.1 are satisfied.

Proof of Lemma B.3.1: We need to prove that the design measure \( \xi^* \) given in Theorem 3.6.1 minimizes

\[
|E^\pi^* \left[ A_d^{-1}HM(\xi)^{-1}H^tA_d^{-1} \right] |
\]  

over \( \Xi_1 \). Let \( f \) be the vector of regression functions of \((u, v)\) in (3.4.6). Then, using Condition ii.) of Theorem 1.3.4, a design measure \( \xi \in \Xi_1 \) is IavD optimal with respect to the prior \( \pi^* \) if and only if

\[
\text{tr} \left\{ E^\pi^* \left[ A_d^{-1}HM(\xi)^{-1}H^tA_d^{-1} \right] \left( E^\pi^* \left[ A_d^{-1}HM(\xi)^{-1}H^tA_d^{-1} \right] \right)^{-1} \right\} \leq k_d \]  

for every \((u, v) \in [-1, 1]^{kd+k_d^*}\) (see Example 1.3.2).

Let us consider the matrix \( E^\pi^* \left[ A_d^{-1}HM(\xi)^{-1}H^tA_d^{-1} \right] \) that is being inverted on the left hand side of (B.3.2). The matrix \( A_d^{-1} \) is constant with respect to the prior \( \pi_d^* \); therefore, we can express this matrix as

\[
E^\pi^* \left[ A_d^{-1}HM(\xi^*)H^tA_d^{-1} \right] = E^{\pi^*}_{b} \left[ A_d^{-1}E^{\pi^*}_{d}[HM(\xi^*)]A_d^{-1} \right]. \]  

From Lemma 3.6.5, we know the matrix \( E^{\pi^*}_{d}[HM(\xi^*)] \) is a scalar multiple of the identity. After some manipulation, the inequality in (B.3.2) can be reexpressed as

\[
 f^tM(\xi^*)E^\pi^* \left[ H^tA_d^{-1} \left( E^{\pi^*}_{b}[A_d^{-2}] \right)^{-1}A_d^{-1}H \right]M(\xi^*)f \leq \frac{k_d \left\{ \frac{1}{4r} + \frac{s_d^{-2} \left( r^2 + (k_d - 2)(t - r^2) \right)}{r - t} + \frac{s_d^{-2}(k_d - 1)}{4t} \right\} }{k_d \left( \frac{r}{r - t} \left( r^2 + (k_d - 1)(t - r^2) \right) \right)}. \]  

(B.3.4)
Notice (B.3.4) is very similar to (3.6.21). If we can verify that
\[ E^\pi^* \left[ H^t A_d^{-1} \left( E^\pi^* [A_d^{-2}] \right)^{-1} A_d^{-1} H \right] = E^\pi^* [H^t H], \] (B.3.5)
then the two equations are identical, and we have proved the lemma.

We have encountered a matrix similar to \( E^\pi^* \left[ H^t A_d^{-1} \left( E^\pi^* [A_d^{-2}] \right)^{-1} A_d^{-1} H \right] \), e.g., the matrix \( E^\pi^* \left[ H^t A_d^{-2} H \right] \). In Appendix A.1, we showed that
\[ E^\pi^* [A_d^{-2}] = \Lambda_Q, \] (B.3.6)
implied the matrix \( E^\pi^* \left[ H^t A_d^{-2} H \right] \) was also diagonal (see Lemma 3.5.2). Furthermore, the values along the diagonal of \( \Lambda_Q \) appeared in the matrix \( E^\pi^* \left[ H^t A_d^{-2} H \right] \) in a well defined way.

For the matrix \( H^t A_d^{-1} \left( E^\pi^* [A_d^{-2}] \right)^{-1} A_d^{-1} H \), we have that
\[ E^\pi^* \left[ A_d^{-1} \left( E^\pi^* [A_d^{-2}] \right)^{-1} A_d^{-1} \right] = E^\pi^* \left[ A_d^{-2} \left( E^\pi^* [A_d^{-2}] \right) \right]^{-1} = I_{kd}. \] (B.3.7)
The first equality is true, since \( A_d^{-1} \) is a completely symmetric matrix (see Corollary 1.3.18). It follows that the matrix \( E^\pi^* \left[ H^t A_d^{-1} \left( E^\pi^* [A_d^{-2}] \right)^{-1} A_d^{-1} H \right] \) equals
\[
\begin{pmatrix}
0 & 0^{t\ell}_{k_d} & 0^{t\ell}_{k_d} & 0^{t(\ell)'\ell}_{(k_d+1)k_\pi} & 0^{t\ell}_{(k_d+1)k_\pi} \\
0_{k_d} & \frac{1}{4} I_{k_d} & 0_{k_d \times k_d} & 0_{k_d \times (\ell)'\ell} & 0_{k_d \times (k_d+1)k_\pi} \\
0_{k_d} & 0_{k_d \times k_d} & s_d^{-2} I_{k_d} & 0_{k_d \times (\ell)'\ell} & 0_{k_d \times (k_d+1)k_\pi} \\
0_{(\ell)'\ell} & 0_{(\ell)'\ell} \times k_d & 0_{(\ell)'\ell} \times k_d & \frac{s_d^{-2}}{4} I_{(\ell)'\ell} & 0_{(\ell)'\ell} \times (k_d+1)k_\pi} \\
0_{(k_d+1)k_\pi} & 0_{(k_d+1)k_\pi} \times k_d & 0_{(k_d+1)k_\pi} \times k_d & 0_{(k_d+1)k_\pi} \times (\ell)'\ell} & 0_{(k_d+1)k_\pi} \times (k_d+1)k_\pi} \\
\end{pmatrix}. \] (B.3.8)
This verifies (B.3.5); hence, this lemma is proved. □
BIBLIOGRAPHY


