The Use of Correlated Simulation Experiments in Response Surface Optimization

by

Joan M. Donohue

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Management Science

APPROVED:

Ernest C. Houck, Chairman

Laurence J. Moore

Loren P. Rees

Raymond H. Myers

Bernard W. Taylor, III

August, 1988

Blacksburg, Virginia

The Use of Correlated Simulation Experiments in Response Surface Optimization

by

Joan M. Donohue

Ernest C. Houck, Chairman

Management Science

(ABSTRACT)

Response surface methodology (RSM) provides a useful framework for the optimization of stochastic simulation models. The sequential experimentation and model fitting procedures of RSM enable prediction of the response and location of the optimum operating conditions. In a simulation environment, the experimentation phase of RSM involves selecting the input variable levels for each simulation run and assigning pseudorandom number streams to the stochastic model components. Through an appropriate assignment of random number streams to simulation runs, correlation among the simulated responses can be induced, thereby affecting reductions in the variances of certain model coefficients. Three methods of correlation induction are considered in this research: (i) no correlation induction, achieved through the use of independent streams, (ii) positive correlation induction, achieved through the use of common streams, and (iii) a combination of positive and negative correlation induction, achieved through the use of the assignment rule blocking strategy.

The performance of the correlation induction strategies is evaluated in terms of two mean squared error design criteria; *MSE of response* and *MSE of slope*. The MSE of slope criteria is useful in the early stages of RSM, when the experimental objective is location of the region containing the optimum. The MSE of response criteria is useful in the latter stages of RSM, when the experimental objective is prediction of the optimum response. The correlation induction strategies are evaluated under two experimental situations; fitting a first order model while protecting against quadratic curvature in the response surface, and fitting a second order model while protecting against cubic curvature. In the case of fitting a first order model, two-level factorial designs are used

to evaluate the correlation induction strategies, and in the second order case, four design classes are considered; central composite designs, Box-Behnken designs, three-level factorial designs, and small composite designs.

The findings of this research indicate that the assignment rule blocking strategy generally performs the best of the three strategies under both MSE criteria, and the performance of this strategy improves as the magnitudes of the induced correlations increase. The independent streams strategy is a poor choice when the design criteria is MSE of slope and the common streams strategy is a poor choice when the design criteria is MSE of response. The central composite and Box-Behnken designs were found to perform the best of the four second order design classes. The three-level factorial designs performed poorly under MSE of response criteria and the small composite designs performed poorly under the MSE of slope criteria.

Acknowledgements

The author would like to express sincere appreciation to all of her committee members for guidance and encouragement throughout her graduate studies at Virginia Tech. She is also grateful to the Graduate School for awarding her a Cunningham Dissertation Fellowship for the 1987-1988 academic year. Many individuals have implicitly contributed to this research, and at this time the author would like to extend her appreciation to each of them.

The author would like to thank her chairman, Dr. Ernest C. Houck, for the encouragement and support he has provided throughout her graduate studies at Virginia Tech. He has been an excellent course advisor and committee chairman, and couldn't have been more trusting, patient, and understanding during the dissertation stage of her doctoral program.

The author thanks Dr. Bernard W. Taylor, III, Head of the Department of Management Science, for providing her with excellent research facilities and monetary support for her graduate work at Virginia Tech. Dr. Taylor's achievements in management science, particularly his research and publication vita, will serve as an inspiration to the author in her future academic career.

The author thanks Dr. Laurence J. Moore for his guidance and assistance throughout her graduate studies at Virginia Tech. Dr. Moore's friendly and congenial nature, along with his commendable achievements in the field of management science, are certainly inspirational to the author.

The author thanks Dr. Loren P. Rees for generating her initial interest in management science during her first year as an M.B.A. student at Virginia Tech. Dr. Rees was influential in the author's decision to earn a doctorate degree and in the author's choice of a dissertation topic.

The author thanks Dr. Raymond H. Myers for sharing his knowledge and enthusiasm for response surface methodology with her. Dr. Myers' expertise was indispensable when it came to deciding on a topic for her dissertation research.

The author thanks Dr. Robert S. Schulman and Dr. Jeffrey B. Birch for generating her initial interest in statistics. They gave the author confidence in her statistical abilities, which eventually led her to pursue a statistically-related dissertation topic.

Finally, the author would like to thank for his friendship and support throughout graduate school.

Table of Contents

List of Illustrations

List of Tables

Chapter 1

INTRODUCTION

Computer simulation studies offer a useful approach to modeling business and economic systems when it is impossible or impractical to perform controlled experiments on the actual system. A real-world system can be imitated, or *simulated,* on a computer using a model of the actual system. A model is a set of assumptions about how the system works and usually takes the form of mathematical or logical relationships. If the relationships are fairly simple, then an exact analytic solution may be possible. However, models are frequently too complex to be solved analytically and, therefore, computer simulation can be a useful alternative for studying many real-world systems.

Banks and Carson (1) define simulation as the imitation of the operation of a real-world system over time. A computer is programmed to perform the procedural steps of a simulation model and the output data represents an artificial history of the actual system. The simulation-generated data can be used to estimate the true characteristics of the system, draw inferences concerning the operating characteristics of the system, and estimate performance measures of the system.

Law and Kelton (37) distinguish between types of computer simulation models on the basis of three dichotomous characteristics:

1. Static / Dynamic models:

   Static simulation models are a representation of a system at a particular point in time, whereas dynamic models represent a system as it evolves over time.

2. Deterministic / Stochastic models:

   Deterministic simulation models contain no random variables and produce a unique set of output data for a given set of input variables. Stochastic models contain one or more random variables and the random output data only estimates the true characteristics of the system.

3. Discrete / Continuous models:

   Discrete and continuous models are types of dynamic simulation models in which the system is modeled over time. In discrete models, the variables can only change at specific points in time, whereas the variables can change continuously over time in continuous models.

The simulation models of concern in this paper are dynamic, stochastic, and discrete, frequently referred to as *discrete-event simulation models.*

The often-seen term, Monte Carlo, is defined by Hammersley and Handscomb (22) as the branch of experimental mathematics which is concerned with experiments on random numbers. Distribution sampling refers to experiments which use Monte Carlo techniques but do not involve the passage of time. These sampling experiments are static in nature and are not considered in this research. Stochastic simulation, on the other hand, refers to experiments which utilize Monte Carlo techniques and are dynamic in nature. The focus of this research is on stochastic simulation experiments involving both the passage of time and the use of Monte Carlo techniques.

The starting point for any computer simulation study is the development of a mathematical model of the actual system. Data for the simulation study can then be generated using Monte Carlo techniques and a mathematical model of the real-world system. Naylor (50) lists four basic steps involved in any computer simulation study:

1. Formulation of the computer program.
2. Validation of the mathematical model.
3. Choice of the experimental design plan.
4. Analysis of the simulated data.

This research is primarily concerned with the third step: choice of the *experimental design* for computer simulation studies. In a simulation context, the *experiment* is the execution of the computer simulation model for a specified setting of the input variables and the *design* refers to the plan used for selection of the input variable settings. The choice of experimental design depends on the goals of the experimenter, or equivalently, the criteria chosen for evaluation of the design. Because of the method by which random behavior is achieved in simulation studies, it is possible to induce a correlation structure among simulated responses which, in turn, affects the choice of the experimental design plan. This capability of inducing correlations among response observations is generally not available in other experimental situations and represents an added dimension in the design of computer simulation experiments.

Schruben and Margolin (56) suggest that authors of textbooks on Monte Carlo methods (20), (22), (33), (39), (49) appear to have accepted the premise that either experimental design procedures or correlation induction techniques can be beneficially employed in simulation studies, but not both simultaneously. Schruben and Margolin demonstrate the benefits that accrue when the two techniques are combined in stochastic simulation studies. Their research evaluated three methods of inducing correlations among simulated responses in the context of parameter estimation of the general linear model. Two variance-related design criteria were used to evaluate the three methods of correlation induction. This research extends the work of Schruben and Margolin by further exploring the opportunities that arise when incorporating both experimental design and correlation induction techniques as an integral part of simulation studies.

A simulation study based on random behavior naturally requires a mechanism for the generation of random responses. The response generating mechanism that is generally used in simulation produces a set of independent random variates, each with a uniform distribution on the [0,1] interval. The most commonly employed random number generation algorithm produces a nonrandom sequence of numbers, each number being determined by its predecessor, and consequently all numbers being determined by the initial number, or *seed* value. If the parameters of the generator are chosen carefully, then the streams of numbers are sufficiently uniform and independent for many practical purposes. To emphasize the inherently nonrandom characteristic of the uniform variates, these randomly generated streams of numbers are termed *pseudorandom* number streams.

In a simulation environment, the pseudorandom number streams used to generate values for the stochastic model components can be selectively controlled. A technique which beneficially utilizes this capability is termed a variance reduction technique (VRT). By appropriately assigning sets of random number streams to design points, a correlation structure among the simulated responses can be induced, thereby affecting reductions in the variances of certain parameter estimates. A VRT does not rely on a statistical analysis of the input and output variables of a simulation study, but instead relies on a reorganization of the simulation itself. If a simulation model is thought of as a black box representing a real-world system, then a statistical analysis looks at the black box from the outside only, but a VRT gets inside the black box and tampers with the internal mechanisms. The two VRTs which have received the greatest attention in the simulation literature are the assignment of common random number streams (CRN) and the assignment of antithetic random number streams (ARN) to the process generating mechanisms of a simulation model on different simulation runs. These two techniques, as well as the assignment of independent random number streams (IRN), are the basis of the correlation induction strategies evaluated in this research.

The IRN technique uses a unique set of pseudorandom number streams to generate each of the simulated responses, and therefore no correlation is induced among the sample responses. Suppose $\overline{Y}_1$ and $\overline{Y}_2$ denote the mean sample responses on two different simulation runs, then for the IRN technique the variance of the difference in these two mean responses becomes

$$\text{Var}(\overline{Y}_1 - \overline{Y}_2) = \text{Var}(\overline{Y}_1) + \text{Var}(\overline{Y}_2) \ . \qquad [1.1]$$

Suppose $\overline{Y}$ denotes the average of the mean responses, $\overline{Y}_1$ and $\overline{Y}_2$, then the variance of this average sample response becomes

$$\text{Var}(\overline{Y}) = \tfrac{1}{4}\left[\text{Var}(\overline{Y}_1) + \text{Var}(\overline{Y}_2)\right] . \qquad [1.2]$$

The variances of these sample response measurements do not include any covariance terms because no correlation is induced with the IRN technique and the individual sample responses are considered to be statistically independent.

The CRN technique induces positive correlation among the sample responses by using a common set of pseudorandom number streams to generate two or more sample responses. The use of CRN is based on the premise that experiments should be performed under as nearly homogeneous circumstances as possible. The technique is applicable when an experimenter is interested in comparing mean responses or estimating linear contrasts. The statistical justification for CRN is that the variance of the difference between two mean responses,

$$\text{Var}(\overline{Y}_1 - \overline{Y}_2) = \text{Var}(\overline{Y}_1) + \text{Var}(\overline{Y}_2) - 2\,\text{Cov}(\overline{Y}_1, \overline{Y}_2) \ , \qquad [1.3]$$

is reduced due to the positive correlation induced; that is, $\text{Cov}(\overline{Y}_1, \overline{Y}_2) > 0$, when the CRN technique is used.

The ARN technique induces negative correlation between pairs of sample responses through the use of complementary sets of pseudorandom number streams. The (0,1) uniform variate $\{r\}$ is used to generate one observation and its *antithetic* variate $\{1 - r\}$ is used to generate the paired observation, thereby inducing a negative correlation between the pair of responses. The logic behind the use of antithetic random number streams is that one observation in the pair should be large and the other observation should be small, thus offsetting each other and inducing a negative correlation between the pair of responses. The ARN technique applies to the estimation of the average response $\overline{Y}$ at a particular setting of the input variables. Suppose $\overline{Y}_1$ and $\overline{Y}_2$ are two estimates of the mean response at a specific setting of the input variables, then the variance of the average sample response $\overline{Y}$,

$$\mathrm{Var}(\overline{Y}) = \tfrac{1}{4}\left[\mathrm{Var}(\overline{Y}_1) + \mathrm{Var}(\overline{Y}_2) + 2\,\mathrm{Cov}(\overline{Y}_1, \overline{Y}_2)\right], \qquad [1.4]$$

is reduced due to the negative correlation induced; that is, $\mathrm{Cov}(\overline{Y}_1, \overline{Y}_2) < 0$, when the ARN technique is used.

It should be noted that there are many other VRTs applicable to simulation experiments, such as control variates, importance sampling, regression sampling, stratified sampling, and quasi-random numbers, which are not addressed in this research. See references (23), (33), (37), (44), (61) for discussions of these techniques.

In addition to the applicability of VRTs to simulation studies, experimental design procedures can be beneficially employed as well. Experimental design, the process of planning the experiment so that appropriate data will be collected, is often thought of as a subject dealing with problems and techniques for designing real-world experiments. However, the techniques used in experimental design are also relevant to the design of computer simulation experiments. In a simulation context, experimental design procedures provide a way of deciding beforehand, "where to take" the simulation runs, so that the desired information can be obtained. Thus, experimental design enables the experimenter to achieve the desired goals through specific choice of the input variable settings.

The general purpose of an experiment (to obtain the desired information) is not sufficient for designing an experimental plan. Therefore, statisticians have developed various *design criteria*, and the experimenter can choose one according to the purposes of the study. The design criteria might be to provide economy in the number of experimental runs, achieve desirable confounding patterns, obtain minimum variance estimators, minimize the mean squared error of prediction, or other objectives (45), (51), (57). As the experimenter learns more about the behavior of the actual system through the simulation experiments, his or her goals often become more precise, with the final objective frequently being a search for the conditions that maximize or minimize response.

The problem of optimization in simulation is an extremely difficult one and has received increased attention in recent years (2), (11), (41), (42), (43), (53). One optimization strategy that is often cited is response surface methodology (RSM); a collection of statistical and mathematical procedures specifically designed for the empirical estimation of optimum system operating conditions. A response surface is simply the expected response of a system (output measure of the system's performance) viewed as a function of the factors (input variables of the simulation model). Once a simulation model describing the actual system has been developed, RSM can be an effective vehicle for constructing and parameterizing optimization models. Through sequential experimentation and model fitting, the values of the input variables that optimize the simulated response can be estimated. The experimental investigation of a response surface is generally divided into two phases; a design phase and an analysis phase. In the design phase, decisions are made to set the factors at certain levels for a given simulation run. This is equivalent to the selection of the design points. In the analysis phase, the response data from the experimental runs is used to answer questions regarding the operating conditions of the system. Since the eventual goal of RSM is usually optimization, it is extremely important that proper experimental design plans be chosen so that an accurate estimate of the true optimum can be obtained.

Myers (46) gives two typical goals of an experimenter using response surface methods:

1. Find a suitable approximating function for the purpose of predicting future response.
2. Determine the values of the factors which are optimum as far as the response is concerned.

Achieving either of these goals requires determining the relative contributions of the input variables to the response. Response surface methods assume that the relationship between the response and the input variables can be approximated with low-order polynomial models within restricted regions of the factors. First order polynomial models are typically used to study the response data in small regions of the factor space where little curvature in the response function is expected. Second order polynomial models are generally used in larger regions where there may be quadratic curvature in the response function, or when the region containing the optimum response has been tentatively located. Models of order greater than two are sometimes used when the region of the optimum response has been tentatively identified and cubic curvature in the response function is indicated.

As noted previously, RSM experiments are initially performed in small regions of the factor space using first order polynomial models to fit the response data. These experiments are typically performed sequentially until the region which contains the optimum response has been located. Then, to enable more accurate prediction and/or optimization, a second order polynomial model is fit to the response data. An appropriate experimental design plan must be developed for each sequential set of experiments performed during the optimization process. The choice of an experimental design plan is affected by a variety of factors. These include:

1. Order of the fitted response model.
2. Design criteria chosen by the experimenter.
3. Structure of the correlation matrix of the sample responses.

The focus of this research is the evaluation of three strategies for inducing correlation among the sample responses of a simulation model. These strategies are evaluated for the cases of fitting first and second order polynomial models using two *mean squared error* design criteria.

The design criteria of integrated mean squared error of response, which takes variance of the predicted response and bias due to model misspecification into account, was first proposed by Box and Draper (5) for a spherically-shaped region in the factor space. Later, Draper and Lawrence (16) extended the criteria to cuboidal-shaped regions and recently Draper and Guttman (15) extended the criteria to a family of flexible region shapes. A generalization of the mean squared error criteria, proposed by Myers and Lahoda (48), involves minimizing the integrated mean squared error of the slopes of the response function for either a spherical or cuboidal region of interest. Throughout this research the Box and Draper criteria is referred to as the *MSE of response* criteria and the Myers and Lahoda criteria is referred to as the *MSE of slope* criteria.

Three specific correlation induction strategies are evaluated in this research using these two MSE design criteria; minimum MSE of response and minimum MSE of slope. The first correlation induction strategy is that of inducing no correlation among the sample responses, achieved by using independent random number streams (IR) for all design points. The second strategy involves the induction of positive correlations among the sample responses, achieved by using common random number streams (CR) for all non-replicated design points. The third strategy is the *assignment rule* (AR), a method proposed by Schruben and Margolin (56) which can be applied to orthogonally blockable experimental designs. The AR strategy involves the induction of positive correlations among responses in the same blocks and the induction of negative correlations between responses in different blocks. The strategy is achieved by using common sets of random number streams for non-replicated design points within orthogonal blocks, and using antithetic sets of streams for non-replicated design points in the opposite blocks. Hussey, Myers, and Houck (28), (29) evaluated these three correlation induction strategies for both first and second order response surface models using four variance-related design criteria. This research extends their work in the following two ways:

1. Incorporation of bias into the design criteria.
2. Inclusion of designs with replicated center runs.

The objective of this research is to investigate the performance of the three correlation induction strategies on response surface designs which are applicable to a wide variety of simulation studies. Relevant research findings in the areas of simulation and response surface methodology are reviewed in Chapter 2. The correlation induction strategies and MSE design criteria are fully presented in that chapter. In Chapter 3 the correlation induction strategies are evaluated for first order response surface designs and in Chapter 4 second order designs are considered. An overview of this research is presented in Chapter 5.

Chapter 2

LITERATURE REVIEW

2.1 Computer Simulation Methodology

Simulation is a computer-based numerical technique for the experimental study of a stochastic or deterministic process over time. Simulation studies are generally undertaken with a goal of learning something about the actual process. Once the goal is clearly identified, the experimental process is thoroughly researched and a valid simulation model is developed. The simulation model can then be used to generate the data necessary to determine the relationship between the input variables and some specified output variable. The values of the input variables, or *factors*, are specified by the experimenter and can be easily changed from one simulation run to the next. The output variable, or *response*, is generated by the simulation experiment and is generally a performance measure related to the real-world system being studied. Typically, the goal of an experimenter is the prediction or optimization of response, which can be accomplished by determining the relationship between the factors and the response variable. Krasnow (36) distinguishes among three methods of observing a simulated response variable:

1. The response can reflect the *model state* at certain prespecified times, or under certain prespecified conditions, such as the value of the response when the simulated time is t minutes or the queue size is equal to n.

2. The response can be observed as a *time series*, such as the value of the response every t minutes, and the results related to simulated time using time series analysis techniques.

3. The response can be some *statistical evaluation* of observations over time, such as the average inventory in the system during a specified time period.

This paper is primarily concerned with the third type of response variable measurement, and specific attention is given to the *average* value of the response observations during the simulation runs. Statistical evaluations of the response variable are generally easy to obtain because most computer simulation packages include, as standard output, statistical summary measures of the response observations.

2.1.1 Generation of Simulated Responses

The output from a discrete-event simulation consists of a random sample of simulated responses for each specific setting of the input variables. The values of the input variables, or factors, are fixed by the experimenter and are easily changed from one run to the next. The values of the response, or output variable, are randomly generated and recorded throughout the simulation runs. When the simulated responses are recorded at equally spaced time intervals during the simulation runs, the simulation can be viewed as a discrete time series, say

$$\{ Y_{ut} \;; \quad u = 1, 2, \ldots, N \quad t = 1, 2, \ldots, T \}$$

where N is the number of specific settings of the factors, or equivalently, the number of simulation runs,

T is the number of observed responses during each simulation run, and

Y_{ut} is the t^{th} observed response for the u^{th} setting of the factors.

The statistical evaluation of interest is frequently the average of the observations in the series. The average value of *T* response observations, denoted by $\overline{Y}_u$, or simply Y_u , is the sample response for the u^{th} setting of the factors. For processes which are second order stationary (expected value and covariances of the Y_{ut}'s are finite), Schruben and Margolin (56) suggest that the expected value of the response be estimated by the average of the series. The sample response for the u^{th} simulation run of a second order stationary process becomes

$$Y_u = \frac{1}{T} \sum_{t=1}^{T} Y_{ut} . \qquad [2.1.1]$$

This sample response is not necessarily computed using the first *T* time periods. The experimenter may choose to record the first response observation far enough into the run to eliminate transient start-up conditions, or for variance considerations, the experimenter may choose to record observations during a specific time period after steady-state conditions have been reached. The sample response values obtained from a number of simulation runs can then be used as the values of the dependent variable in the regression models developed for optimization. The independent variables of the regression models are the set of *k* input variables, whose values for the u^{th} simulation run are specified by the design point

$$\mathbf{x}_u' = [\, x_{u1} , x_{u2} , \ldots , x_{uk} \,] .$$

The selection of *N* experimental design points constitutes the traditional experimental design. A simulation run is made for each design point and a set of *N* sample response values is obtained. The set of *N* sample responses and corresponding design points are used to estimate the parameters of the regression models developed for prediction of the system response.

2.1.2 Pseudorandom Number Streams

A simulation study which models a stochastic process involves the use of random variables generated from one or more probability distributions. For example, the simulation of a multiple-server queueing system may require specification of the distribution of customer interarrival times, the probability of taking a particular path through the system, and the distribution of service times. Assuming that observed data on the actual system is available, the probability distributions for the stochastic model components can be specified using one of two general approaches:

1. Fit the observed data to a theoretical distribution form, such as an exponential, normal, gamma, or Poisson distribution.

2. Use the observed data to define an empirical distribution (typically used when a theoretical distribution which adequately fits the data cannot be found).

For a complete discussion on the selection of input probability distributions, see Law and Kelton (37). This research assumes that the experimenter has chosen appropriate input distributions for the simulation model. Random variables from these distributions can then be generated by transformations on independent, uniformly-distributed random variables on the [0,1] interval. Most computer installations and simulation languages have algorithms for affecting these transformations. Fishman (20) discusses the subroutines available in many simulation languages and presents theoretically exact methods for about twenty-five probability distributions.

Although the uniform (0,1) probability distribution is the simplest continuous distribution, its role in computer simulation studies is extremely important. The prominent role of this distribution stems from the fact that random variables from all other distributions, as well as realizations of various random processes, can be obtained from uniform (0,1) deviates through the use of appropriate transformations. The uniform deviates produced by random number generators are called *pseudorandom* numbers because of the inherently nonrandom characteristics of the generation processes. Typically, the random numbers are generated in sequence, and each number is determined by one or more of its predecessors according to a fixed mathematical formula. Therefore, once an initial seed value is chosen, the entire sequence of random numbers is completely determined. Most arithmetic generators can produce random numbers which appear to be independent draws from the uniform (0,1) distribution. Descriptions of various random number generators and statistical techniques for testing random number sequences for uniformity and independence are presented in references (20), (22), (37).

Simulation studies generally utilize many pseudorandom number streams simultaneously. Each stream of numbers generated is assumed to be a set of random draws from the uniform (0,1) probability distribution. However, because of the methods by which these uniform variates are generated, each sequence of numbers is completely determined by its seed value. For a given seed value, a stream of pseudorandom numbers is produced, denoted as

$$R = (r_1\ ,\ r_2\ ,\ r_3\ , \ldots\ldots) .$$

If a simulation study involves the use of g pseudorandom number streams to drive the stochastic components of the model, this *set* of streams, based on g seed values, can be denoted as

$$\mathbf{R} = (R_1\ ,\ R_2\ , \ldots,\ R_g) .$$

In an experimental design context, simulation runs are considered design points, and each design point utilizes a set of g pseudorandom number streams to generate values for the g stochastic model components. One sample response value is generated for each simulation run. For an experimental plan involving N simulation runs, N sets of g pseudorandom number streams must be chosen. If a different set of streams is used on each run, then the simulated responses should be independent from one run to the next. However, if an experimenter appropriately assigns specific sets of streams, then positive and/or negative correlation can be induced between the responses. Techniques which involve the manipulation of stream assignments to simulation runs are commonly known as variance reduction techniques (VRTs) because *planned* correlation induction can affect a reduction in the variability of the sample responses, thereby allowing for improved parameter estimation in the statistical models developed for estimation of the true system response.

2.1.3 Variance Reduction Techniques

Simulation experiments are frequently complicated by the presence of large experimental error variances. Variance reduction techniques (VRTs) are strategies developed for the resolution of this problem. These procedures enable the experimenter to realize increased precision in the estimation of parameters, while using the same number of simulation runs. For example, if two simulation studies are performed on the the same model, using the same parameters for the input distributions, the same method of observing responses, and the same number and length of simulation runs, then the simulation study utilizing an appropriate VRT should result in smaller response variances than the same study performed without utilizing the VRT. The focus of the VRT research in the literature is typically restricted to the estimation of the average response value and, in appropriate situations, these techniques are capable of reducing the variance of the estimated response without altering its expected value. Some of the VRTs which have been developed include: common random numbers, antithetic random numbers, control variates, importance sampling, regression sampling, stratified sampling, and quasi-random numbers. See references (20), (23), (33), (37), (44), (61) for descriptions of these techniques. These references also indicate the types of situations in which the techniques can be beneficially employed; that is, situations in which reductions in the variability of the simulated responses can be realized. The two most commonly used VRTs, and the two which are considered in this research, are common random numbers (CRN) and antithetic random numbers (ARN).

2.1.3.1 Independent Random Numbers

Frequently the reduction in variance achievable with a VRT is ascertained through a comparison of the response variance obtained using independent random numbers (IRN). The technique of IRN involves the use of a different set of pseudorandom number streams to generate values for the stochastic model components on each simulation run. Assuming that the random number generator used in the simulation study produces *independent* uniform (0,1) deviates, each of the simulated response values should be statistically independent. The IRN technique uses different sets of randomly-selected seed values for each simulation run and, therefore, the resulting time series samples, and hence their means, should be statistically independent. For example, if a simulation model consists of g stochastic model components, and a corresponding set of g different pseudorandom number streams are used to drive the first simulation run, then these streams can be denoted as

$$\mathbf{R}_1 = (R_{11}, R_{21}, \ldots, R_{g1})$$

where R_{j1} ($j = 1, \ldots, g$) is the pseudorandom number stream which generates values for the j^{th} stochastic model component on the 1^{st} simulation run (setting of the factors). If the second simulation run utilizes a different set of g streams to generate values for the stochastic model components, then these streams can be denoted as

$$\mathbf{R}_2 = (R_{12}, R_{22}, \ldots, R_{g2}).$$

The observed time series, $\{Y_{11}, \ldots, Y_{1t}\}$ and $\{Y_{21}, \ldots, Y_{2t}\}$, and the mean sample responses, Y_1 and Y_2, are typically uncorrelated when two different sets of pseudorandom number streams, $\mathbf{R}_1$ and $\mathbf{R}_2$, are used to drive the two simulation runs (20), (30), (33). These findings can be extended to a simulation study involving N simulation runs. If each of the N runs is made using a different set of pseudorandom number streams, $\mathbf{R}_u$ ($u = 1, \ldots, N$), then the sample responses from each pair of simulation runs are assumed to be uncorrelated, so that $\mathrm{Cov}(Y_u, Y_v) = 0$, $(u, v = 1, 2, \ldots, N;\ u \neq v)$. Therefore, the variance of the average sample response for the N runs becomes

$$\operatorname{Var}(\overline{Y}) = \operatorname{Var}\left[\frac{1}{N}\sum_{u=1}^{N} Y_u\right] = \frac{1}{N^2}\left[\sum_{u=1}^{N} \operatorname{Var}(Y_u)\right]. \qquad [2.1.2]$$

This average response variance does not include any covariance terms because the sample responses were generated using the IRN technique, which is generally assumed to produce independent responses.

2.1.3.2 Common Random Numbers

The first variance reduction technique used in the correlation induction strategies evaluated in this research, common random numbers (CRN), can be thought of as a way of introducing positive correlation between sample responses. The technique involves the use of the same set of pseudorandom number streams, or seed values, for generating values for the stochastic model components on two or more simulation runs. For example, if a simulation model consists of g stochastic model components, and a set of g different pseudorandom number streams are used to drive each simulation run, then the u^{th} set of streams can be denoted as

$$\mathbf{R}_u = (R_{1u}, R_{2u}, \ldots, R_{gu}).$$

By using the same set of g streams to generate values for the stochastic model components on a number of simulation runs, the sample responses tend to be positively correlated, so that $\operatorname{Cov}(Y_u, Y_v) > 0$, for each pair of sample responses (21), (25), (33), (34), (37), (56).

Law and Kelton (37) suggest that the CRN technique can be beneficially employed when an experimenter is interested in comparing two alternative systems, where Y_{1u} and Y_{2u} are the sample responses from the first and second systems on the u^{th} simulation run. For a simulation study consisting of N runs, the same set of N pseudorandom number streams would be used to drive both systems, denoted as

$$R = (R_1\ ,\ R_2\ , \ldots ,\ R_N\ )$$

where R_u is a set of g independent streams used to drive the u[th] simulation run of each system. If each system is simulated N times and an experimenter is interested in comparing the systems by estimating the average difference in system responses, say

$$\overline{Z} = \frac{1}{N}\left[\sum_{u=1}^{N}(Y_{1u} - Y_{2u})\right], \qquad [2.1.3]$$

then the variance of the average difference between system responses becomes

$$\mathrm{Var}(\overline{Z}) = \frac{1}{N^2}\sum_{u=1}^{N}\left[\mathrm{Var}(Y_{1u}) + \mathrm{Var}(Y_{2u}) - 2\,\mathrm{Cov}(Y_{1u}, Y_{2u})\right]. \qquad [2.1.4]$$

By using the same sets of streams to generate paired responses for the two systems, the response pairs will tend to be positively correlated, so that $\mathrm{Cov}(Y_{1u}, Y_{2u}) > 0$, and a reduction in the variance of the average difference between system responses can be affected.

2.1.3.3 Antithetic Random Numbers

The second variance reduction technique used in the correlation induction strategies evaluated in this research, antithetic random numbers (ARN), is a method of inducing negative correlation between sample responses. The technique involves the use of complementary sets of pseudorandom number streams on pairs of simulation runs, frequently the pairs being replicated design points. Given a stream of uniform (0,1) variates, denoted by $R = (r_1, r_2, r_3, \ldots)$, the complementary, or *antithetic*, stream is defined as

$$\overline{R} = (1 - r_1\ ,\ 1 - r_2\ ,\ 1 - r_3\ ,\ \ldots\ ).$$

For a simulation model consisting of g stochastic model components, in which common and antithetic streams are used to drive paired replications of simulation runs, the sets of pseudorandom number streams used to drive the u^{th} paired replicate would be denoted as

$$\mathbf{R}_u = (R_{1u} ,\ R_{2u} , \ldots ,\ R_{gu})$$

$$\overline{\mathbf{R}}_u = (\overline{R}_{1u} ,\ \overline{R}_{2u} , \ldots ,\ \overline{R}_{gu}) .$$

When paired replicates of simulation runs are made with stream sets $\{ \mathbf{R}_u \}$ and $\{ \overline{\mathbf{R}}_u \}$, the sample response pairs, denoted by $Y_{u(1)}$ and $Y_{u(2)}$, tend to be negatively correlated (21), (23), (33), (37), (39), (44), (56), so that $\mathrm{Cov}\,(Y_{u(1)} , Y_{u(2)}) < 0$. The ARN technique is applicable to the situation in which an experimenter is interested in estimating the mean response of a single system. If N paired replicates of a simulation run are made, and each pair is driven with common and antithetic stream sets, then the sample response pairs $\{ (Y_{1(1)} , Y_{1(2)}) , \ldots , (Y_{N(1)} , Y_{N(2)}) \}$ would be obtained. The average response of the system would be estimated as

$$\overline{Y} = \frac{1}{N} \sum_{u=1}^{N} \left[\frac{(Y_{u(1)} + Y_{u(2)})}{2} \right] . \qquad [2.1.5]$$

The variance of this average response for the two systems becomes

$$\mathrm{Var}\,(\overline{Y}) = \frac{1}{4N^2} \sum_{u=1}^{N} \left[\mathrm{Var}\,(Y_{u(1)}) + \mathrm{Var}\,(Y_{u(2)}) + 2\,\mathrm{Cov}\,(Y_{u(1)}, Y_{u(2)}) \right] . \qquad [2.1.6]$$

By using antithetic sets of streams to generate pairs of responses, paired observations will tend to be negatively correlated, and a reduction in the variance of the average response can be affected.

2.1.3.4 VRT Complications in Simulation Studies

The application of the CRN and ARN variance reduction techniques to simulation studies may lead to some complications when the techniques are implemented improperly. Originally VRTs were applied in Monte Carlo sampling experiments, where systems are modeled at one point in time. The application of VRTs in simulation, where systems are modeled over time, has raised some complications which did not arise in sampling experiments. Because stochastic simulation involves the use of pseudorandom number streams to generate responses over time, the responses in the observed time series may be serially correlated. However, Kleijnen (33) notes that when the sample response measurements of interest are the average values of the time series (not the individual responses within the series), the complications of serial correlation should not affect the experimental analysis of the response data. In this research, the sample response measurements of interest are the average values of the time series and, therefore, the issue of serial correlation is not addressed.

Another complication which has arisen in the application of VRTs to simulation studies concerns the estimation of the magnitudes of the induced correlations. Simulation models of real-world systems are usually complicated, and therefore the response variable generally cannot be expressed as a simple function of the input variables. In order to determine the amount of correlation induced when using the CRN and ARN techniques, such a function relating the input and output variables would be needed. Thus, in most simulation studies, the magnitudes of the induced correlations cannot be determined analytically. Schruben and Margolin (56) note that even though no general technique exists for inducing correlations of a specified *magnitude*, it is possible to induce correlations of a particular *sign* (positive, negative, or zero). Schruben and Margolin assume empirically reasonable signs for the induced correlations and note that these signs are consistent with a number of empirical studies, such as references (33) and (55) and the numerous references cited therein. The subject of estimating the magnitude of induced correlations is not addressed in this paper; instead the focus of this research is on the signs and relative magnitudes of the induced correlations.

2.1.4 Correlation Induction Strategies

An important characteristic of simulation studies is the experimenter's ability to selectively choose the pseudorandom number streams used to generate values for the stochastic model components. Through an appropriate choice of streams, it is possible to affect a reduction in the variability of the sample responses. Several VRTs have been developed for exploiting this opportunity to obtain improved (less variable) estimates of the true response. The induction of positive correlation through the use of CRN and the induction of negative correlation through the use of ARN are the two most easily applied VRTs used in simulation studies. Both analytical and empirical studies have substantiated, in specific contexts, the ability of these two techniques to affect a reduction in the variability of system responses (12), (21), (28), (29), (33), (37), (41), (44), (54), (55), (56), (60), (63).

In addition to the individual use of CRN and ARN in simulation studies, the simultaneous use of these two VRTs has been considered by several authors (18), (21), (33), (56), (61). A procedure proposed by Schruben and Margolin (56), termed the *assignment rule*, involves the simultaneous use of CRN and ARN in orthogonally blockable experimental designs. The procedure involves the induction of positive correlations among sample responses within the same block and the induction of negative correlations among responses in opposite blocks. Unlike most simultaneous applications of CRN and ARN, the assignment rule uses the two techniques within an experimental design rather than for replications of design points. The procedure incorporates classical experimental design procedures, as well as variance reduction techniques, into the study of simulated systems.

2.1.4.1 The Assignment Rule

The assignment rule, developed by Schruben and Margolin (56), is a VRT which involves a union of statistical design procedures and computer simulation techniques. By combining experimental design procedures with variance reduction techniques, the assignment rule strategy is aimed at improving the parameter estimates in the functional relationships developed to describe the system response as a function of the input variables. The technique requires the use of an *orthogonally blockable* experimental design (to be discussed in Section 2.2.2). The simulation runs, or design points, within each orthogonal block utilize CRN to generate values for the stochastic model components, and the design points in opposite blocks utilize ARN, thereby inducing a specific correlation structure among the simulated responses.

Schruben and Margolin (56) develop the assignment rule within the framework of an experimental design consisting of N simulation runs, yielding a set of N sample responses, denoted as

$$\underline{y} = [\, Y_1 , Y_2 , \ldots\ldots , Y_N \,]' .$$

Schruben and Margolin assume homogeneity of variance; that is, all pairs of sample responses have a common variance denoted by σ^2. Under this assumption, the variance-covariance matrix of the sample responses, apart from σ^2, reduces to a correlation matrix. The variance-covariance matrix of $\underline{y}$ becomes

$$\mathrm{Var}(\underline{y}) = \sigma^2 V \qquad [2.1.7]$$

where V is an N×N matrix of the correlations between sample response pairs. This correlation matrix can be written as

$$V = \begin{bmatrix} 1 & \rho_{12} & & \cdots & \rho_{1N} \\ \rho_{21} & 1 & 1 & & : \\ : & & & . & : \\ : & & & 1 & \rho_{ij} \\ \rho_{N1} & & \cdots & \rho_{ij} & 1 \end{bmatrix} \qquad [2.1.8]$$

where ρ_{ij} is the correlation between sample responses Y_i and Y_j $(\, i, j = 1,\ 2,\ \ldots , N\,)$.

The general structure of a simulation, as discussed by Schruben and Margolin, assumes that a relationship between the response and the levels of the input variables exists. Typically the experimenter postulates some functional relationship between the response and the factors, and due to the inability of the postulated function to determine the value of the true response, an experimental error term, denoted by ε, is included in the functional relationship. Denoting the postulated function of the input variables by $\mu(x)$, the form of the relationship can be written as

$$\underline{y}_{(x)} = \mu(x) + \underline{\varepsilon}_{(x)} \,. \qquad [2.1.9]$$

It is generally assumed that the expected value of ε is zero for each setting of the input variables. The variance of $\underline{\varepsilon}$ is then equal to the variance of $\underline{y}$ because μ is a function of the input variables only. Therefore, the variance-covariance matrix of the vector of error terms becomes

$$\mathrm{Var}\,(\underline{\varepsilon}) = \sigma^2\, V \,. \qquad [2.1.10]$$

Schruben and Margolin (56) have made the following eight assumptions concerning the signs of the ρ_{ij}'s and the components of $\underline{\varepsilon}$ in their statistical development of the assignment rule strategy:

I. A positive correlation of unknown magnitude ρ_+ is induced between the mean responses of two simulation runs when CRN is used to generate the pair of responses.

II. A negative correlation of unknown magnitude ρ_- is induced between the mean responses of two simulation runs when ARN is used to generate the pair of responses.

III. Zero correlation is induced between the mean responses of two simulation runs when IRN is used to generate the pair of responses.

IV. The positive correlation induced by using CRN on two simulation runs is a constant ρ_+ which does not depend on the specific set of pseudorandom number streams, $\mathbf{R} = (R_1, R_2, \ldots, R_g)$, or the specific pair of responses, Y_i and Y_j. Similarly the magnitude of the negative correlation induced by using ARN on two simulation runs is a constant ρ_- which does not depend on the specific pair of responses nor the specific sets of random number streams, R and $\overline{\text{R}}$. The relationship between the magnitudes of the induced correlations, which is consistent with empirical findings, is

$$0 \leq \rho_- \leq \rho_+ < 1 .$$

V. When a set of N experimental runs are performed, random block effects are produced. These effects are the result of using the same sets of randomly assigned pseudorandom number streams within each block on each simulation run (39). The model describing the dependence of experimental error on the randomly chosen pseudorandom number streams is

$$\underline{\varepsilon} = \underline{b} + \underline{\varepsilon}^*$$

where $\underline{\varepsilon}$, $\underline{b}$, and $\underline{\varepsilon}^*$ are N×1 vectors whose u^{th} elements correspond to the u^{th} simulation run, and $\underline{\varepsilon}$ is a vector of total experimental errors, $\underline{b}$ is a vector of block effects due to randomly assigning pseudorandom number streams within a block, and $\underline{\varepsilon}^*$ is a vector of the remaining unexplained portion of the experimental error.

In the assignment rule strategy, it is assumed that corresponding elements in $\underline{b}$ and $\underline{\varepsilon}^*$ are uncorrelated.

VI. The elements of $\underline{\varepsilon}^*$ are assumed to be uncorrelated.

VII. The expected value of both $\underline{b}$ and $\underline{\varepsilon}^*$ are assumed to be zero.

VIII. The i^{th} element of $\underline{b}$ depends on the i^{th} simulation run only through the set of pseudorandom number streams.

Under these assumptions, the elements of the variance-covariance matrix of the N×1 vector of block effects, $\underline{b}$, becomes

$$\operatorname{Var}(\underline{b}) = \sigma^2 \begin{bmatrix} \rho_+ & \rho_{12} & \cdots & & \rho_{1N} \\ \rho_{21} & \rho_+ & & & \vdots \\ \vdots & & & \cdot & \vdots \\ \vdots & & & \rho_+ & \rho_{ij} \\ \rho_{N1} & & \cdots & \rho_{ij} & \rho_+ \end{bmatrix}$$

where
$$\begin{aligned} \rho_{ij} &= 0 && \text{if IRN is used to generate } Y_i \text{ and } Y_j , \\ &= \rho_+ && \text{if CRN is used to generate } Y_i \text{ and } Y_j , \\ &= -\rho_- && \text{if ARN is used to generate } Y_i \text{ and } Y_j . \end{aligned}$$

The variance-covariance matrix of the N×1 vector of the unexplained experimental error, $\underline{\varepsilon}^*$, becomes

$$\mathrm{Var}(\underline{\varepsilon}^*) = \sigma^2 \begin{bmatrix} 1-\rho_+ & 0 & \cdots & & 0 \\ 0 & 1-\rho_+ & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & 1-\rho_+ & 0 \\ 0 & & \cdots & 0 & 1-\rho_+ \end{bmatrix}.$$

Therefore, the variance-covariance matrix of the total experimental error, $\underline{\varepsilon}$, as shown in equation [2.1.10] on page 23, becomes

$$\begin{aligned} \mathrm{Var}(\underline{\varepsilon}) &= \mathrm{Var}(\underline{b}) + \mathrm{Var}(\underline{\varepsilon}^*) \\ &= \sigma^2 \begin{bmatrix} 1 & \rho_{12} & \cdots & & \rho_{1N} \\ \rho_{21} & 1 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & 1 & \rho_{ij} \\ \rho_{N1} & & \cdots & \rho_{ij} & 1 \end{bmatrix} \\ &= \sigma^2 V. \end{aligned}$$

Assumptions I and II state that the magnitudes of the induced correlations are equal to ρ_+ for all pairs of positively correlated responses, and equal to ρ_- for all pairs of negatively correlated responses. Assumption IV implies that the signs and relative magnitudes of the induced correlations are known, but that the actual magnitudes are unknown. Schruben and Margolin (56) consider many possible structures, or sign patterns, for the correlation matrix V, and these authors suggest that three forms of this matrix are of practical importance in simulation studies. The three forms of V (discussed in the following section) are achieved through appropriate assignments of pseudorandom number streams to simulation runs.

2.1.4.2 *Three Correlation Induction Schemes*

Schruben and Margolin (56) suggest three methods of practical importance for inducing correlation among the response observations of a simulation study. These three procedures are the basis of the correlation induction schemes evaluated in this research and are defined as follows:

1. The use of an *independent* set of pseudorandom number streams for each simulation run (no correlation induction). This method is denoted as **IR** in this paper.

2. The use of a *common* set of pseudorandom number streams for non-replicated simulation runs (positive correlation induction). Replications of design points utilize independent random numbers because the use of common streams would yield the same observations. This method is denoted as **CR** throughout this paper.

3. The use of a *common* set of pseudorandom number streams for non-replicated design points within orthogonal blocks, and the use of an *antithetic* set of streams for non-replicated design points in opposite blocks. Replications of design points within blocks utilize independent random numbers because the use of the same streams would yield the same observations. In addition, for the designs which partition into an odd number of blocks, independent streams are used for the design points in the unpaired block. This method is termed the assignment rule and is denoted as **AR**.

The research of Schruben and Margolin examined the three aforementioned correlation induction schemes for designs which did not necessitate the replication of design points for the CR scheme, nor the replication of design points within blocks for the AR scheme. Therefore, the authors examined *pure* CR and AR correlation induction strategies. Utilization of independent random number streams for replicated design points (as will be done in this research), allows for added flexibility in the choice of an experimental design plan by combining the *pure* CR and AR strategies with the IR strategy. Schruben and Margolin examined the three correlation induction schemes in terms of two variance-related experimental design criteria, D-optimality and A-optimality, which minimize the determinant and trace of the variance-covariance matrix of the parameter estimates, respectively. Their research focused on the problem of estimating the parameters of a linear model consisting of first order and interaction terms.

Recently Hussey, Myers, and Houck (28), (29) examined the performance of the three aforementioned correlation induction schemes for the problem of estimating polynomial response surface models of orders one and two. Their research utilized four variance-related design criteria (generalized variance, prediction variance, integrated variance, and variance of slopes) to evaluate the performance of the three schemes. Similar to Schruben and Margolin, these authors considered designs in which there were no replications of design points for the CR strategy and no replications of design points within blocks for the AR strategy, thereby evaluating the *pure* CR and AR strategies.

This paper represents an extension of the research cited in the previous paragraph. Attention is restricted to experimental strategies for the estimation of response surface models of orders one and two, and the two criteria used to evaluate the three correlation induction strategies incorporate bias, as well as variance, into the performance measures. Also, designs with replicated points are evaluated through the use of independent random number streams in the CR and AR strategies (in order to generate different response values for the replicated design points).

The structure of the correlation matrix shown in equation [2.1.8] on page 22 depends on the correlation induction strategy used to generate the sample responses. Schruben and Margolin (56) concisely define V, the correlation matrix of the sample responses, for the IR, *pure* CR, and *pure* AR correlation induction strategies. In order to allow for replicated design points in the CR and AR strategies, some modifications to the authors' equations are necessary. The generalized form of the V matrices are derived in Appendix A (pages 273-277) and are shown in equations [2.1.11], [2.1.12], and [2.1.13] which follow.

The first strategy considered in this research, IR, involves the use of independent pseudorandom number streams on each simulation run. The V matrix for this scheme, denoted as V_{IR}, becomes

$$V_{IR} = I_N$$

$$= \begin{bmatrix} 1 & 0 & \dots & & 0 \\ 0 & 1_1 & & & : \\ : & & . & & : \\ : & & & 1 & 0 \\ 0 & \dots & & 0 & 1 \end{bmatrix} \qquad [2.1.11]$$

where I_N is the N×N identity matrix and N is the number of design points, or equivalently, the number of simulation runs.

The second correlation induction strategy used in this research, CR, involves the use of common streams for non-replicated design points and the use of independent streams for replicated design points. The V matrix for this scheme, denoted as V_{CR}, becomes

$$V_{CR} = I_N + \rho_+ \underline{u}\,\underline{u}' - \rho_+ U \qquad [2.1.12]$$

where U is an N×N diagonal matrix of the vector $\underline{u}$, and $\underline{u}$ is an N×1 vector whose i^{th} element is

u_i = 1 if a common *or* antithetic random number stream is used for design point *i*,
= 0 if an independent random number stream is used for design point *i*.

The correlation matrix can be written in a partitioned form if the vector of sample responses is partitioned into two parts, say

$$\underline{y}' = [\, \underline{y}'_{CRN} \mid \underline{y}'_{IRN} \,]$$

where $\underline{y}_{CRN}$ contains the responses generated using common random number streams (non-replicated design points) and $\underline{y}_{IRN}$ contains the responses generated using independent random number streams (replicated design points). The partitioned form of V_{CR} becomes

$$
V_{CR} = \left[\begin{array}{ccccc|c} 1 & \rho_+ & \rho_+ & \cdots & \rho_+ & \\ \rho_+ & 1 & & \rho_+ & \vdots & \\ \vdots & \rho_+ & 1 & & \vdots & \mathbf{0} \\ \vdots & & & \cdot & \rho_+ & \\ \rho_+ & & \cdots & \rho_+ & 1 & \\ \hline & & \mathbf{0} & & & \mathbf{I} \end{array} \right]_{(N \times N)}
$$

where the $\mathbf{0}$'s represent null matrices and $\mathbf{I}$ represents an identity matrix.

The third strategy used in this research, AR, involves the use of common streams for non-replicated design points within orthogonal blocks, the use of antithetic streams for non-replicated design points in opposite blocks, and the use independent streams for replications of design points within blocks and for all design points in any unpaired blocks. The V matrix for this scheme, denoted as V_{AR}, becomes

$$
V_{AR} = I_N + \tfrac{1}{2}(\rho_+ - \rho_-)\,\underline{u}\,\underline{u}' + \tfrac{1}{2}(\rho_+ + \rho_-)\,\underline{v}\,\underline{v}' - \rho_+ U \qquad [2.1.13]
$$

where $\underline{v}$ is an $N \times 1$ vector whose i^{th} element is

v_i = 1 if a common random number stream is used for design point i,
= 0 if an independent random number stream is used for design point i,
= -1 if an antithetic random number stream is used for design point i.

The correlation matrix can be written in a partitioned form if the vector of sample responses is partitioned into three parts, say

$$
\underline{Y}' = [\, \underline{Y}'_{CRN} \mid \underline{Y}'_{ARN} \mid \underline{Y}'_{IRN} \,]
$$

where $\underline{y}_{CRN}$ contains the responses generated using common random number streams (non-replicated design points within orthogonal blocks), $\underline{y}_{ARN}$ contains the responses generated using antithetic streams (non-replicated design points in opposite blocks), and $\underline{y}_{IRN}$ contains the remaining responses which are generated using independent streams. The partitioned form of V_{AR} becomes

$$
V_{AR} = \left[\begin{array}{ccccc|ccccc|c}
1 & \rho_+ & & \cdots & \rho_+ & & & & & & \\
\rho_+ & 1 & & & \vdots & & & & & & \\
\vdots & \cdot & 1 & & \vdots & & & -\boldsymbol{\rho}_- & & & \mathbf{0} \\
\vdots & & & \cdot & \rho_+ & & & & & & \\
\rho_+ & \rho_+ & \cdots & \rho_+ & 1 & & & & & & \\
\hline
& & & & & 1 & \rho_+ & & \cdots & \rho_+ & \\
& & -\boldsymbol{\rho}_- & & & \rho_+ & 1 & & & \vdots & \mathbf{0} \\
& & & & & \vdots & & & \cdot & \rho_+ & \\
& & & & & \rho_+ & \cdots & & \rho_+ & 1 & \\
\hline
& & \mathbf{0} & & & & & \mathbf{0} & & & \mathbf{I}
\end{array}\right]
$$

where $-\boldsymbol{\rho}_-$ is a matrix in which each element is equal to $-\rho_-$.

The next two sections present the fundamental concepts of response surface methodology (RSM). Section 2.2.1 describes the analysis phase of RSM and section 2.2.2 describes the design phase.

2.2 *Response Surface Methodology*

Response surface methodology (RSM) is a set of experimental design and optimization procedures useful in exploring the relationship between a set of input variables and one or more response variables. A response surface is a k-dimensional surface depicting the functional relationship between a response variable and a set of k input variables. RSM was initially developed by Box and Wilson (10) in the 1950's for applications in the chemical and processing industries. The breadth of applications of RSM, as it evolved during the 1960's and 1970's, expanded to many areas, including computer simulation studies. Three comprehensive literature surveys, covering both practical applications and statistical developments in RSM, are given by Hill and Hunter (26), Mead and Pike (38), and Myers (45), and thorough presentations of the statistical techniques used in RSM are given in textbooks by Box and Draper (7), Khuri and Cornell (31), and Myers (46).

RSM offers a useful approach to the prediction and optimization problems encountered in simulation studies. The design and analysis techniques used in RSM can provide an effective vehicle for constructing and parameterizing models needed for the optimization of a simulated response variable. The remainder of this Chapter describes the techniques which are commonly used in RSM.

2.2.1 The Analysis Phase of RSM

The experimental strategy of RSM revolves around the assumption that a response, denoted by η, is a function of k design variables, denoted by $\{ x_1, x_2, \ldots, x_k \}$; that is,

$$\eta = f(x_1, x_2, \ldots, x_k). \qquad [2.2.1]$$

The actual form of the function f is generally unknown, but in RSM it is assumed that the function can be approximated in small regions of the design variables by low-order polynomial models.

Prominent among the models considered are first order polynomial models and second order polynomial models, respectively written as

$$\eta = \beta_0 + \sum_{i=1}^{k} \beta_i x_i \qquad [2.2.2]$$

$$\eta = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i<j}\sum \beta_{ij} x_i x_j \qquad [2.2.3]$$

where the β coefficients are unknown model parameters which can be estimated from the data. The estimation technique generally used is ordinary least squares, however, weighted least squares is an appropriate technique when the response observations are correlated.

In addition to the use of polynomial models in RSM, there are situations in which nonlinear models are used to approximate the functional relationship between the response and the design variables. For a discussion on the use nonlinear models in the estimation of response surfaces, the reader is referred to Myers (45) and Khuri and Cornell (31). This research focuses on the estimation of response surfaces using first and second order polynomial models, and therefore nonlinear models are not addressed.

An experimenter using RSM procedures collects data from the system to estimate the β coefficients of the response surface model. The data collected are the observed responses, y, which deviate from the true response, η, due to random error. This random error, denoted by ε, is assumed to have an expected value of zero. The error term is used primarily to account for one's inability to describe the true form of the function shown in equation [2.2.1] on page 31. Therefore, the response surface model is generally written in terms of the *observed* response, rather than the *true* response, and the error term is included in the model to account for the inherent discrepancies between y and η. The first and second order polynomial models, shown in equations [2.2.2] and [2.2.3], respectively, can be written in terms of the observed response as

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon \qquad [2.2.4]$$

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i<}\sum_{j} \beta_{ij} x_i x_j + \varepsilon \, . \qquad [2.2.5]$$

The design variables $\{ x_1, x_2, \ldots, x_k \}$ are continuous, quantitative variables whose values are set by the experimenter. If N experimental runs are taken for specific combinations of the design variables, then the u^{th} design point can be written as

$$\underline{x}_u = [x_{1u}, x_{2u}, \ldots, x_{ku}]' \, .$$

The collection of design points from each of the N experimental runs constitutes the experimental design, and the design matrix, denoted by D, becomes

$$D = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{1N} & x_{2N} & \cdots & x_{kN} \end{bmatrix}$$

where each row of D corresponds to a design *point*, and each column of D corresponds to a design *variable*. For convenience, the levels of each of the k design variables are centered and scaled so that the first order design moment (the mean) of each variable is zero and the second order design moment of each variable is one; that is

$$[i] = \frac{1}{N} \sum_{u=1}^{N} x_{iu} = \bar{x}_i = 0 \qquad [2.2.6]$$

$$(i = 1, 2, \ldots, k)$$

$$[ii] = \frac{1}{N} \sum_{u=1}^{N} x_{iu}^2 = \overline{x_i^2} = 1 \, . \qquad [2.2.7]$$

The centered and scaled values of the i^{th} design variable (i^{th} column of *D*) are obtained by subtracting the mean value from the actual value and dividing by the square root of the second order design moment. Letting ξ_{iu} denote the actual value of the u^{th} level of the i^{th} input variable, then the *coded* values of the design variables become

$$x_{iu} = \frac{\xi_{iu} - \bar{\xi}_i}{S_i} \qquad [2.2.8]$$

where $\bar{\xi}_i$ is the average value of the i^{th} variable, computed as

$$\bar{\xi}_i = \frac{1}{N} \sum_{u=1}^{N} \xi_{iu}$$

and S_i is the quantity needed to scale the i^{th} variable, computed as

$$S_i = \sqrt{\frac{1}{N} \sum_{u=1}^{N} (\xi_{iu} - \bar{\xi}_i)^2} \; .$$

2.2.1.1 *Region of Interest*

In addition to the centering and scaling of individual design *variables* using the coding convention shown in equation [2.2.8], there are restrictions which can be placed on the design *points* that affect the shape of the experimental region. The centering and scaling convention is applied to the columns of the design matrix *D* for mathematical convenience, while restrictions applied to the rows of *D* affect the shape of the *k*-dimensional region in the coded design variables. The operability region, denoted by *O*, is the region in the ξ space (uncoded variables) in which experiments can actually be performed. The boundaries of *O* are usually vague and experimenters rarely explore the whole region. Typically a sequential group of experiments is used to explore small regions within *O*, each termed a region of interest, and denoted by *R*. An experimenter generally performs RSM strategies in the region, *R*, believed to contain the optimum response, and frequently exploration in the current region of interest leads to further exploration in a different region.

The shape of R is defined by the restrictions placed on the coded design points. The works of Box and Draper (5) and Draper and Lawrence (16) suggest the importance of two region shapes; spherical and cuboidal. These two general forms of R can be defined as follows:

1. R is spherical or ellipsoidal; that is, a spherical region in the coded variables can always be constructed from an ellipsoidal region in the uncoded variables, so that mathematically only a sphere needs to be considered. A spherical region is attained by requiring that each of the coded design points satisfy the inequality

$$\sum_{i=1}^{k} x_{iu}^2 \leq 1 \qquad (u = 1,\ 2,\ \ldots,N).$$

   A coded design point satisfying this requirement as an equality would fall on the perimeter of the spherical region.

2. R is cuboidal or some deformation of a k-dimensional cube; that is, a cuboidal region in the coded variables can always be attained through changes in scale, so that mathematically only a cube needs to be considered. A cuboidal region is attained by restricting the levels of the coded variables to satisfy the inequality

$$|x_{iu}| \leq 1 \qquad (i = 1,\ 2,\ \ldots,k), \qquad (u = 1,\ 2,\ \ldots,N).$$

   A coded design point satisfying this requirement as an equality for each of the k design variables would fall on an outer vertice of the cuboidal region.

If an experimenter is interested in the *corners* of the region of interest, then a cuboidal region would be an appropriate choice. On the other hand, a spherical region would be an appropriate choice for an experimenter less interested in the corners and more interested in the *axial* directions of the region. In a recent article by Draper and Guttman (15), the authors unite the spherical and cuboidal situations, and simultaneously extend the k-dimensional region shapes to an infinite range of possibilities. The range of region types that the authors develop extends from a region consisting of all axial points, to a cuboidal region consisting of all outer vertice points, with the spherical region falling on the continuum between these two extremes. This continuum of region shapes, termed *flexible* regions, is defined by the inequality

$$\sum_{i=1}^{k} |x_{iu}|^m \leq 1 \qquad (0 < m < \infty), \qquad (u = 1, 2, \ldots, N)$$

where m is a parameter which defines the shape of the experimental region; for example, the inequality defines a hypersphere if $m = 2$ and a k-dimensional cuboid if $m = \infty$. Fractional values of m define regions which are elongated in the axial directions. Therefore, a small value of m would be recommended for an experimenter interested in the axial directions, and a large value of m would be recommended for an experimenter interested in the corners of the experimental region. The reader is referred to the research of Draper and Guttman (15) for a thorough discussion of flexible design regions. In practice, the spherical and cuboidal regions are the most frequently employed region shapes, and these are the two region shapes used in the region-dependent statistical analyses of this research.

2.2.1.2 Optimization Techniques

In addition to the use of RSM techniques for modeling the relationship between a response variable and a set of input variables, RSM techniques are often used to find the levels of the design variables that yield the optimum value of the predicted response. The following three RSM procedures are frequently used in attaining this optimum:

1. Method of steepest ascent
2. Canonical analysis
3. Ridge analysis.

Each of these procedures are used at different stages of the RSM optimization process. RSM can be considered a method of hill climbing, with an objective of locating the summit of a mountain. Similar to a mountain-climbing expedition, the starting point of the optimization process is generally far below the summit. The objective of the method of steepest ascent is to move swiftly up the mountain without stopping to explore lower peaks. The method assumes that the summit region can be reached by moving along a rising path. Once the summit region is reached, canonical analysis is used to explore the region and search for the highest peak. If the peak cannot be located in the summit region because of a rising ridge or a saddle system, then ridge analysis is used to estimate the location of the highest peak.

The method of steepest ascent (or descent), initially presented by Box and Wilson (10), is an important and useful procedure for locating the region R containing the optimum. The method involves a series of experiments, each designed from the results of the preceding experiments, and each aimed at finding the path in which the value of the response increases most rapidly. First order designs are generally used in the steepest ascent method, but as the region of the optimum is approached it is sometimes necessary to use second order designs to account for curvature. The reader is referred to Box and Draper (7) and Myers (46) for complete coverage of the steepest ascent method. A modification to the method of steepest ascent, suggested by Myers and Khuri (47), provides the method with a formal stopping rule, protects against taking too many observations when the true response is decreasing, and protects against premature stopping when the true response is increasing. A different method of searching for the region of the optimum, suggested by Spendley, Hext, and Himsworth (58), is based on sequential exploration of the system using simplex search procedures. The simplex search method does not provide insight into the nature of the surface, however it is fast and efficient at locating the region of the optimum.

Canonical analysis, also initially presented by Box and Wilson (10), is a method used to locate the optimum when the current experimental region is in the general vicinity of the optimum. The method involves fitting a second order response model and rewriting the fitted model in terms of *canonical* variables to determine the nature of the local surface. The canonical form of the model is obtained by rotating the axes to form a set of orthogonal axes (removes interaction terms) and translating the origin to the stationary point of the response surface (removes first order terms). The new canonical variables are linear combinations of the original design variables, and the new origin is located at the stationary point of the fitted surface. If the stationary point is a maximum (or minimum) within the experimental region, then the experimenter can determine the levels of the design variables yielding the optimal value of the fitted response variable. If, however, the stationary point is saddle point or lies outside the experimental region, then the experimenter can use the method of ridge analysis to estimate the optimum point. The usefulness of the canonical form of the model is that it clearly reveals the nature of the response surface and gives an indication of the sensitivity of the response to changes in the values of the input variables. See references (7), (31), (46) for complete coverage of canonical analysis.

Ridge analysis, initially presented by Hoerl (27), and later incorporated into RSM by Draper (13), is a method of examining a second order response surface when canonical analysis has indicated that the optimum lies outside the experimental region. This situation occurs when the stationary point is a saddle point or when a rising, falling, or stationary ridge system is indicated. In ridge analysis, the estimated response function is subjected to additional analysis to aid the experimenter in determining the optimum point. Experimentation begins in the current experimental region and moves along the path of the rising (or falling) ridge, augmenting the canonical analysis with additional observations taken on spheres of varying radii about the ridge. Generally, the experimenter chooses the optimum point as the one with the maximum (or minimum) predicted response within the current experimental region. For thorough presentations of ridge analysis, see references (7), (31), (46).

2.2.1.3 Parameter Estimation

An experimenter utilizing RSM techniques needs a method for fitting the response function to the observed data. In particular, the β coefficients of the first and second order polynomials model need to be estimated in order to predict the response. The method of least squares is a useful estimation procedure for this research because induced correlations among responses can be taken into account. If the fitted model adequately describes the response surface and the random error terms have expected values of zero and equal variances, then the least squares estimators of the β coefficients, denoted by $\hat{\beta}$ or b, are unbiased and have the minimum variance of all linear unbiased estimators. However, if the fitted model is inadequate, then the $\hat{\beta}$ estimators may be biased. Ordinary least squares (OLS) is an appropriate estimation technique when the error terms are uncorrelated and weighted least squares (WLS) is an appropriate technique when the errors are correlated. If an OLS analysis is performed, but WLS is called for, then the $\hat{\beta}$ estimators would no longer be minimum variance estimators (17).

To obtain the least squares estimators of the β coefficients it is convenient to rewrite the first and second order polynomial models, shown in equations [2.2.4] and [2.2.5] on page 33, using matrix notation. Both polynomial models can be written in one *general linear model* form as

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon} \qquad [2.2.9]$$

where $\underline{y}$ is the vector of responses and $\underline{\varepsilon}$ is the vector of error terms. For both first and second order polynomial models, these vectors can be written as

$$\underline{y} = [y_1, y_2, \ldots, y_N]'$$

$$\underline{\varepsilon} = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N]' .$$

The β coefficients and regressor terms included in the model are different for the first and second order models. For the first order polynomial model of equation [2.2.4] on page 33, the $\underline{\beta}$ and X terms in the general linear model are

$$\underline{\beta} = [\beta_0, \beta_1, \ldots, \beta_k]'$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{1N} & x_{2N} & \cdots & x_{kN} \end{bmatrix} .$$

For the second order polynomial model of equation [2.2.5] on page 33, the $\underline{\beta}$ and X terms in the general linear model are

$$\underline{\beta} = [\ \beta_0, \beta_1, \ldots, \beta_k,\quad \beta_{11}, \beta_{22}, \ldots, \beta_{kk},\quad \beta_{12}, \beta_{13}, \ldots, \beta_{k-1,k}\ ]'$$

$$X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{k1} & x_{11}^2 & \cdots & x_{k1}^2 & x_{11}\,x_{21} & \cdots & x_{k-1,1}\,x_{k1} \\ 1 & x_{12} & \cdots & x_{k2} & x_{12}^2 & \cdots & x_{k1}^2 & x_{12}\,x_{22} & \cdots & x_{k-1,2}\,x_{k2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{1N} & \cdots & x_{kN} & x_{1N}^2 & \cdots & x_{kN}^2 & x_{1N}\,x_{2N} & \cdots & x_{k-1,N}\,x_{kN} \end{bmatrix} .$$

The ordinary least squares estimator of $\underline{\beta}$ is the vector which results in a minimum value for the sum of squares of the errors $(\underline{\varepsilon}'\,\underline{\varepsilon})$. The weighted least squares estimator of $\underline{\beta}$ is the vector which results in a minimum value for the sum of squares of the weighted errors $(\underline{\varepsilon}'\,V^{-1}\,\underline{\varepsilon})$, where V is the variance-covariance matrix of the error terms, apart from σ^2. The ordinary least squares estimator of $\underline{\beta}$, denoted as $\underline{b}_{OLS}$, and the weighted least squares estimator, denoted as $\underline{b}_{WLS}$, become

$$\underline{b}_{OLS} = (X'\,X)^{-1}\,X'\,\underline{y} \qquad [2.2.10]$$

$$\underline{b}_{WLS} = (X'\,V^{-1}\,X)^{-1}\,X'\,V^{-1}\,\underline{y}\ . \qquad [2.2.11]$$

Under the assumptions that $E[\,\underline{\varepsilon}\,] = \underline{0}$ and $Var[\,\underline{\varepsilon}\,] = \sigma^2\,V$, the variance-covariance matrices of the ordinary and weighted least squares estimators become

$$Var[\,\underline{b}_{OLS}\,] = (X'\,X)^{-1}\,X'\,V\,X\,(X'\,X)^{-1}\,\sigma^2 \qquad [2.2.12]$$

$$Var[\,\underline{b}_{WLS}\,] = (X'\,V^{-1}X)^{-1}\,\sigma^2\ . \qquad [2.2.13]$$

If $\mathrm{Var}[\underline{\varepsilon}] = \sigma^2 I_N$, implying that V is the identity matrix and the error terms are uncorrelated, then the OLS and WLS estimators are equivalent and the variance-covariance matrix of the least squares estimators can be written as

$$\mathrm{Var}[\underline{b}_{WLS}] = \mathrm{Var}[\underline{b}_{OLS}] = (X'X)^{-1}\sigma^2 . \qquad [2.2.14]$$

Therefore, the equivalence of OLS and WLS applies to the IR strategy because $V_{IR} = I_N$. Schruben and Margolin (56) have also established the equivalence of OLS and WLS for the pure CR and AR strategies. However, the OLS and WLS techniques are not equivalent when the CR and AR strategies are modified to allow for replications of design points.

The general linear model can now be rewritten in *fitted* form by omitting the error term, replacing the vector of β coefficients with the least squares estimators, and replacing the response vector with the *fitted* response vector, denoted by $\hat{\underline{y}}$. The fitted form of the general linear model becomes

$$\hat{\underline{y}} = X\underline{b} \qquad [2.2.15]$$

and the fitted forms of the first and second order polynomial models of equations [2.2.4] and [2.2.5] on page 33 can be written as

$$\hat{y} = b_0 + \sum_{i=1}^{k} b_i x_i \qquad [2.2.16]$$

$$\hat{y} = b_0 + \sum_{i=1}^{k} b_i x_i + \sum_{i=1}^{k} b_{ii} x_i^2 + \sum_{i<j}\sum b_{ij} x_i x_j . \qquad [2.2.17]$$

The least squares estimators of the β coefficients are unbiased; that is $E[\underline{b}] = \underline{\beta}$, if the fitted model is the *correct* model. However, if an experimenter fits a first order model and there is actually quadratic curvature in the response function, then the least squares estimators of one or more of the fitted model coefficients may be biased. Similarly, if a second order model is fit to the data when cubic curvature in the response function exists, then the least squares estimators of the β coefficients may be biased. The bias in the vector of estimated model coefficients is defined as

$$\text{Bias}[\underline{b}] = E[\underline{b}] - \underline{\beta} \,. \qquad [2.2.18]$$

The elements in the vector, *Bias* $[\underline{b}]$, are a function of the following items:

1. Least squares technique (ordinary or weighted).
2. Orders of the fitted and protection models (d_1 and d_2).
3. Experimental design plan (levels of the variables in the design matrix).

The protection model refers to a model of a higher degree than the fitted polynomial model, whose curvature the experimenter desires protection against. The bias in the estimated model coefficients resulting from an inadequate fitted model can be conveniently defined when the polynomial portion of the protection model is partitioned into two parts; a fitted part and an unfitted part. The partitioned form of the protection model becomes

$$\underline{y} = X_1 \underline{\beta}_1 + X_2 \underline{\beta}_2 + \underline{\varepsilon} \qquad [2.2.19]$$

where $X_1 \underline{\beta}_1$ is the fitted part of the model which includes p_1 regressor terms and is of order d_1, and $X_2 \underline{\beta}_2$ is the unfitted part which includes p_2 regressor terms and is of order $d_2 - d_1$.

The postulated polynomial model which the experimenter uses to *fit* the response data, can therefore be written as

$$\underline{y} = X_1 \underline{\beta}_1 + \underline{\varepsilon} \,. \qquad [2.2.20]$$

If the experimenter has *underfit* the response relationship, and the unfitted terms in the protection model are needed to describe the response surface curvature, then the least squares estimates of the β coefficients may be biased. The bias in the coefficient estimators resulting from the unfitted terms in X_2 can be written as

$$\text{Bias}[\underline{b}_1] = E[\underline{b}_1] - \underline{\beta}_1$$

$$= A\,\underline{\beta}_2 \qquad [2.2.21]$$

where A is the *alias*, or bias, matrix whose form depends on the least squares analysis (ordinary or weighted) used by the experimenter. The alias matrix under ordinary least squares, denoted by A_{OLS}, and the alias matrix under weighted least squares, denoted by A_{WLS}, can be written as

$$A_o = (X'_1 X_1)^{-1} X'_1 X_2 \qquad [2.2.22]$$

$$A_w = (X'_1 V^{-1} X_1)^{-1} X'_1 V^{-1} X_2 \;. \qquad [2.2.23]$$

The regressor terms included in the X_1 and X_2 matrices, and therefore the coefficients included in the $\underline{\beta}_1$ and $\underline{\beta}_2$ vectors, depend on the orders of the fitted and protection models. The fitted models used in response surface analysis are generally of order one or two, and the protection models are generally of order two or three; that is, the protection models are one degree higher than the fitted models. Denoting the order of the fitted model by d_1 and the order of the protection model by d_2, the following two fit-protection situations are examined in this research:

1. $d_1 = 1$, $d_2 = 2$.
2. $d_1 = 2$, $d_2 = 3$.

For the first situation, in which the experimenter fits a first order model ($d_1 = 1$) and desires protection against a true second order model ($d_2 = 2$), the *fit* and *protection* models, as shown in equation [2.2.16] on page 41 and equation [2.2.5] on page 33, respectively, become

$$\hat{y} = b_0 + \sum_{i=1}^{k} b_i x_i$$

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i<j}\sum \beta_{ij} x_i x_j + \varepsilon \;.$$

In the first order case, the $\underline{\beta}_1$ and $\underline{\beta}_2$ vectors of the partitioned general linear model can therefore be written as

$$\underline{\beta}_1 = [\beta_0, \beta_1, \ldots, \beta_k]'$$

$$\underline{\beta}_2 = [\beta_{11}, \beta_{22}, \ldots, \beta_{kk}, \quad \beta_{12}, \beta_{13}, \ldots, \beta_{k-1,k}]'$$

where $\underline{\beta}_1$ has $p_1 = k+1$ coefficients, and $\underline{\beta}_2$ has $p_2 = k + \binom{k}{2} = \frac{1}{2}k(k+1)$ coefficients. The corresponding X_1 and X_2 matrices can be written as

$$X_1 = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{1N} & x_{2N} & \cdots & x_{kN} \end{bmatrix} (N \times p_1).$$

$$X_2 = \begin{bmatrix} x_{11}^2 & \cdots & x_{k1}^2 & x_{11}\,x_{21} & \cdots & x_{k-1,1}\,x_{k1} \\ x_{12}^2 & \cdots & x_{k2}^2 & x_{12}\,x_{22} & \cdots & x_{k-1,2}\,x_{k2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{1N}^2 & \cdots & x_{kN}^2 & x_{1N}\,x_{2N} & \cdots & x_{k-1,N}\,x_{kN} \end{bmatrix} (N \times p_2).$$

For the $d_1 = 2$, $d_2 = 3$ fit-protection situation, the fitted second order model shown in equation [2.2.17] on page 41 and the third order protection model, respectively, become

$$\hat{y} = b_0 + \sum_{i=1}^{k} b_i x_i + \sum_{i=1}^{k} b_{ii} x_i^2 + \sum_{i<j}\sum b_{ij} x_i x_j$$

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k} \beta_{iii} x_i^3 + \sum_{i<}\sum_{j} \beta_{ij} x_i x_j + \sum_{i<}\sum_{j} \beta_{iij} x_i^2 x_j$$
$$+ \sum_{i<}\sum_{j} \beta_{ijj} x_i x_j^2 + \sum_{h<}\sum_{i<}\sum_{j} \beta_{hij} x_h x_i x_j + \varepsilon . \qquad [2.2.24]$$

In the second order case, the $\underline{\beta}_1$ and $\underline{\beta}_2$ vectors of the partitioned general linear model can therefore be written as

$$\underline{\beta}_1 = [\beta_0, \ldots, \beta_k \quad \beta_{11}, \ldots, \beta_{kk} \quad \beta_{12}, \ldots, \beta_{k-1,k}]'$$

$$\underline{\beta}_2 = [\beta_{111}, \ldots, \beta_{kkk} \quad \beta_{112}, \ldots, \beta_{k-1,k-1,k} \quad \beta_{122}, \ldots, \beta_{k-1,k,k} \quad \beta_{123}, \ldots, \beta_{k-2,k-1,k}]'$$

where $\underline{\beta}_1$ has $p_1 = 1 + 2k + \binom{k}{2} = \frac{1}{2}(k+1)(k+2)$ coefficients, and $\underline{\beta}_2$ has $p_2 = k + 2\binom{k}{2} + \binom{k}{3} = k^2 + \binom{k}{3} = \frac{1}{6}k(k+1)(k+2)$ coefficients. The X_1 matrix, which contains the fitted first and second order model terms, becomes

$$X_1 = \begin{bmatrix} 1 & x_{11} & \ldots & x_{k1} & x_{11}^2 & \ldots & x_{k1}^2 & x_{11}x_{21} & \ldots & x_{k-1,1}x_{k1} \\ 1 & x_{12} & \ldots & x_{k2} & x_{12}^2 & \ldots & x_{k1}^2 & x_{12}x_{22} & \ldots & x_{k-1,2}x_{k2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{1N} & \ldots & x_{kN} & x_{1N}^2 & \ldots & x_{kN}^2 & x_{1N}x_{2N} & \ldots & x_{k-1,N}x_{kN} \end{bmatrix} (N \times p_1)$$

and the X_2 matrix, which contains the unfitted third order terms in the protection model, can be written as

$$X_2 = \begin{bmatrix} x_{11}^3 \cdots x_{k1}^3 & x_{11}^2 x_{21} \cdots x_{k-1,1}^2 x_{k1} & x_{11} x_{21}^2 \cdots x_{k-1,1} x_{k1}^2 & x_{11} x_{21} x_{31} \cdots x_{k-2,1} x_{k-1,1} x_{k1} \\ x_{12}^3 \cdots x_{k2}^3 & x_{12}^2 x_{22} \cdots x_{k-1,2}^2 x_{k2} & x_{12} x_{22}^2 \cdots x_{k-1,2} x_{k2}^2 & x_{12} x_{22} x_{32} \cdots x_{k-2,2} x_{k-1,2} x_{k2} \\ \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \\ x_{1N}^3 \cdots x_{kN}^3 & x_{1N}^2 x_{2N} \cdots x_{k-1,N}^2 x_{kN} & x_{1N} x_{2N}^2 \cdots x_{k-1,N} x_{kN}^2 & x_{1N} x_{2N} x_{3N} \cdots x_{k-2,N} x_{k-1,N} x_{kN} \end{bmatrix}$$

$$(N \times p_2) \ .$$

In this research, two least squares analyses techniques (ordinary and weighted) and two fit-protection situations ($d_1 = 1, d_2 = 2$ and $d_1 = 2, d_2 = 3$) are used in the evaluation of the correlation induction strategies. The next section of this paper presents the experimental design plans that are used to evaluate the correlation induction strategies for response surface models of orders one and two.

2.2.2 The Design Phase of RSM

The experimental investigation of a response surface involves both an analysis phase and a design phase, which the experimenter must treat together in the total analysis. The experimental strategies used in the analysis phase were discussed in section 2.2.1. The design phase of RSM, which deals with the choice of experimental plans for fitting first and second order polynomial models, are discussed in this section. First and second order response surface designs are generally treated separately because linear response surfaces can be estimated with fewer model parameters than quadratic response surfaces. Designs for fitting first order models utilize $N \geq k + 1$ design points, where k is the number of regressor variables. There must be at least two levels of each factor in order to estimate the linear coefficients. Second order designs utilize $N \geq \frac{1}{2}(k + 1)(k + 2)$ design points and must contain at least three levels of each factor in order to estimate the quadratic coefficients.

The experimental designs used to fit second order models could be used to fit first order models, but these designs would be unnecessarily large and consist of more factor levels than needed to fit a first order model. For this reason, first order experimental designs are generally less complicated than designs used to fit second order models.

The remainder of this section is divided into three parts; section 2.2.2.1 is devoted to first order designs, section 2.2.2.2 discusses the important properties of second order designs, and section 2.2.2.3 presents the specific classes of second order designs used to evaluate the correlation induction strategies examined here.

2.2.2.1 First Order Designs

First order experimental designs are those designs which are useful for estimating the linear coefficients of first order response surface models. These designs require that each factor be present at two or more levels and that there be at least one design point for each of the estimated linear coefficients. When these designs possess the property of *orthogonality*, the coefficients of first order models can be estimated with maximum precision. Orthogonality was first recognized as an important design property in the research of Fisher (19) and Yates (64). Later, Box (3) motivated the use of orthogonal designs in RSM. Orthogonal designs are a class of designs in which the *N*-dimensional regressor variable vectors are at right angles to each other, or equivalently, the columns of the X matrix are orthogonal to each other. A first order design is considered to be an orthogonal design when the regressor variables are orthogonal to each other. For design variables which have been centered and scaled using the coding convention shown in equation [2.2.8] on page 34, the requirement for first order orthogonality becomes

$$\sum_{u=1}^{N} x_{iu} x_{ju} = 0 \qquad (i \neq j = 1, \ldots, k)$$

where x_{iu} and x_{ju} refer to the u^{th} levels of the i^{th} and j^{th} coded regressor variables. When this condition holds, the X'X matrix for a first order design is a diagonal matrix, and the variances of the least squares estimators of the β coefficients are minimized.

The orthogonal designs most frequently used in fitting first order models are the two-level factorial and fractional factorial designs, discussed at length by Box and Hunter (9). Full two-level factorial designs require 2^k design points and fractional designs require 2^{k-p} design points, where p is a positive integer less than k and 2^{-p} is the size of the fractional design relative to the full factorial design. Simplex designs (3) and Plackett-Burman designs (52) are two other types of orthogonal designs useful in fitting first order models, both of which are economical in terms of the number of runs required. These designs utilize two levels of each factor but require the use of only $N = k + 1$ design points. Designs such as these which utilize the minimum number of design points needed to estimate the model coefficients are termed *saturated* designs. Another type of first order design is the non-orthogonal, one-factor-at-a-time Koshal design (35). However, in this research, only orthogonal designs are considered in the first order case because the property is so easily obtained and because the assignment rule correlation induction strategy requires the use of an experimental design which blocks in an orthogonal manner.

Orthogonal first order designs possess many desirable properties. Those of particular importance in this research include:

1. Orthogonal first order designs minimize the variances of the least squares estimates of the β coefficients as compared to all other designs with the same number of design points.

2. The designs have the property of *rotatability*; that is, the variance of the predicted response is constant on spheres about the center of the design.

3. Without loss of orthogonality or rotatability, the designs can be augmented with multiple center runs to afford a convenient check for quadratic curvature.

4. The two-level factorial designs can be *blocked orthogonally*; that is, these designs can be partitioned into blocks in such a manner that the block effects do not affect the usual estimates of the linear coefficients.

The assignment rule correlation induction strategy, discussed on pages 21-25, necessitates the use of an orthogonally blockable experimental design. Therefore, in this research, the two-level factorial and fractional factorial designs are used to evaluate the correlation induction strategies for the case of fitting a first order polynomial model. The factorial designs can be blocked orthogonally by making the vectors of block effects orthogonal to the regressor variable vectors. This is done by confounding block effects with interaction terms and by making the levels of each regressor variable sum to zero within each block. For example, a full 2^3 factorial design can be partitioned into two orthogonal blocks by confounding the $x_1x_2x_3$ interaction term with the block effects. The resulting centered and scaled design matrix, augmented with two center runs, becomes

$$
\begin{array}{cc}
 & \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\
D = & \left[\begin{array}{ccc}
-1 & -1 & -1 \\
1 & 1 & -1 \\
1 & -1 & 1 \\
-1 & 1 & 1 \\
\hdashline
1 & -1 & -1 \\
-1 & 1 & -1 \\
-1 & -1 & 1 \\
1 & 1 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{array} \right] \times g
\end{array}
$$

where g is a scaling factor, frequently used to scale the design to the $[ii] = 1$ coding convention shown in equation [2.2.7] on page 33. If experimental runs were costly, an experimenter might choose to perform the runs in only the first block, thus utilizing a ½ fraction, or equivalently, a 2^{3-1} fractional factorial design. This design would be saturated, but unlike the full 2^3 design, it could not be partitioned into orthogonal blocks. Fractional factorial designs are typically used in response surface modeling to reduce the number of experimental runs required when $k \geq 5$ regressor variables are involved, and fractions of these larger designs generally lend themselves to orthogonal blocking. Center points can easily be accommodated in any of the orthogonally blocked two-level factorial designs, and the coefficients of the first order model remain orthogonal even if the number of center runs in each block are not the same.

2.2.2.2 Second Order Design Properties

First order designs are predominantly used in RSM for locating the region in the factor space which contains the optimum response. Once an experimenter has tentatively found this region, a second order design is generally used to fit a second order polynomial model. In addition to the linear terms in a first order model, second order models contain quadratic and two-way interaction terms. These additional terms allow for accommodation of curvature in the fitted response surface, and therefore a second order model can more accurately approximate the true response surface and yield a better estimate of the optimum. However, an experimenter must pay for the added precision gained with a second order design. When compared to first order designs, the number of experimental runs and the number of parameters estimated are both significantly larger for second order designs.

Second order designs are experimental plans useful in the estimation of the coefficients of second order polynomial models. The designs require at least three levels of each factor and at least one design point for each of the estimated $\frac{1}{2}(k+1)(k+2)$ coefficients, of which $k+1$ are linear, k are quadratic, and $\binom{k}{2}$ are two-way interaction coefficients. As in the first order case, orthogonality is an important property for second order designs, but rotatability, uniform precision, and orthogonal blocking are often more important properties in the second order case. A second order design which is orthogonal in its initial orientation in the design region often loses its orthogonality when rotated. For this reason, orthogonality is less important, and second order designs are frequently chosen to achieve some other design property, such as stable variances of the coefficients as the design is rotated, or uniform variance of the predicted response within the region of interest.

Orthogonality

In the case of a first order design, the property of orthogonality is achieved by requiring that the columns of the X matrix be orthogonal to each other, thus yielding a diagonal X'X matrix. In the case of a second order design, it is impossible to achieve orthogonality in the first order sense because the model terms

$$x_{iu},\ x_{iu}x_{ju},\ x_{iu}^2,\ x_{iu}^2x_{ju},\ x_{iu}x_{ju}x_{ku},\ x_{iu}^3,\ x_{iu}^2x_{ju}^2,\ x_{iu}^2x_{ju}x_{ku},\ \text{and}\ x_{iu}x_{ju}x_{ku}x_{lu}$$

are not all functionally independent, and a diagonal X'X matrix cannot be obtained unless all of the x_{iu} are equal to zero. However, Box and Hunter (8) note that an infinite variety of *sort-of* orthogonal designs can be produced by redefining the independent variables in terms of orthogonal polynomials. The resulting designs are orthogonal in the sense that no two first order coefficients are correlated; that is,

$$\mathrm{Cov}[b_0, b_i] = \mathrm{Cov}[b_i, b_j] = 0$$

and no two second order order coefficients are correlated; that is,

$$\mathrm{Cov}[b_{ii}, b_{jj}] = \mathrm{Cov}[b_{ii}, b_{ij}] = \mathrm{Cov}[b_{ii}, b_{jk}] = \mathrm{Cov}[b_{ij}, b_{jk}] = \mathrm{Cov}[b_{ij}, b_{kl}] = 0\,.$$

However, the b_0 estimator is inevitably correlated with the quadratic coefficients; that is, $\mathrm{Cov}[b_0, b_{ii}] \neq 0$. All other first and second order coefficients are uncorrelated in these second order orthogonal designs.

In order to achieve a diagonal X'X matrix, the quadratic model terms can be corrected for their means by replacing the $b_{ii}\,x_{iu}^2$ terms of the second order model with $b_{ii}\,(x_{iu}^2 - \overline{x_{iu}^2})$. However, when this type of second order orthogonal design is rotated, a different variance structure for all but the linear coefficients results; that is, the variance of b_i remains constant, but the variances of b_{ii} and b_{ij} and the covariances between b_{ii}, b_{jj}, and b_{ij} change as the design is rotated. Therefore, the second order design property of orthogonality refers to orthogonality in a specific orientation only, and the property is generally lost upon rotation of the design. For this reason, the property of rotatability is often more desirable than orthogonality in the second order case.

Rotatability

Rotatable designs are those designs in which the variances and covariances of the coefficients are unaffected by rotations of the design relative to the response surface. A design has the property of rotatability when the variance of the predicted response is constant on spheres about the center of the design. The use of a rotatable design assures a spherically uniform distribution of information on the response variable. The conditions necessary for rotatability involve the *moments* of the design, which are denoted as

$$[\,1^{\delta_1}\ 2^{\delta_2}\ \ldots\ldots\ k^{\delta_k}\,] = \frac{1}{N}\sum_{u=1}^{N} x_{1u}^{\delta_1}\ x_{2u}^{\delta_2}\ \ldots\ldots\ x_{ku}^{\delta_k} \qquad [2.2.25]$$

where $\delta = \sum_{i=1}^{k} \delta_i$ is the *order* of the design moment,

$\lambda_2 = [\mathrm{ii}] = \frac{1}{N}\sum_{u=1}^{N} x_{iu}^2$ is the *pure* second order design moment, and

$\lambda_4 = [\mathrm{iijj}] = \frac{1}{N}\sum_{u=1}^{N} x_{iu}^2\ x_{ju}^2$ is the *mixed* fourth order design moment.

Box and Hunter (8) derived the conditions for rotatability and found that for a second order design to be rotatable, the following three conditions must hold:

1. All of the design moments through order four with *any* δ_i equal to an odd number must be equal to zero,

$$[\mathrm{i}] = [\mathrm{ij}] = [\mathrm{iii}] = [\mathrm{iij}] = [\mathrm{ijk}] = [\mathrm{iiij}] = [\mathrm{iijk}] = [\mathrm{ijkl}] = 0\,.$$

2. *Pure* second order design moments must all be equal,

$$\lambda_2 = [ii] = [jj] .$$

3. *Pure* fourth order design moments must be equal to three times the *mixed* fourth order moments,

$$3\,\lambda_4 = 3\,[iijj] = [iiii] .$$

Some second order rotatable designs have the capability of being orthogonal designs also. However, the variance of the predicted response is often large at the perimeter of second order orthogonal designs, and therefore the design property of uniform precision is often of greater importance than orthogonality in the second order case.

Uniform Precision

A uniform precision second order design is one in which the variance, or precision, of the predicted response is the same at the perimeter of the design region as it is at the center of the design region. It is assumed that the design region is centered and scaled such that the points on the perimeter of the design fall on a radius of $\sqrt{\sum_{i=1}^{k} x_{iu}} = 1$ from the center of the design. Uniform precision designs can be constructed from rotatable second order designs by making $\lambda_2 = 1$ and requiring that λ_4 be equal to a specific value which is a function of the number of regressor variables. Box and Hunter (8) specify the following values of λ_4 needed for a second order *rotatable* design to have the property of uniform precision:

k	2	3	4	5	6	7	8
$\lambda_4 = [iijj]$	.7844	.8385	.8704	.8918	.9070	.9184	.9274

A rotatable second order design utilizing the uniform precision value of the mixed fourth moment achieves roughly uniform precision in the predicted responses over a centered and scaled *spherical* region of radius one. Outside this region, however, the variances of the predicted responses generally increase rapidly.

Orthogonal Blocking

In addition to the properties of orthogonality, rotatability, and uniform precision, designs which have the capability of blocking in an orthogonal manner are frequently useful to an experimenter. When all of the experimental runs cannot be made under homogeneous conditions, blocking becomes an essential part of the experimental design procedure, and partitioning the experimental design into orthogonal blocks enables the coefficients to be estimated with greater accuracy than would otherwise be possible. Second order designs that admit to orthogonal blocking are those designs in which the block effects do not alter the usual estimates of the coefficients of a second order polynomial model. Box and Hunter (8) show that two conditions must be fulfilled to achieve this blocking property. For a second order design in which there are N_b design points in the b^{th} block, the conditions for orthogonal blocking are as follows:

1. Each block itself must be a first order orthogonal design,

$$\sum_{u}^{N_b} x_{iu} x_{ju} = 0 \ . \qquad [2.2.26]$$

2. The sum of squares for each design variable within each block must be proportional to the block size,

$$\frac{\sum_{u}^{N_b} x_{iu}^2}{\sum_{u=1}^{N} x_{iu}^2} = \frac{N_b}{N} \ . \qquad [2.2.27]$$

In the next section, examples of orthogonal blocking for each of the design classes used in this research are given. Because the assignment rule correlation induction strategy (discussed on pages 21-25) requires the use of an orthogonally blockable experimental design, only second order designs classes which lend themselves to orthogonal blocking are considered here.

2.2.2.3 Second Order Design Classes

Myers (45) notes that the central composite design is the most flexible and most popular family of second order designs used in response surface modeling. However, many alternative second order designs exist as well, such as the Box-Behnken designs, three-level factorial designs, and various small composite designs. Each of these four design classes are discussed in this section and each is used to evaluate the correlation induction strategies examined in this research.

Central Composite Design

The central composite design (CCD), introduced by Box and Wilson (10), is a five-level design consisting of a two-level factorial design, or fraction, augmented with $2k$ axial points and a chosen number of center points. With appropriate choices for the levels of the design variables, the CCDs can achieve one or more of the second order design properties of orthogonality, rotatability, uniform precision, and orthogonal blocking. For example, a $k=3$ CCD with one center run, partitioned into two orthogonal blocks, would have the design matrix

$$
D = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \left[\begin{array}{ccc}
-1 & -1 & -1 \\
-1 & -1 & 1 \\
-1 & 1 & -1 \\
-1 & 1 & 1 \\
1 & -1 & -1 \\
1 & -1 & 1 \\
1 & 1 & -1 \\
1 & 1 & 1 \\
\hdashline
\alpha & 0 & 0 \\
-\alpha & 0 & 0 \\
0 & \alpha & 0 \\
0 & -\alpha & 0 \\
0 & 0 & \alpha \\
0 & 0 & -\alpha \\
0 & 0 & 0
\end{array} \right] \end{array} \times g
$$

where g is a scaling factor and α is a parameter representing the distance of the axial points from the origin. This design is fairly economical, utilizing 15 design points to estimate the 10 coefficients of the second order model. The axial portion of the design is a three-level, one-factor-at-a-time augmentation to the two-level factorial portion of the design and it allows for the estimation of the β_{ii} quadratic model coefficients. Through appropriate choices for the values of α and the number of center runs, the above blocking arrangement can be an orthogonal one.

The CCDs are augmented with a chosen number of center run(s) to avoid singularity of the $X'X$ matrix, which occurs when all of the design points lie on a common sphere about the origin. When replicated, the center runs provide an internal estimate of the pure experimental error variance. Much of the flexibility of the CCDs is derived from the experimenter's ability to choose both the number of center runs and the value of α needed to achieve the desired design properties.

In this research, the notation used to denote number of design points in the CCDs is as follows:

Number of factorial points: $F = 2^k$ or 2^{k-p}

Number of axial points: $n_a = 2k$

Number of center points: N_c

Total number of design points: $N = F + n_a + N_c$.

The *odd* design moments of the CCDs are all equal to zero, and the even moments through order four can be written as

$$[ii] = \lambda_2 = \frac{F g^2 + 2 g^2 \alpha^2}{N}$$

$$[iijj] = \lambda_4 = \frac{F g^4}{N}$$

$$[iiii] = \frac{F g^4 + 2 g^4 \alpha^4}{N}$$

where g is a scaling factor, typically used in scaling the design to the $[ii] = 1$ coding convention. The value of g needed to scale the CCDs to $[ii] = 1$ becomes

$$g = \sqrt{\frac{N}{F + 2\alpha^2}} .$$

A CCD can be a second order orthogonal design, for a given number of center runs, by choosing the value of α such that $\lambda_2 = \lambda_4 = 1$. Utilizing the above value of g satisfies the $\lambda_2 = 1$ orthogonality requirement, and setting $[iijj] = 1$ satisfies the $\lambda_4 = 1$ orthogonality requirement. Therefore, an orthogonal CCD can be constructed by choosing α as the value which satisfies the equation

$$\alpha = \sqrt{\frac{1}{2} (\sqrt{F N} - F)} .$$

A CCD is also capable of being a rotatable design through an appropriate choice for the value of α, irregardless of the number of center points. A rotatable CCD must have $[iiii] = 3\,[iijj]$, and therefore a rotatable CCD can be constructed by selecting α as the value which satisfies the equation

$$\alpha = \sqrt[4]{F} .$$

A rotatable CCD can also be an orthogonal design by choosing the appropriate number of center runs needed to make $\lambda_2 = \lambda_4 = 1$. Instead of achieving orthogonality, a rotatable CCD can be a uniform precision design by requiring that $\lambda_2 = 1$ and that λ_4 be equal to the appropriate uniform precision value (see page 53). However, the number of center runs (N_c) required to satisfy either the orthogonality or uniform precision conditions is typically not integer-valued, and therefore the actual design plans are only *near*-orthogonal or *near*-uniform precision CCDs.

The CCD also lends itself well to orthogonal blocking arrangements because the factorial and axial portions are each first order orthogonal designs. By using the value of α which makes the sums of squares within blocks proportional to the block sizes, the axial and factorial portions of a CCD are each an orthogonal block. The assignment rule strategy (discussed in section 2.1.4.1) necessitates the use of an experimental design, such as the CCD, which can be partitioned into orthogonal blocks. In this research, the full CCDs with $k = 2, 3, 4, 5$ factors and the ½ fractional CCDs with $k = 6, 7$ factors are used to evaluate the correlation induction strategies.

Box-Behnken Designs

The Box-Behnken designs (BBDs), introduced in a paper by Box and Behnken (4), are a class of three-level, second order designs formed by combining two-level factorial designs with incomplete block designs and a chosen number of center runs in a particular manner. The BBDs are not as flexible as the CCDs but require the use of only three levels of each variable, compared to the five levels required for a CCD. Like the CCDs, the BBDs are augmented with a chosen number of center runs to avoid singularity of the $X'X$ matrix. For example, the $k = 5$ BBD with six center runs, partitioned into two orthogonal blocks, has the design matrix

$$
D = \begin{array}{c}
\begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\
\left[\begin{array}{ccccc}
\pm 1 & \pm 1 & 0 & 0 & 0 \\
0 & 0 & \pm 1 & \pm 1 & 0 \\
0 & \pm 1 & 0 & 0 & \pm 1 \\
\pm 1 & 0 & \pm 1 & 0 & 0 \\
0 & 0 & 0 & \pm 1 & \pm 1 \\
0 & 0 & 0 & 0 & 0 \\
\hdashline
0 & \pm 1 & \pm 1 & 0 & 0 \\
\pm 1 & 0 & 0 & \pm 1 & 0 \\
0 & 0 & \pm 1 & 0 & \pm 1 \\
\pm 1 & 0 & 0 & 0 & \pm 1 \\
0 & \pm 1 & 0 & \pm 1 & 0 \\
0 & 0 & 0 & 0 & 0
\end{array}\right]
\end{array} \times g
\qquad
\begin{array}{c}
\text{Number} \\ \text{of Points} \\
4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 3 \\ \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 3
\end{array}
$$

$N = 46$

where the ± 1 symbol indicates that all combinations of the $+1$ and -1 levels are to be run. This design is not particularly economical, utilizing 46 design points to estimate the 21 coefficients of the second order model. The exact number of center runs used is not critical, but for the $k=5$ design Box and Behnken recommend three center points in each block in order to achieve a fairly uniform variance profile over the experimental region.

The Box-Behnken designs are second order orthogonal and near-rotatable, with the $k=4, 7$ designs being exactly rotatable. Through an appropriate choice for the number of center runs, the BBDs can be near-uniform precision designs also. No BBDs exist for $k=2, 8$ factors and the $k=4$ design is simply an orthogonal rotation of the four variable CCD. With the exception of the $k=3$ design, orthogonal blocking is possible in all of the BBDs, but an equal number of center runs must be added to each block to ensure that the sum of squares within blocks remain proportional to the block sizes. The $k=4$ BBD partitions into three orthogonal blocks and the $k=5, 7$ designs partition into two blocks. The BBDs can also be scaled to the $[ii]=1$ coding convention using a scaling factor of $g=\sqrt{N/4(k-1)}$ for the $k=4, 5, 7$ designs.

The assignment rule correlation induction strategy (discussed on pages 21-25) requires the use of an experimental design which can be partitioned into orthogonal blocks. Therefore, the BBD with three factors cannot be used in this research. Because the $k=4$ design partitions into an odd number of blocks, a *modified* assignment rule strategy, which involves the use of common and antithetic random number streams in two blocks and the use of independent streams in the third block, is used for this design. In this research, the BBDs with $k=4, 5, 7$ factors are used to evaluate the correlation induction strategies.

Three-Level Factorial Designs

A natural extension to the two-level factorial designs are the class of three-level factorial designs, utilizing k factors at each of three levels. The 3^k factorial designs can be used to estimate the coefficients of a second order model since each factor is present at three levels. Full 3^k designs require a large number of experimental runs even for moderate values of k, so the concept of fractional replication is particularly attractive for these designs. Like the 2^k factorial designs, the 3^k designs fall into the class of orthogonal designs, however fractional replications are not second order orthogonal. Unlike the CCDs, the 3^k designs cannot be rotatable or uniform precision designs, but similar the CCDs, the 3^k designs can be blocked orthogonally. For example, a 3^3 factorial design can be partitioned into three orthogonal blocks by confounding the $x_1x_2x_3$ interaction term with the block effects. Utilizing two of the three orthogonal blocks as the experimental design, the design matrix for the $\frac{2}{3}$ fraction of the 3^3 factorial design becomes

$$D = \begin{array}{c} \\ \end{array}\begin{array}{ccc} x_1 & x_2 & x_3 \end{array}$$

$$D = \left[\begin{array}{ccc}
-1 & -1 & -1 \\
-1 & 0 & 1 \\
-1 & 1 & 0 \\
0 & -1 & 1 \\
0 & 0 & 0 \\
0 & 1 & -1 \\
1 & -1 & 0 \\
1 & 0 & -1 \\
1 & 1 & 1 \\
\hdashline
-1 & -1 & 0 \\
-1 & 0 & -1 \\
-1 & 1 & 1 \\
0 & -1 & -1 \\
0 & 0 & 1 \\
0 & 1 & 0 \\
1 & -1 & 1 \\
1 & 0 & 0 \\
1 & 1 & -1
\end{array}\right] \times g \, .$$

This design is fairly economical, utilizing 18 design points to estimate the 10 coefficients of the second order model. If the full factorial design were used, the design would be second order orthogonal but an additional 9 design points would be required.

Center points can be accommodated in the orthogonally blocked three-level factorial designs, but an equal number of center runs must be added to each block to ensure that the sum of squares within blocks remain proportional to the block sizes. The full three-level factorial designs always contain one center run, representing the intermediate levels of each factor and, therefore, the addition of center runs represent replications of this design point. The 3^k designs can be scaled to the coding convention of $[ii] = 1$ through the use of a scaling factor of $g = \sqrt{2/3\,(N/F)}$.

The 3^k designs are one of the second order design classes used to evaluate the correlation induction strategies because of their ability to block orthogonally. However, the 3^k designs partition into three orthogonal blocks, and therefore a modified AR strategy (as used for the $k = 4$ BBD) is required for the 3^k designs. Also, because of the excessively large size of the full 3^k designs, $1/3$ fractional replications of the $k = 6, 7$ designs are considered in this research.

Small Composite Designs

Hartley (24), Westlake (62), and Draper (14) have each developed economical designs based on the CCDs, termed small composite designs. These designs use a reduced number of design points in the factorial portions of the CCDs to produce designs which are either saturated or near-saturated. Hartley's designs (24) are essentially "small" CCDs, consisting of fractional 2^k factorial designs rather than full 2^k designs for the factorial portion of the CCD. For the case of $k=3$ factors, the small composite design suggested by Hartley utilizes the $x_1x_2x_3$ term as the defining contrast, and the resulting ½ fraction of the 2^3 design is augmented with the axial portion of a CCD and one center run, yielding the design matrix

$$D = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \left[\begin{array}{ccc} 1 & 1 & 1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ 0 & 0 & 0 \\ \hdashline \alpha & 0 & 0 \\ -\alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & -\alpha & 0 \\ 0 & 0 & \alpha \\ 0 & 0 & -\alpha \end{array}\right] \end{array} \times g \; .$$

This design utilizes 11 design points to estimate the 10 coefficients of the second order model. A center run is included to avoid singularity of the X'X matrix and the values of α and N_c can be chosen by the experimenter to achieve orthogonal blocking, rotatability, or uniform precision. Similar to a CCD, the design can be scaled to the $[ii]=1$ coding convention through the use of a scaling factor of $g=\sqrt{N/(F+2\alpha^2)}$.

Hartley recommends the use of ½ fractions for the $k=2, 3, 4, 5$ designs, ¼ fractions for the $k=6, 7$ designs, and ⅛ fractions for the $k=8, 9$ designs. The 2^k fractions suggested by Hartley insure that two-way interaction terms are not aliased with each other, but allow for aliasing of two-way interaction terms and linear terms because the axial portions of the designs provide additional information on the linear terms. With an appropriate choice for the value of α, Hartley's small composite designs partition into two orthogonal blocks; the factorial portion serves as one block and the axial portion of the design serves as the other block. Each of Hartley's designs are either saturated or near-saturated, with the exception of the $k=5, 7, 9$ designs.

Westlake (62) provided near-saturated $k=5, 7, 9$ small composite designs using *irregular* fractions of the 2^k designs for the factorial portions. However, Westlake's designs do not block orthogonally because of the irregular fractions, and therefore these designs are not considered in this research.

Similar to Westlake, Draper (14) developed $k=5, 7, 9$ small composite designs which are closer to saturation than the designs developed by Hartley. Draper's designs utilize incomplete Plackett-Burman designs (52) augmented with the axial portion of a CCD. Because the Plackett-Burman designs are first order orthogonal, Draper's small composite designs can be blocked in an orthogonal manner. For example, the $k=5$ small composite design suggested by Draper, partitioned into two blocks, has the design matrix

$$
D = \begin{array}{c}
\begin{array}{ccccc} \;x_1 & \;x_2 & \;x_3 & \;x_4 & \;x_5 \end{array} \\
\left[\begin{array}{ccccc}
-1 & -1 & -1 & -1 & -1 \\
1 & 1 & 1 & 1 & 1 \\
-1 & -1 & 1 & 1 & 1 \\
-1 & 1 & 1 & -1 & 1 \\
1 & -1 & -1 & 1 & 1 \\
1 & 1 & -1 & 1 & -1 \\
1 & 1 & 1 & -1 & -1 \\
1 & -1 & 1 & -1 & -1 \\
1 & -1 & -1 & -1 & 1 \\
-1 & 1 & -1 & 1 & -1 \\
-1 & 1 & -1 & -1 & 1 \\
-1 & -1 & 1 & 1 & -1 \\
\hline
\pm\alpha & 0 & 0 & 0 & 0 \\
0 & \pm\alpha & 0 & 0 & 0 \\
0 & 0 & \pm\alpha & 0 & 0 \\
0 & 0 & 0 & \pm\alpha & 0 \\
0 & 0 & 0 & 0 & \pm\alpha
\end{array}\right]
\end{array} \times g
$$

where the $\pm\alpha$ notation indicates that one run is made at the $+\alpha$ level and another at the $-\alpha$ level. This design is near-saturated, utilizing 22 design points to estimate the 21 coefficients of the second order model and the design can easily accommodate center runs. The non-axial portion of the design consists of five columns from a twelve-run Plackett-Burman design and can be used to fit a first order model, but the alias structure is not as *clean* as that of a fractional factorial design. Draper also developed designs utilizing 42 runs for the case of $k = 7$ factors (8 more runs than a saturated design) and 62 runs for the case of $k = 9$ factors (7 more runs than a saturated design). Similar to the CCDs, Draper's designs can be scaled to the $[ii] = 1$ coding convention by using a scaling factor of $g = \sqrt{N/(F + 2\alpha^2)}$.

In summary, the second order designs used to evaluate the correlation induction strategies examined this research (each of which admits to orthogonal blocking) are the full and ½ fractional central composite designs, the Box-Behnken designs, the full and ⅓ fractional three-level factorial designs, and the small composite designs developed by Hartley and Draper.

2.2.3 RSM Design Criteria

There are numerous statistical design criteria available for comparing the three correlation induction strategies identified earlier. Optimality design criteria, often referred to as alphabetic optimality criteria (7), are based on making the variance-covariance matrix of the parameter estimates *small* in some sense. Optimal design theory is an important component in the development of designs for the case of regression models, but there has been much controversy concerning the applicability of optimal design theory to response surface analysis (59).

Many of the commonly-used optimality design criteria, motivated in particular by Kiefer (32), assume that the proposed model, stating the effect of the input variables on the response variable, is exactly correct. Box and Draper (5) were the first authors to consider in depth the effect of model misspecification on experimental design. These authors suggest that a mean squared error criteria is more appropriate than a variance-based optimality criteria because it takes bias error due to model misspecification, as well as variance error, into account. Mean squared error criteria are those which minimize the sum of the variance and squared bias errors.

Consideration of bias error is particularly important in response surface analysis because the fitted, low-order polynomial models are only meant to approximate the true response surface in small regions of the factor space. For cases in which a first order model is used, the question arises as to whether the experimental design can offer proper protection against the possibility of quadratic curvature in the response surface. In the second order case, the question arises as to whether the design affords protection against the existence of cubic curvature in the response surface. The mean squared error design criteria consider variance and bias simultaneously, thereby allowing for protection against large variances and affording protection against biases due to inadequacy of the fitted model. Frequently variance and bias are competing effects; that is, when the variance error is small, the bias error is large, and vice versa. Therefore, a criteria which seeks to minimize the average mean squared error of the fitted response equation is particularly useful when the experimenter suspects that both variance and bias error exist. In some situations, the mean squared error criteria can be used to find the optimal placement of design points for the current region of interest.

The design criteria of minimum average mean squared error of the response variable (termed the *MSE of response* criteria) was originally developed by Box and Draper (5). These authors focused on the problem of fitting a first order polynomial model ($d_1 = 1$) while protecting against bias in the fitted response if, in fact, the true response could be more adequately modeled in the region of interest using a second order polynomial model ($d_2 = 2$). In other words, the authors looked at the situation in which a linear model was fit to the data but the possibility of quadratic curvature in the response system existed. The authors developed the criteria for a spherically-shaped region of interest and found that the designs which minimized the MSE of response were similar to those which minimized the bias component alone, but were quite different from those which minimized the variance component alone.

In a later paper, Box and Draper (6) addressed the problem of fitting a second order polynomial model while protecting against third order bias ($d_1 = 2$, $d_2 = 3$). Again, using the MSE of response criteria and a spherically-shaped region of interest, the authors found that designs which minimized the MSE of response were similar to those which minimized the bias component alone. The authors suggested a rule of thumb for situations in which no information about the amount of bias error was available. The recommendation was that [ii], the pure second order design moment, be about 10% larger than would be suggested for the appropriate design which minimizes the bias component alone.

Draper and Lawrence (16) extended the work of Box and Draper by developing the MSE of response criteria for a cuboidal region of interest. The basic conclusions which emerged for cuboidal regions were consistent with the previous results for spherical regions. However, unlike the results for spherical regions, the best choice for the value of [ii] in a cuboidal region of interest is fairly insensitive to the number of factors in the model.

In a recent paper by Draper and Guttman (15), the previous developments for the MSE of response criteria in spherical and cuboidal regions were united. The authors developed a *family* of region shapes, ranging from all axial points to the cuboidal region. The conclusions reached for these flexible regions were again consistent with the previous results for spherical regions; that is, minimum MSE of response designs are slightly expanded versions of the designs which minimize the bias error alone. The authors noted that for any of the flexible design regions, if variance error and a moderate amount of bias error exist, then the design points should be placed to the inside of the boundary of the region of interest.

A generalization of the MSE of response criteria was proposed by Myers and Lahoda (48), in which the average mean squared error of a set parametric functions involving the response equation is minimized. Specific attention was given to the situation in which the parametric functions are the partial derivatives of the response equation, or equivalently, the slopes of the response function. This criteria (termed the *MSE of slope* criteria) has particular appeal in response surface optimization when the experimenter is interested in finding the *direction* of greatest change in the estimated response. In both the steepest ascent and ridge analysis procedures of RSM, it is important that the slopes of the response function be estimated well, and therefore an experimenter may find useful a criteria which takes both variance and bias of the slopes into account. Myers and Lahoda developed the MSE of slope criteria for both spherical and cuboidal regions of interest and for the fit-protection situations of $d_1 = 1, d_2 = 2$ and $d_1 = 2, d_2 = 3$. For the case of fitting a first order model, the authors found that designs which minimize the MSE of slope are orthogonal designs with [ii] as large as possible. For the case of fitting a second order model, the authors found that designs which minimized the MSE of slope were rotatable designs derivable from a class of designs which minimized the bias component alone.

The performance of three correlation induction strategies are evaluated in this research using the MSE of response and MSE of slope criteria for both spherical and cuboidal regions of interest. Chapter 3 is devoted to fitting first order response surface models while protecting against second order bias ($d_1 = 1$, $d_2 = 2$) and Chapter 4 focuses on fitting second order models while protecting against third order bias ($d_1 = 2$, $d_2 = 3$).

In the following two sections the mathematical developments of the MSE of response criteria (section 2.2.3.1) and the MSE of slope criteria (section 2.2.3.2) are presented.

2.2.3.1 MSE of Response Criteria

The assumption that underlies most research work in response surface methodology is that the *response* system can be adequately described by an equation of the form

$$response = model + error \qquad [2.2.28]$$

where the *model* states the effect of the input variables on the response variable and the *error* describes the general form of departures from the model. In response surface experimentation, it is generally assumed that the true model describing the entire system is unknown and perhaps extremely complicated. However, RSM techniques assume that within restricted regions of the input variables, low-order polynomial models (typically first or second order models) can be used to approximate the true model. The general linear model form of a polynomial model, as shown in equation [2.2.9] on page 39, is written as

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon}.$$

Low-order polynomial models are economical in terms of the number of experimental runs needed to estimate the model parameters, but these models can only approximate the true response function. Therefore, an experimenter may be uncertain as to whether or not the fitted polynomial model can adequately describe the true response surface. When model uncertainty exists, it is convenient to write the general linear model in a form which partitions the polynomial part of the equation into two parts; the $X_1\underline{\beta}_1$ fitted part and the $X_2\underline{\beta}_2$ unfitted part. The partitioned general linear model, as shown in equation [2.2.19] on page 42, is written as

$$\underline{y} = X_1\underline{\beta}_1 + X_2\underline{\beta}_2 + \underline{\varepsilon}.$$

The fitted part of the model includes the p_1 model coefficients in $\underline{\beta}_1$ and is of order d_1. The unfitted part of the model includes the p_2 model coefficients in $\underline{\beta}_2$ and is of order $d_2 - d_1$. The order of the *protection* model is d_2; that is, the highest order polynomial term which has been left out of the *fitted* model, but might be needed to adequately describe the response surface curvature in the region of interest. Often an experimenter cannot afford the number of experimental runs needed to fit a high-order polynomial model when the region containing the optimum response has not been located. Therefore, it may be desirable to use an experimental plan which affords protection against bias in the fitted model coefficients if higher order model terms are needed to adequately describe the response surface curvature.

The fitted polynomial model which the experimenter uses to *fit* the response data becomes

$$\hat{\underline{y}} = X_1\underline{b}_1 \qquad [2.2.29]$$

where $\hat{\underline{y}}$ is an $N\times1$ vector of the fitted responses at the data locations, X_1 is an $N\times p_1$ matrix of the regressor terms in the fitted model, and $\underline{b}_1$ is a $p_1\times1$ vector of the estimated model coefficients. The experimenter can estimate the value of the response variable at the u^{th} design point using the equation

$$\hat{y}_{(\underline{x}_u)} = \underline{x}_{1u}'\,\underline{b}_1 \qquad [2.2.30]$$

where $\hat{y}_{(\underline{x}_u)}$ is the fitted response at the u^{th} design point and $\underline{x}_{1u}$ is the u^{th} row of X_1.

However, if the experimenter has underfit the model, then the polynomial equation needed to describe the response surface would also include the unfitted model terms in X_2. Letting $\underline{x}_{2u}'$ denote the u^{th} row of X_2, the value of the true response at the u^{th} design point, $y_{(\underline{x}_u)}$, which includes both fitted and unfitted model terms, becomes

$$y_{(\underline{x}_u)} = \underline{x}_{1u}'\underline{\beta}_1 + \underline{x}_{2u}'\underline{\beta}_2 + \varepsilon_u . \qquad [2.2.31]$$

The MSE of response criteria developed by Box and Draper (5) calls for minimizing the average of the expected squared deviations of the true response, $y_{(\underline{x}_u)}$, from the estimated response, $\hat{y}_{(\underline{x}_u)}$, over the region of interest. The mean squared error of $\hat{y}_{(\underline{x}_u)}$ can be written as

$$\begin{aligned} \text{MSE}[\hat{y}_{(\underline{x}_u)}] &= E\{\hat{y}_{(\underline{x}_u)} - E[y_{(\underline{x}_u)}]\}^2 \\ &= \text{Var}[\hat{y}_{(\underline{x}_u)}] + \text{Bias}^2[\hat{y}_{(\underline{x}_u)}] \end{aligned} \qquad [2.2.32]$$

where the variance and bias terms are defined as

$$\text{Var}[\hat{y}_{(\underline{x}_u)}] = E\{\hat{y}_{(\underline{x}_u)} - E[\hat{y}_{(\underline{x}_u)}]\}^2 \qquad [2.2.33]$$

$$\text{Bias}^2[\hat{y}_{(\underline{x}_u)}] = \{E[\hat{y}_{(\underline{x}_u)}] - E[y_{(\underline{x}_u)}]\}^2 . \qquad [2.2.34]$$

The criteria proposed by Box and Draper (5) utilizes an average value of MSE $[\hat{y}_{(\underline{x}_u)}]$, obtained by uniformly weighting the mean squared error of the fitted responses at each of the data locations. This average value is computed by integrating the MSE $[\hat{y}_{(\underline{x}_u)}]$ over the region of interest and dividing by the volume of the region. The criteria also calls for normalizing the average MSE $[\hat{y}_{(\underline{x}_u)}]$ with respect to the number of design points, N, and the experimental error variance, σ^2. Box and Draper denote this average, normalized MSE of response as **J**, and define **J** as

$$
\begin{aligned}
\mathbf{J} &= \frac{N\Omega_r}{\sigma^2} \int_R \mathrm{MSE}\,[\hat{y}_{(\mathbf{x}_u)}]\, d\mathbf{x} \\
&= \frac{N\Omega_r}{\sigma^2} \int_R E\left\{\hat{y}_{(\mathbf{x}_u)} - E\,[y_{(\mathbf{x}_u)}]\right\}^2 d\mathbf{x} \\
&= \frac{N\Omega_r}{\sigma^2} \int_R E\left\{\hat{y}_{(\mathbf{x}_u)} - E\,[\hat{y}_{(\mathbf{x}_u)}]\right\}^2 d\mathbf{x} + \frac{N\Omega_r}{\sigma^2} \int_R \left\{E\,[\hat{y}_{(\mathbf{x}_u)}] - E\,[y_{(\mathbf{x}_u)}]\right\}^2 d\mathbf{x} \qquad [2.2.35] \\
&= \frac{N\Omega_r}{\sigma^2} \int_R \mathrm{Var}\,[\hat{y}_{(\mathbf{x}_u)}]\, d\mathbf{x} + \frac{N\Omega_r}{\sigma^2} \int_R \mathrm{Bias}^2\,[\hat{y}_{(\mathbf{x}_u)}]\, d\mathbf{x} \\
&= \mathbf{V} + \mathbf{B}
\end{aligned}
$$

where **V** is the average, normalized variance of $\hat{y}_{(\mathbf{x}_u)}$ $(u = 1, \ldots\ldots, N)$ over the region of interest and **B** is the average, normalized squared bias of $\hat{y}_{(\mathbf{x}_u)}$. **V** is termed the *variance error* and **B** is termed the *bias error.* Ω_r^{-1} is the volume of *R*, the centered and scaled region of interest, defined as

$$\Omega_r^{-1} = \int_R d\mathbf{x} \ .$$

The variance and bias error terms can be conveniently written in terms of the *region moment matrices*, which Box and Draper define as

$$\mu_{11} = \Omega_r \int_R \mathbf{x}_1{}'\, \mathbf{x}_1 \; d\mathbf{x} \qquad [2.2.36]$$

$$\mu_{21} = \Omega_r \int_R \mathbf{x}_2{}'\, \mathbf{x}_1 \; d\mathbf{x} \qquad [2.2.37]$$

$$\mu_{22} = \Omega_r \int_R \mathbf{x}_2{}'\, \mathbf{x}_2 \; d\mathbf{x} \ . \qquad [2.2.38]$$

The region moment matrices are obtained by integrating the $\mathbf{x}$ vectors over a specified region of interest. The $\mathbf{x}_1$ vector contains the p_1 regressor terms in X_1, and the $\mathbf{x}_2$ vector contains the p_2 regressor terms in X_2. It is assumed that the vectors are weighted uniformly over the region of interest; that is, the weight assigned to each design point is constant and equal to Ω_r, the inverse of the volume of the centered and scaled design region. The elements of the region moment matrices are of the form

$$w_{1^{\delta_1} 2^{\delta_2} \ldots k^{\delta_k}} = \Omega_T \int_R x_1^{\delta_1} x_2^{\delta_2} \ldots\ldots x_k^{\delta_k} \, d\underline{x} \qquad [2.2.39]$$

where $\sum_{i=1}^{k} \delta_i = \delta$ is the order of the region moment.

The form of the elements of the region moment matrices depend on the order of the fitted model (d_1), the order of the protection model (d_2), and the shape of the region of interest. The most frequently used regions shapes are the spherical and cuboidal regions. Appendix B (pages 278-282) illustrates the region moment matrices for the case of fitting a first order model and Appendix J (pages 348-352) illustrates the region moment matrices for the case of fitting a second order model. Appendix D (pages 287-291) illustrates the *elements* of the region moment matrices for both spherical and cuboidal regions of interest.

Utilizing the region moment notation, the variance error, **V**, becomes

$$\mathbf{V} = \frac{N\Omega_T}{\sigma^2} \int_R \mathrm{Var}[\hat{y}_{(\underline{x}_0)}]\, d\underline{x}$$

$$= \frac{N}{\sigma_2} \mathrm{Trace}\{\mathrm{Var}[\underline{b}_1]\, \mu_{11}\} \qquad [2.2.40]$$

where the form of Var [$\underline{b}_1$] depends on the least squares estimation technique used to estimate $\underline{\beta}_1$. Denoting the variance of $\underline{b}_1$ under ordinary least squares as Var [$\underline{b}_{1,OLS}$], and the variance under weighted least squares as Var [$\underline{b}_{1,WLS}$] , the two forms of the variance of $\underline{b}_1$ become

$$\mathrm{Var}[\underline{b}_{1,OLS}] = (X'_1 X_1)^{-1} X'_1 \; V \; X_1 (X'_1 X_1)^{-1} \sigma^2 \qquad [2.2.41]$$

$$\mathrm{Var}[\underline{b}_{1,WLS}] = (X'_1 V^{-1} X_1)^{-1} \sigma^2 \qquad [2.2.42]$$

where V is the variance-covariance matrix of the error terms apart from the experimental error; that is, $\text{Var}[\underline{\varepsilon}] = \sigma^2 V$.

The bias error, **B**, written in terms of the region moment matrices, becomes

$$\mathbf{B} = \frac{N\Omega_r}{\sigma^2} \int_R \text{Bias}^2 [\hat{y}_{(\underline{x}_u)}] \, d\underline{x} \qquad [2.2.43]$$

$$= \frac{N}{\sigma^2} \underline{\beta}'_2 \{A' \mu_{11} A - 2\mu_{21} A + \mu_{22}\} \underline{\beta}_2$$

where A is the *alias matrix* whose form depends on the least squares estimation technique used to estimate the model parameters. Denoting the alias matrix under ordinary least squares estimation as A_{OLS}, and under weighted least squares as A_{WLS}, the two forms of the alias matrix, as shown in equations [2.2.22] and [2.2.23] on page 43, are

$$A_{OLS} = (X'_1 X_1)^{-1} X'_1 X_2$$

$$A_{WLS} = (X'_1 V^{-1} X_1)^{-1} X'_1 V^{-1} X_2 .$$

Therefore, the MSE of response criteria developed by Box and Draper (5), and defined in equation [2.2.35] on page 71, becomes the minimization of

$$\mathbf{J} = \frac{N}{\sigma_2} \text{Trace}\{\text{Var}[\underline{b}_1] \, \mu_{11}\} + \frac{N}{\sigma^2} \underline{\beta}'_2 \{A' \mu_{11} A - 2\mu_{21} A + \mu_{22}\} \underline{\beta}_2 . \qquad [2.2.44]$$

2.2.3.2 *MSE of Slope Criteria*

A primary objective of many simulation studies is to find the conditions of the input variables which result in the optimum estimated response. The optimization procedures used in the analysis of response surfaces (discussed in section 2.2.1.2) utilize the partial derivatives of the response function for determination of the *direction* of the greatest change in the estimated response. Generally a series of sequential experiments is performed, each using the gradient information obtained in the previous experiment to determine the direction in which the response changes most rapidly. Before the region containing the optimum response is located, the experimenter's objective is to attain an equation for predicting the rate of change of the response variable with respect to each of the input variables (as opposed to the ultimate goal of attaining an equation for predicting the value of the response variable itself). Therefore, during the sequential experimentation stage of RSM, an experimenter may find useful a design criteria developed specifically for the purpose of estimating the slopes of the response function.

Myers and Lahoda (48) developed a generalization of the MSE of response criteria for the case of estimating a set of parametric functions of the response equation. The *MSE of slope* criteria is a specific application of the authors' criteria in which the parametric functions are the first derivatives, or slopes, of the response equation. The MSE of slope criteria seeks to minimize the average, normalized mean squared error of the slopes of the fitted response equation. Denoting the partial derivative of the response equation with respect to the i^{th} input variable as $\gamma_{(x_i)}$, the i^{th} partial derivative of the first order polynomial model shown in equation [2.2.4] on page 33 becomes

$$\gamma_{(x_i)} = \frac{\partial y}{\partial x_i} = \beta_i \ . \qquad [2.2.45]$$

The i^{th} partial derivative of the second order response function shown in equation [2.2.5] on page 33 can be written as

$$\gamma_{(x_i)} = \beta_i + 2\beta_{ii} x_i + \sum_{j \neq i} \beta_{ij} x_j \ . \qquad [2.2.46]$$

Similarly, the i^{th} partial derivative of the third order polynomial model shown in equation [2.2.24] on page 45 becomes

$$y_{(x_i)} = \beta_i + 2\beta_{ii} x_i + 3\beta_{iii} x_i^2 + \sum_{j \neq i} \beta_{ij} x_j + 2\sum_{j \neq i} \beta_{iij} x_i x_j$$
$$+ \sum_{j \neq i} \beta_{ijj} x_j^2 + \sum_{h \neq i} \sum_{< j \neq i} \beta_{hij} x_h x_j \ . \qquad [2.2.47]$$

The general linear model form of the response equation (shown in equation [2.2.9] on page 39) can be extended to the parametric functions involving the slopes of the response equation. Denoting the partial derivative of the response vector with respect to the vector of k input variables as $\underline{\gamma}_{(\underline{x})}$, the general linear model form of the parametric function involving the slopes of the response equation can be written as

$$\underline{\gamma}_{(\underline{x})} = \frac{\partial \underline{y}}{\partial \underline{x}} = \Lambda'_{(\underline{x})} \, \underline{\beta} \qquad [2.2.48]$$

where $\Lambda'_{(\underline{x})}$ is a matrix of the partial derivatives of the regressor terms in the X matrix, or equivalently, the partial derivatives of the response equation, apart from the β coefficients. The elements in the i[th] row of $\Lambda'_{(\underline{x})}$ are the partial derivatives of the i[th] row of X with respect to x_i.

When there is uncertainty concerning the ability of the fitted polynomial model to adequately describe the true response surface, it is useful to partition the general linear model into two parts (as shown in equation [2.2.19] on page 42). The corresponding partitioned form of the parametric function involving the slopes of the response equation can be written as

$$\underline{\gamma}_{(\underline{x})} = \Lambda'_{1\,(\underline{x})} \, \underline{\beta}_1 + \Lambda'_{2\,(\underline{x})} \, \underline{\beta}_2 \qquad [2.2.49]$$

where $\Lambda'_{1(\underline{x})}$ is a $k \times p_1$ matrix whose elements are the partial derivatives of the regressor terms in fitted X_1 portion of the model, and $\Lambda'_{2(\underline{x})}$ is a $k \times p_2$ matrix whose elements are the partial derivatives of the regressor terms in unfitted X_2 portion of the model. Denoting the partial derivative of the *fitted* response function shown in equation [2.2.29] on page 69 as $\hat{\underline{\gamma}}_{(\underline{x})}$, the equation for the slopes of the fitted response function becomes

$$\hat{\underline{\gamma}}_{(x)} = \frac{\partial \hat{\underline{y}}}{\partial \underline{x}} = \Lambda'_{1(x)} \, \underline{b}_1 \; . \qquad [2.2.50]$$

The form of the elements of the $\Lambda'_{1(x)}$ and $\Lambda'_{2(x)}$ matrices depend on the order of the fitted model (d_1) and the order of the protection model (d_2). For the case of fitting a first order response model with protection against a true second order model ($d_1 = 1$, $d_2 = 2$), the $\underline{\beta}$ vectors and the $\Lambda'_{(x)}$ matrices are shown in Appendix C (pages 283-286). For the case of fitting a second order model with protection against a true third order model ($d_1 = 2$, $d_2 = 3$), the $\underline{\beta}$ vectors and the $\Lambda'_{(x)}$ matrices are shown in Appendix K (pages 353-358).

The MSE of slope criteria developed by Myers and Lahoda (48) calls for minimizing the average of the expected squared deviations of the partial derivatives of the true response function, $\underline{\gamma}_{(x)}$, from the estimated partial derivatives, $\hat{\underline{\gamma}}_{(x)}$, over the region of interest, and the criteria is normalized with respect to the number of design points, N, and the experimental error variance, σ^2. Myers and Lahoda denote this average, normalized MSE of $\hat{\underline{\gamma}}_{(x)}$ as $\mathbf{J}^*$, and define $\mathbf{J}^*$ as

$$\begin{aligned} \mathbf{J}^* &= \frac{N\Omega_r}{\sigma^2} \int_R \mathrm{MSE}\,[\hat{\underline{\gamma}}_{(x)}]\, d\underline{x} \\ &= \frac{N\Omega_r}{\sigma^2} \int_R \mathrm{E}\left\{ [\hat{\underline{\gamma}}_{(x)} - \underline{\gamma}_{(x)}]' \, [\hat{\underline{\gamma}}_{(x)} - \underline{\gamma}_{(x)}] \right\} d\underline{x} \\ &= \frac{N\Omega_r}{\sigma^2} \int_R \mathrm{Var}\,[\hat{\underline{\gamma}}_{(x)}]\, d\underline{x} + \frac{N\Omega_r}{\sigma^2} \int_R \mathrm{Bias}^2\,[\hat{\underline{\gamma}}_{(x)}]\, d\underline{x} \\ &= \mathbf{V}^* + \mathbf{B}^* \end{aligned} \qquad [2.2.51]$$

where $\mathbf{V}^*$ is the average, normalized variance of $\hat{\underline{\gamma}}_{(x)}$, $\mathbf{B}^*$ is the average, normalized squared bias of $\hat{\underline{\gamma}}_{(x)}$, and Ω_r^{-1} is the volume of the region of interest (defined on page 71).

The variance and bias components of $\mathbf{J}^*$ can be written in terms of the region matrices of the partial derivatives of the response function. The region matrices are obtained by integrating the

$\Lambda_{(x)}$ matrices with respect to the vectors of input variables over a specified region of interest and dividing by the volume of the region. The region matrices involving the partial derivatives of the fitted and unfitted model terms are defined as

$$\mu^*_{11} = \Omega_r \int_R \Lambda_{1\,(\underline{x})}\, \Lambda'_{1\,(\underline{x})}\; d\underline{x} \qquad [2.2.52]$$

$$\mu^*_{21} = \Omega_r \int_R \Lambda_{2\,(\underline{x})}\, \Lambda'_{1\,(\underline{x})}\; d\underline{x} \qquad [2.2.53]$$

$$\mu^*_{22} = \Omega_r \int_R \Lambda_{2\,(\underline{x})}\, \Lambda'_{2\,(\underline{x})}\; d\underline{x}\ . \qquad [2.2.54]$$

The form of the elements of these region matrices depend on the order of the fitted model, the order of the protection model, and the shape of the region of interest. Appendix C (pages 283-286) illustrates the region matrices of the partial derivatives of the response function for the case of fitting a first order model and Appendix K (pages 353-358) illustrates these matrices for the case of fitting a second order model. Appendix D (pages 287-291) illustrates the *elements* of the region matrices for both spherical and cuboidal regions of interest.

The variance component of $\mathbf{J}^*$, written in terms of the region matrices, becomes

$$\begin{aligned}
\mathbf{V}^* &= \frac{N\Omega_r}{\sigma^2} \int_R \mathrm{Var}\,[\,\hat{\underline{y}}_{(\underline{x})}\,]\, d\underline{x} \\
&= \frac{N\Omega_r}{\sigma^2} \int_R E\left\{ [\,\hat{\underline{y}}_{(\underline{x})} - E(\hat{\underline{y}}_{(\underline{x})})]'\ [\,\hat{\underline{y}}_{(\underline{x})} - E(\hat{\underline{y}}_{(\underline{x})})\,] \right\} d\underline{x} \qquad [2.2.55] \\
&= \frac{N}{\sigma_2}\ \mathrm{Trace}\left\{ \mathrm{Var}\,[\,\underline{b}_1\,]\ \mu^*_{11} \right\}
\end{aligned}$$

where the form of $\mathrm{Var}\,[\,\underline{b}_1\,]$ depends on the least squares estimation technique used to estimate $\underline{\beta}_1$ (ordinary or weighted), as shown in equations [2.2.41] and [2.2.42] on page 72. The bias component of $\mathbf{J}^*$, written in terms of the region matrices, becomes

$$\mathbf{B}^* = \frac{N\Omega_r}{\sigma^2} \int_R \text{Bias}^2 [\hat{\underline{y}}_{(\underline{x})}] \, d\underline{x}$$

$$= \frac{N\Omega_r}{\sigma^2} \int_R \left\{ [E(\hat{\underline{y}}_{(\underline{x})}) - \underline{y}_{(\underline{x})}]' \; [E(\hat{\underline{y}}_{(\underline{x})}) - \underline{y}_{(\underline{x})}] \right\} d\underline{x} \qquad [2.2.56]$$

$$= \frac{N}{\sigma^2} \, \underline{\beta}'_2 \, \{ A' \mu^*_{11} A - 2 \mu^*_{21} A + \mu^*_{22} \} \, \underline{\beta}_2$$

where the form of A, the alias matrix, depends on the least squares estimation technique used to estimate the model parameters, as shown in equations [2.2.22] and [2.2.23] on page 43.

Therefore, the MSE of slope criteria developed by Myers and Lahoda (48), and defined in equation [2.2.51] on page 76, becomes the minimization of

$$\mathbf{J}^* = \frac{N}{\sigma_2} \, \text{Trace} \, \{ \text{Var} [\, \underline{b}_1 \,] \, \mu^*_{11} \} + \frac{N}{\sigma^2} \, \underline{\beta}'_2 \, \{ A' \mu^*_{11} A - 2 \mu^*_{21} A + \mu^*_{22} \} \, \underline{\beta}_2 \, . \qquad [2.2.57]$$

The design criteria of MSE of response and MSE of slope are used in this research to evaluate the performance of the three correlation induction strategies. The strategies are compared in terms of **J** and **J*** for a variety of first and second order response surface designs. Chapter 3 evaluates the strategies using first order response surface designs and Chapter 4 considers second order designs.

Chapter 3

FIRST ORDER DESIGNS

First order response surface designs are experimental plans useful in the estimation of the parameters of first order polynomial models. Such models require few design points and are generally used in restricted regions of the factor space to obtain preliminary estimates of the relationship between the response variable and a set of input variables. Using first order designs, these preliminary experiments are typically performed in a sequential manner in order to locate the region in the factor space in which the response appears to be optimized. Once this region is located, the experimenter abandons the first order models and fits a large, full-scale, second order design to enable more accurate estimation of the optimum response.

The properties and specific classes of first order designs were introduced in section 2.2.2.1 on pages 47-50. The first order designs which are the focus of the research in this Chapter are the two-level factorial designs. These designs are the most widely used class of first order designs because they are simple, yet provide highly efficient estimates of the model parameters. In addition, the two-level factorial designs are capable of being blocked orthogonally (a requirement of the assignment rule correlation induction strategy), and therefore these designs can be used to evaluate the three correlation induction strategies examined in this research.

The first section of this Chapter examines the two-level factorial design class and its inherent properties of importance to this research (orthogonal blocking, accommodation of center runs, and fractional replication). The second section utilizes the two-level factorial designs to evaluate the three correlation induction strategies in terms of the MSE of response criteria, and the third section evaluates these strategies in terms of the MSE of slope criteria. The final section presents a summary of the research findings in the first order case.

3.1 Two-level Factorial Designs

Factorial designs are useful experimental plans for simulation studies in which the joint effect of several input variables on a simulated response variable is to be studied. The two-level factorial designs involve the use of k input variables at each of two levels, thereby requiring 2^k design points for a complete factorial arrangement. Because there are only two levels of each factor, these designs are only useful in fitting first order models. Thus, an experimenter using a 2^k design must be willing to assume that the response is approximately linear over the range of the factor levels chosen. Despite their inability to fit curvature, the 2^k designs are extremely useful during the steepest ascent stage of RSM. With only a small number of design points, the 2^k designs provide efficient estimates of the linear coefficients of first order models and, in turn, are useful for location of the region containing the optimum response.

When fitting a first order response surface model using a 2^k design, the input variables are generally coded for each region of experimentation such that -1 represents the low level of each variable and $+1$ represents the high level of each variable. Each time the experimenter moves to a new region of interest, the input variables must be re-coded to reflect changes in the size and/or central location of the design relative to the response surface. The levels of the i^{th} coded design variable in the current region of interest, denoted by x_i $(i = 1, 2, \ldots, k)$, are determined from the uncoded, or natural, variables using the transformation shown in equation [2.2.8] on page 34. The two levels of the coded design variables used in the 2^k experiments are not necessarily the ± 1, or high/low, levels. Depending on the design criteria chosen by the experimenter, the design may be slightly larger or smaller than the cuboid defined by the factorial arrangement of the ± 1 levels. Variance-related design criteria generally suggest the use of larger designs and bias-related design criteria generally suggest the use of smaller designs.

The remainder of this section discusses the ability of the 2^k designs to block orthogonally, accommodate center runs, and be fractionally replicated. The following two-level factorial design, with $k = 3$ factors and $N_c = 1$ center run, is used to illustrate the properties of the 2^k designs:

$$
D = \begin{array}{c}
\begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\
\left[\begin{array}{ccc}
-1 & -1 & -1 \\
1 & 1 & -1 \\
1 & -1 & 1 \\
-1 & 1 & 1 \\
\hdashline
1 & -1 & -1 \\
-1 & 1 & -1 \\
-1 & -1 & 1 \\
1 & 1 & 1 \\
0 & 0 & 0
\end{array} \right]
\end{array} \times g
$$

where g is the scaling factor, or equivalently, the levels of the coded design points. The factorial design points lie outside the current region of experimentation when the value of g is larger than 1, and these design points lie inside the boundaries of the current region when the value of g is smaller than 1.

3.1.1 Orthogonal Blocking in the Two-level Factorial Designs

First order designs which block orthogonally are those designs which can be partitioned into two or more blocks in such a manner that the block effects do not affect the usual estimates of the linear coefficients. It is generally assumed that each block contributes only an *additive* effect to the first order model and, therefore, any model terms which are orthogonal to the block effects are unaffected by the blocking scheme. In order to partition a 2^k design into 2^p blocks, an experimenter must choose $p < k$ defining contrasts and, in turn, $2^p - p - 1$ generalized interactions will result. The defining contrasts and generalized interactions are indistinguishable from, or *confounded* with, the block effects because these terms are not orthogonal to the block effects. Thus, the experimenter must choose the defining contrasts in such a manner that the usual estimates of the first order model coefficients are not confounded with the blocks effects; that is, the linear model terms cannot be defining contrasts nor generalized interactions. See references (40) and (46) for complete coverage of blocking in the 2^k designs.

Any of the two-level factorial designs can be partitioned into $2^p = 2$ orthogonal blocks through the use of an interaction term as the $p = 1$ defining contrast. For example, the 2^3 design shown on page 81 was partitioned into two orthogonal blocks by using the $x_1x_2x_3$ term as defining contrast. Only the β_{123} coefficient is not estimable because it is confounded with the block effects, and therefore the β_0, β_1, β_2, and β_3 linear coefficients of the first order model can be estimated. The center run can be placed in either block because the coefficients of the first order model remain orthogonal even if the number of center runs in each block is not the same. (The addition of center runs to the 2^k designs is further discussed in section 3.1.2.)

For the case of $2^p = 4$ orthogonal blocks, $p = 2$ defining contrasts must be chosen and $2^p - p - 1 = 1$ generalized interaction results. To insure that no first order terms are confounded with the block effects, there must be at least $k = 3$ factors in the design. For example, the 2^3 design shown on page 81 can be partitioned into four orthogonal blocks through the use of the x_1x_2 and

x_2x_3 terms as defining contrasts, resulting in the x_1x_3 generalized interaction term. The partitioned design matrix becomes

$$
D = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \left[\begin{array}{ccc} -1 & -1 & -1 \\ 1 & 1 & 1 \\ \hline 1 & -1 & -1 \\ -1 & 1 & 1 \\ \hline -1 & -1 & 1 \\ 1 & 1 & -1 \\ \hline -1 & 1 & -1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{array} \right] \times g \end{array} .
$$

This design allows for the estimation of the linear coefficients of the first order model because only the β_{12} and β_{23} coefficients of the defining contrasts and the β_{13} coefficient of the generalized interaction are confounded with the block effects. Table 1 summarizes the number of factors required, the number of defining contrasts which must be chosen, and the number of generalized interactions which result when a 2^k design is partitioned into 2^p = *2, 4, 8, 16, 32, 64* blocks. To insure that the blocking is orthogonal, the linear terms of the model cannot be defining contrasts nor generalized interactions.

Table 1. Orthogonal Blocking in the Two-Level Factorial Designs.
Required number of factors and defining contrasts.

Number of Orthogonal Blocks (2^p)	Required Number of Factors ($k \geq p+1$)	Number of Defining Contrasts (p)	Number of Generalized Interactions ($2^p - p - 1$)
2	$k \geq 2$	1	0
4	$k \geq 3$	2	1
8	$k \geq 4$	3	4
16	$k \geq 5$	4	11
32	$k \geq 6$	5	26
64	$k \geq 7$	6	57

In the remainder of this Chapter it is assumed that an experimenter utilizing the assignment rule correlation induction strategy has partitioned the 2^k design into blocks in such a manner that the first order effects are orthogonal to the block effects.

3.1.2 Addition of Center Runs to the Two-level Factorial Designs

A center run is a design point taken at the center of the current experimental region; that is, at the intermediate level $(x_i = 0)$ of each coded design variable. By definition, the two-level factorial designs consist of all combinations of *two* levels of each factor. Therefore, by augmenting the 2^k designs with center runs, these designs become three-level designs. However, the disadvantages of an added level of each factor are often outweighed by the advantages obtained through the addition of center runs. The addition of multiple center runs enables estimation of the experimental error variance, provides a means of checking the adequacy of the first order model, and affords protection against bias in the model coefficients.

One of the disadvantages of the two-level factorial designs is that they do not permit estimation of the pure experimental error variance, σ^2. The estimate of the error variance obtained from fitting a first order model with a 2^k design contains both pure experimental error and error due to inadequacy of the model. Through the addition of multiple center runs, the *error sum of squares* (the usual estimate of the error variance, apart from its degrees of freedom) can be partitioned into three component parts:

1. Pure experimental error
2. Lack-of-fit error due to missing quadratic terms
3. Lack-of-fit error due to missing interaction terms.

The pure experimental error variance is estimated from the variability of the responses obtained on replicated center runs and the lack-of-fit error due to missing quadratic terms is estimated using the difference between the average response at the center of the design and the average response at the factorial points. By taking a ratio of the quadratic lack-of-fit error to the pure experimental error, an experimenter can perform a formal hypothesis test for overall quadratic curvature, thereby assessing whether or not the first order model adequately describes the response surface in the current region of interest. The remaining portion of the error sum of squares is attributed to lack-of-fit error due to missing interaction terms. Thus, by augmenting a 2^k design with multiple center runs, the experimenter can estimate the pure experimental error variance and assess the adequacy of the fitted first order model.

Another disadvantage of the two-level factorial designs is that they do not offer protection against bias in the model coefficients when all of the design points are set at the ± 1 levels. When these designs are not augmented with center runs and all of the factorial points are placed at the ± 1 extremes, the designs are generally optimal in terms of most variance-related design criteria, but they do not offer protection against bias in the coefficients when the model is inadequate. Bias-related design criteria, which take model inadequacy into account, generally suggest the use of smaller designs; that is, placing the factorial points at $\pm g$, where $g < 1$. Mean squared error design criteria, which take both variance and bias error into account, typically suggest the use of designs which are slightly expanded versions of the designs which consider bias alone.

The *size* of factorial designs can be made smaller by moving the factorial points closer to the center of the design and/or by augmenting the designs with center runs. Box and Draper (5) define the size of a design as the root mean square distance of the design points from the center of the region, which for a factorial design is simply the square root of k [ii]. Thus, for a given number of factors in the model, k, the size of a 2^k design is determined by the value of [ii], the pure second order design moment, which is computed as

$$[ii] = \frac{1}{N} \sum_{u=1}^{N} x_{iu}^2 = \frac{F g^2}{F + N_c} \qquad [3.1.1]$$

where F is the number of factorial points, N_c is the number of center runs, and g is the level of the coded factorial design points. From equation [3.1.1], it is evident that either increasing the value of N_c or decreasing the value of g will reduce the value of [ii] and, therefore, reduce the size of the design.

The addition of center runs for the purpose of reducing the size of the design, thereby affording protection against bias in the model coefficients, is viable technique if the experimenter can afford the number of center runs required. It is generally more efficient to reduce the size of the design by reducing the value of g; that is, by moving the factorial points closer to the center of the design. The number of center runs required to provide proper protection against bias is frequently more than the experimenter can afford, yet only a slight reduction in the value of g easily reduces the design to the desired size. For example, to achieve $[ii] = 0.4$ without the addition of center runs requires that the factorial points be placed at levels of $\pm .63$. To achieve $[ii] = 0.4$ without altering the ± 1 levels requires the use of 12 center runs for a 2^3 design and 24 center runs for a 2^4 design, each of which represents a 150% increase in the total number of design points. However, augmenting the 2^k designs with just a few center runs reduces the size of the design slightly. For example, without altering the value of ± 1 factorial points, the addition of $N_c = 3$ center runs to a 2^3 design reduces the size of the design to $[ii] = .73$ and reduces the size of a 2^4 design to $[ii] = .84$. These reductions in the value of [ii] through the addition of three center runs provide protection against biased model coefficients and enable a formal test for quadratic curvature to be performed.

Table 2 on page 87 summarizes the value of g required to achieve [ii] = *1.0, .8, .6, .4, .2* without the addition of center runs in the 2^k designs. The table also indicates the number of center runs needed for an equivalent reduction in the size of the design without altering the value of g. Because the number of center runs depends on the number of factors in the model, the table indicates the required values of N_c for 2^k designs with $k = 2, 3, 4, 5$ factors.

Table 2. Altering the Size of Two-level Factorial Designs.
The required levels of the factorial points or the required number of center runs.

Size of the Design [ii]	Level of the Factorial Points When $N_c = 0$ (g)	Required Number of Center Runs When $g = 1$ (N_c)			
		k=2	k=3	k=4	k=5
1.0	1.00	0	0	0	0
.8	.89	1	2	4	8
.6	.77	2.7	5.3	10.7	21.3
.4	.63	6	12	24	48
.2	.45	16	32	64	128

When the factorial points are placed at the ± 1 high/low levels, the addition of center runs affects slight reductions in the second order design moments. However, to achieve large reductions in the size of the design, say reducing [ii] by $^1/_3$ or more, requires that at least $^1/_3$ of the design points be center runs, and therefore an increase of at least 50% in the total number of design points. Thus, as the desired value of [ii] decreases, the required number of center runs increases rapidly, and it becomes necessary to move the factorial points closer to the center of the experimental region and away from the ± 1 extremes in order to achieve the desired design size. Also, the required number of center runs required is not necessarily integer-valued, and the experimenter may need to alter the value of g even if the design is augmented with center runs.

The advantages of the simultaneous use of center runs and moving the factorial points closer to the center of the design are threefold:

1. An estimate of the pure experimental error variance (σ^2) can be obtained.
2. The adequacy of the first order model can be ascertained through a formal test for quadratic curvature.

3. The experimenter obtains protection against biased coefficients when the fitted model is inadequate.

3.1.3 Fractional Replication of Two-level Factorial Designs

As the number of factors in the two-level factorial designs increase, the number of experimental runs needed for a complete replication rapidly becomes unaffordable. Fractional replications of the 2^k designs, which consist of a fraction of the complete factorial experiment, are particularly useful in response surface analysis when five or more factors are involved. The reduction in the number of design points is accomplished at the expense of interaction terms; that is, the coefficients of certain interaction terms cannot be estimated because they are indistinguishable from the linear coefficients of the first order model.

The procedures for developing a fractional factorial design are similar to those used in the development of orthogonal blocking arrangements (discussed on page 82). Once a 2^k design has been partitioned into orthogonal blocks, a fractional replication of the design is achieved by using the design points in only *one* of the blocks. Recall that when a full 2^k design is run in blocks, the defining contrasts and generalized interaction terms are confounded with the block effects. However, when a 2^k design is fractionally replicated, the defining contrasts and generalized interaction terms define a particular *alias* structure. Associated with each of the linear terms in the first order model are alias effects which are indistinguishable from the main effects. Each defining contrast and generalized interaction term generate an effect that is aliased with the estimable effects of the regression model.

The 2^3 factorial design shown on page 81 was partitioned into two orthogonal blocks utilizing the $x_1x_2x_3$ term as the defining contrast. If only one of the blocks were used as the experimental design, then the design would be considered a ½-fractional replication. Utilizing the second block of the partitioned 2^3 design, the ½-fractional replicate becomes

$$D = \begin{array}{c} \\ \\ \\ \\ \\ \end{array}\begin{array}{ccc} x_1 & x_2 & x_3 \end{array}$$

$$D = \left[\begin{array}{ccc} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right] \times g \, .$$

This fractional replication was generated using the $x_1x_2x_3$ term as the defining contrast, and therefore the alias structure of the design can be written as

$$\begin{aligned} x_1 &\equiv x_2x_3 \\ x_2 &\equiv x_1x_3 \\ x_3 &\equiv x_1x_2 \\ (1) &\equiv x_1x_2x_3 \end{aligned}$$

where $\equiv$ indicates that the effects are aliased with each other. Thus, the β_i linear coefficients are estimable but the β_{ij} two-way interaction coefficients cannot be estimated because they are aliased with (indistinguishable from) the linear coefficients. In addition, the estimate of the β_{123} coefficient of the defining contrast is completely sacrificed because it is aliased with the β_0 intercept term.

Fractional replications of the 2^k designs can be partitioned into orthogonal blocks provided that the generalized interactions between the defining contrasts and blocking contrasts are second order, or higher, interaction terms. Therefore, orthogonal blocking is possible in ½ fractions of the 2^k designs with $k \geq 4$ factors, and ¼ fractions can be blocked orthogonally when there are $k \geq 5$ factors in the design.

The research in the remainder of this Chapter is an evaluation of the correlation induction strategies in the first order case. The first order designs used to evaluate the three schemes are the following full and fractional two-level factorial designs:

1. Full 2^k factorial designs with $k = 1, 2, 3, 4$ factors
2. ½-fractional replicates of the 2^k designs with $k = 5, 6$ factors
3. ¼-fractional replicate of the 2^k design with $k = 7$ factors.

3.2 *MSE of Response Criteria in the First Order Case*

The focus of this section is the evaluation of three correlation induction strategies for the situation in which an experimenter is fitting a first order model using a two-level factorial design. The three strategies (introduced on pages 26-30) are the use of independent random number streams for each design point (IR), the use of common streams for non-replicated design points and independent streams for replicated design points (CR), and the use of the assignment rule blocking strategy for non-replicated design points within blocks and independent streams for replicated design points within blocks (AR).

The criteria used to evaluate the correlation induction strategies are two mean squared error design criteria (introduced on pages 65-78). The first criteria, MSE of response, involves minimizing the integrated mean squared error of the predicted response variable over the region of interest, and the second criteria, MSE of slope, involves minimizing the integrated mean squared error of the slopes of the response function. This section of the paper evaluates the correlation induction strategies in terms of the MSE of response criteria, and the MSE of slope criteria serves as the performance measure in the following section.

The average, normalized, mean squared error of the predicted response, denoted by $\mathbf{J}$, is the sum of the bias and variance errors, denoted by $\mathbf{B}$ and $\mathbf{V}$, respectively. These two individual components of $\mathbf{J}$, as well as their sum, are used to evaluate the correlation induction strategies. The first part of this section deals with the bias component alone, the second part considers with the variance component alone, and in the final part the correlation induction strategies are evaluated in terms of the MSE of response criteria.

The research of this Chapter assumes that a two-level factorial design is being used to model the relationship between a set of input variables and a response variable. The equation which the experimenter uses to fit the data is a first order polynomial regression model; that is, the order of the fitted model is $d_1 = 1$. It is also assumed that the experimenter desires protection against bias in the first order coefficients due to the presence of any quadratic curvature in the response surface; that is, the order of the protection model is $d_2 = 2$. Therefore, the equations for the fitted and true response at the u^{th} setting of the input variables, as shown in equations [2.2.30] and [2.2.31] on pages 69-70, respectively, are

$$\hat{y}_{(\underline{x}_u)} = \underline{x}_{1u}'\underline{b}_1$$

$$y_{(\underline{x}_u)} = \underline{x}_{1u}'\underline{\beta}_1 + \underline{x}_{2u}'\underline{\beta}_2 + \varepsilon_u$$

where the subscript $_1$ indicates that the vectors contain terms from the fitted first order model, and the subscript $_2$ indicates that the vectors contain the interaction and quadratic terms from the protection model. The bias and variance components of the mean squared error of $\hat{y}_{(\underline{x}_u)}$ are presented in the following two sections.

3.2.1 Bias Component of J in the First Order Case

The bias component of **J** is computed as the average squared bias of the predicted responses within the region of interest, normalized with respect to the number of design points and the experimental error variance. Box and Draper (5) present a mathematical development of the bias error, **B**, for the situation in which the responses are *uncorrelated* and the region of interest is spherically-shaped and Draper and Lawrence (16) extend the bias error equation to cuboidal-shaped regions. The research presented herein extends the previous work to situations in which the responses are *correlated.*

Box and Draper (5) define the bias component of the MSE of response criteria as

$$\mathbf{B} = \frac{N\Omega_r}{\sigma^2} \int_R \text{Bias}^2 [\hat{y}_{(\underline{x}_u)}] \, d\underline{x}$$

$$= \frac{N\Omega_r}{\sigma^2} \int_R \{E[\hat{y}_{(\underline{x}_u)}] - E[y_{(\underline{x}_u)}]\}^2 \, d\underline{x} \qquad [3.2.1]$$

$$= \frac{N}{\sigma^2} \underline{\beta}'_2 \{A' \mu_{11} A - 2\mu_{21} A + \mu_{22}\} \underline{\beta}_2$$

where Ω_r^{-1} is the volume of the region of interest (defined on page 71), A is the alias matrix (defined on page 43), and the μ terms are the region moment matrices of the design (defined on page 71).

The equation for **B** indicates that the bias error is a function of the unknown β parameters in the protection model. Therefore, the bias error cannot be computed using equation [3.2.1] unless the β coefficients of the unfitted interaction and quadratic terms can be estimated. For the case of fitting a first order model with protection against a true second order model, estimation of $\underline{\beta}_2$ requires estimation of $p_2 = \frac{1}{2}k(k+1)$ coefficients, and this number increases rapidly as k increases. However, if the value of **B** is averaged over all orthogonal rotations of the response surface, then the number of parameters which need to be estimated can be reduced to two, regardless of the value of k. Box and Draper (5) develop the equations for the average value of **B** using the following two parameters:

$$\theta = \frac{N}{\sigma^2} \left[\sum_{i=1}^{k} \beta_{ii}^2 + \frac{1}{2} \sum_{i} \sum_{<\, j} \beta_{ij}^2 \right] \qquad [3.2.2]$$

$$\phi = \frac{(\sum_{i=1}^{k} \beta_{ii})^2}{\sum_{i=1}^{k} \beta_{ii}^2 + \frac{1}{2} \sum_{i} \sum_{<\, j} \beta_{ij}^2} \quad ; \qquad (0 \le \phi \le k) \qquad [3.2.3]$$

where θ is a standardized measure of the magnitude of the quadratic tendency of the response surface and ϕ is a parameter which indicates the state of conditioning of the response surface. The best state of conditioning corresponds to a response surface with spherical response contours about the canonical axes, occurring when $\phi = k$, and the worst state of conditioning corresponds to a response surface with redundant canonical variables, occurring when $\phi = 0$. The product of these two parameters, $\phi\theta$, is a standardized and squared sum of the quadratic coefficients, which can be written as

$$\phi\,\theta = \frac{N}{\sigma^2}\,(\sum_{i=1}^{k}\beta_{ii})^2 \,. \qquad [3.2.4]$$

Using the ϕ and θ parameters, Box and Draper (5) develop an equation for the average value of $\mathbf{B}$ for situations in which ordinary least squares (OLS) is the appropriate technique for estimation of the β coefficients. However, when the responses are correlated due to the use of the modified CR and AR strategies, weighted least squares (WLS) is the appropriate estimation technique. The parameter estimates are different under OLS and WLS, and these differences are reflected in the bias component of $\mathbf{J}$ through the form of the alias matrix. However, when there are no replicated design points, indicating that the CR and AR strategies are pure, Schruben and Margolin (56) have shown that the OLS and WLS parameter estimates are equivalent.

The equation for the average value of $\mathbf{B}$ developed by Box and Draper (5) can be generalized to the situation in which the responses are correlated and weighted least squares is the appropriate estimation technique. The modified form of the equation for $\mathbf{B}$ becomes

$$\mathbf{B} = \theta\left[\phi\,(z\,[ii] - w_{ii})^2 + 2\,y_r\,w_{ii}\right] \qquad [3.2.5]$$

where $[ii]$ is the pure second order design moment, w_{ii} is the pure second order region moment, y_r is a region-dependent parameter defined for convenience reasons, and z is a coefficient of $[ii]$ which is a function of both the least squares estimation technique and the correlation induction strategy.

For a two-level factorial design, the pure second order design moment (shown in equation [3.1.1] on page 86) is computed as

$$[ii] = \frac{1}{N} \sum_{u=1}^{N} x_{iu}^2$$

$$= \frac{F g^2}{F + N_c} \ .$$

The w_{ii} term in the equation for **B** is the pure second order region moment which is defined in equation [2.2.39] on page 72. As shown in Appendix D (pages 287-291), w_{ii} is a function of k in a spherical region, but w_{ii} is equal to a constant in a cuboidal region. The y_r term, defined for convenience reasons, is also a function of the shape of the region of interest. For a spherical region, the y_r and w_{ii} terms can be expressed as

$$y_r = \frac{k + 2 - \phi}{(k+2)(k+4)} \qquad [3.2.6]$$

$$w_{ii} = \frac{1}{k+2} \qquad [3.2.7]$$

and for a cuboidal region these terms become

$$y_r = \frac{4 + 5k - 3\phi}{15(k+2)} \qquad [3.2.8]$$

$$w_{ii} = \frac{1}{3} \ . \qquad [3.2.9]$$

The only quantity in the equation for **B** which is a function of the least squares estimation technique is the z term. Under OLS estimation, $z = 1$, and the equation for **B** is the same as the equation developed in the research of Box and Draper (5). However, under WLS, the value of z is a function of the correlation induction strategy and the number of center runs in the design.

Before defining z for each correlation induction strategy it is useful to define some notation for the number of design points utilizing independent, common, and antithetic random number streams. In developing the notation, it is assumed that the designs may be augmented with any number of center runs and the factorial points are not replicated. The notation used to designate the number of design points in the 2^k designs is the following:

$$
\begin{aligned}
N &= \text{total number of design points} \\
&= F + N_c \\
&= N_1 + N_2
\end{aligned}
$$

where
F = number of factorial points (2^k or 2^{k-p})
N_c = number of center points
N_1 = number of design points utilizing *common* or *antithetic* streams
N_2 = number of design points utilizing *independent* streams.

The number of design points using independent random number streams (N_2) under each correlation induction strategy becomes

IR Strategy:	$N_2 = N$	
CR Strategy:	$N_2 = N_c - 1$	if $N_c > 1$
	$= 0$	if $N_c \leq 1$
AR Strategy:	$N_2 = N_c - 2$	if $N_c > 1$
	$= N_c$	if $N_c \leq 1$.

The IR strategy uses independent random number streams for each design point and, therefore, the value of N_2 is equal to the the total number of design points. The CR strategy uses common streams for non-replicated design points and independent streams for replications of design points. Therefore, N_2 is equal to the number of replicated center runs, or equivalently, one less than the total number of center runs. The AR strategy uses common and antithetic streams for non-replicated design points within orthogonal blocks and independent streams for replications of design points within blocks. Therefore, N_2 is equal to the total number of replicated center runs within blocks, or equivalently, two less than the total number of center runs.

The z term in the equation for **B** is equal to one for all three correlation induction strategies under OLS estimation. Under WLS, however, the values of z depend on the number of design points using independent, common, and antithetic random number streams. The WLS values of z for the **IR**, **CR**, and **AR** correlation induction strategies, respectively, become

$$z_{IR} = 1 \qquad [3.2.10]$$

$$z_{CR} = \frac{N}{N - N_2\,\rho_+ \; + N_1 N_2\,\rho_+} \qquad [3.2.11]$$

$$z_{AR} = \frac{N}{N - N_2\,\rho_+ \; + \tfrac{1}{2} N_1 N_2\,(\rho_+ - \rho_-)} \; . \qquad [3.2.12]$$

Equation [3.2.10] indicates that z, and therefore the equation for **B**, are the same under OLS and WLS estimation for the **IR** strategy. This result is due to the use of independent random number streams, thereby generating independent sample responses. Equations [3.2.11] and [3.2.12] indicate that **B** is a function of N_1, N_2, ρ_+, and ρ_- for the modified **CR** and **AR** strategies. However, when $N_2 = 0$, and the **CR** and **AR** strategies are pure, the equations for z simplify to $z_{CR} = z_{AR} = 1$ and **B** is the same under OLS and WLS estimation. This result is in agreement with the research of Schruben and Margolin (56), in which the authors show the equivalence of the OLS and WLS estimators for the pure **CR** and **AR** strategies. When $N_2 > 0$, indicating the use of the modified **CR** and **AR** strategies, the z term, and therefore the equation for **B**, are different under OLS and WLS due to differences in the alias matrix. Appendix E (pages 292-296) illustrates the alias matrix for the two-level factorial designs and shows that the relationship between the OLS and WLS alias matrices is the following: $A_{WLS} = z\,A_{OLS}$. The equation for the average value of **B** (equation [3.2.5] on page 93) can be derived from equation [3.2.1] on page 92 through the use of the WLS alias matrix (instead of the OLS alias matrix which Box and Draper utilized).

Equation [3.2.5] on page 93 can now be used to find the value of [ii] which results in a Min-**B** design. The optimal value of [ii] is found by setting the partial derivative of **B** with respect to [ii] equal to zero, yielding

$$\frac{\partial \mathbf{B}}{\partial\,[\mathrm{ii}]} \;=\; 2\,\phi\theta\, z\,(z\,[\mathrm{ii}] - w_{ii}) \;=\; 0 \;. \qquad [3.2.13]$$

The second derivative of **B** with respect to [ii], $\frac{\partial^2 \mathbf{B}}{\partial\,[\mathrm{ii}]^2} = 2\,\phi\theta\, z^2$, is always a positive quantity, and therefore the solution of equation [3.2.13] provides the value of [ii] which minimizes **B**. The **B**-optimal value of the pure second order design moment for a 2^k design becomes

$$[\mathrm{ii}] \;=\; \frac{w_{ii}}{z} \;. \qquad [3.2.14]$$

Under OLS estimation, $z = 1$, and **B** is minimized when [ii] is equal to w_{ii}. Therefore, the Min-**B** value of [ii] is $1/3$ in a cuboidal region and $1/(k+2)$ in a spherical region. These optimal values of [ii] are also the optimal values under WLS whenever $z = 1$; that is, for the IR, pure CR, and pure AR strategies. However, if replicated center runs necessitate the use of the modified CR and AR strategies, then the optimal value of [ii] under WLS depends on the value of z.

Incorporating the z_{CR} term on the previous page into equation [3.2.14] indicates that the optimal value of [ii] for the CR strategy under WLS is slightly larger than the OLS value, and the optimal value of [ii] increases as ρ_+ increases. Incorporating the z_{AR} term into equation [3.2.14] indicates that the optimal value of [ii] for the AR strategy under WLS is also slightly larger than the OLS value, except the reverse is true when ρ_+ and ρ_- are equal in magnitude. The optimal value of [ii] increases as ρ_+ increases, and decreases as ρ_- increases. Thus, the size of the Min-**B** designs tend to be slightly larger for the modified CR and AR strategies than for the pure strategies.

The relationships between the B-optimal value of [ii] and ρ_+, ρ_-, k, and N_c for a weighted least squares analysis are plotted in Appendix G (pages 303-318). Plots for both the CR and AR strategies are shown and the results are indicative of the IR strategy whenever $\rho_+ = \rho_- = 0$.

3.2.2 Variance Component of J in the First Order Case

The variance component of **J** is computed as the average variance of the predicted responses within the region of interest, normalized with respect to the number of design points and the experimental error variance. Box and Draper (5) present a mathematical development of the variance error, **V**, for the situation in which the responses are uncorrelated and the region of interest is spherical, and Draper and Lawrence (16) present an extension of this work to cuboidal regions. Hussey, Myers, and Houck (28) expand on the previous work by allowing for correlated responses and develop equations for the variance error under the IR, pure CR, and pure AR correlation induction strategies. The research presented herein expands on their work by considering situations in which there are replicated center runs; that is, the modified CR and AR strategies.

Box and Draper (5) define the variance component of the MSE of response criteria as

$$\begin{aligned} \mathbf{V} &= \frac{N\Omega_r}{\sigma^2} \int_R \mathrm{Var}\,[\,\hat{y}_{(\mathbf{x}_u)}\,]\, d\mathbf{x} \\ &= \frac{N\Omega_r}{\sigma^2} \int_R \mathrm{E}\left\{\hat{y}_{(\mathbf{x}_u)} - \mathrm{E}\,[\hat{y}_{(\mathbf{x}_u)}]\right\}^2 d\mathbf{x} \qquad [3.2.15] \\ &= \frac{N}{\sigma_2}\ \mathrm{Trace}\,\{\,\mathrm{Var}\,[\,\mathbf{b}_1\,]\ \mu_{11}\}\ . \end{aligned}$$

where Ω_r is the inverse of the volume of the centered and scaled region of interest (defined on page 71), $\mathrm{Var}\,[\,\mathbf{b}_1]$ is the variance-covariance matrix of the coefficient estimators (defined on page 72) and μ_{11} is a region moment matrix of the design (defined on page 71).

Equation [3.2.15] indicates that **V** depends on the shape of the region of interest through the μ_{11} term. For the case of fitting a first order model, the μ_{11} matrix is illustrated in Appendix B (page 282), and the elements of the matrix are defined Appendix D (pages 287-291) for both spherical and cuboidal regions of interest. Because the elements of region matrices which involve odd powers of x_i are equal to zero, the μ_{11} matrix for a 2^k design becomes

$$\mu_{11} = \text{Diag}\ (1,\ w_{ii}, w_{ii}, \ldots\ldots, w_{ii}) \qquad [3.2.16]$$

where μ_{11} is a $p_1 \times p_1$ diagonal matrix, with $p_1 = k+1$ in the first order case. For a cuboidal region of interest, $w_{ii} = 1/3$, and for a spherical region of interest, $w_{ii} = 1/(k+2)$.

Equation [3.2.15] also indicates that **V** depends on the variances of the linear coefficients of the first order model. The variances of the β estimators depend on the least squares estimation technique and the correlation induction strategy. The form of $\text{Var}[\underline{b}_1]$ under OLS and WLS is shown in equations [2.2.41] and [2.2.42] on page 72. Both of these variance-covariance matrices depend on the correlation induction strategy through the form of V, the correlation matrix of the sample responses. The form of V for the three correlation induction schemes is shown in equations [2.1.11], [2.1.12], and [2.1.13] on pages 28 and 29.

For the situation in which an experimenter is fitting a first order model using a two-level factorial design, the covariance between any two estimated model coefficients is zero. Therefore, the variance-covariance matrix of $\underline{b}_1$ is a diagonal matrix and can be written as

$$\text{Var}[\underline{b}_1] = \text{Diag}\Big(\text{Var}(b_0), \text{Var}(b_i), \text{Var}(b_i), \ldots\ldots, \text{Var}(b_i)\Big) \qquad [3.2.17]$$

where $\text{Var}[\underline{b}_1]$ is a $p_1 \times p_1$ diagonal matrix. The variance of the b_0 intercept term depends on both the least squares estimation technique and the correlation induction strategy. However, the variances of the b_i $(i = 1, 2, \ldots\ldots, k)$ linear coefficients (which are equivalent for all i) depend on the correlation induction strategy but not on the least squares technique. The variance of the b_i estimators, as shown in Appendix F (pages 297-302), can be written as

$$\mathrm{Var}\,(b_i) = \frac{1-\rho_+}{N\,[ii]}\ \sigma^2 \qquad [3.2.18]$$

where $\rho_+ = 0$ for the **IR** correlation induction strategy. Therefore, the variances of the estimated linear coefficients (with the exception of the intercept term) are the same under the **CR** and **AR** strategies for a given value of ρ_+.

The equation for **V**, the variance component of **J**, can now be defined in terms of the variances of the coefficients for each correlation induction strategy and each least squares estimation technique by incorporating equations [3.2.16] and [3.2.18] into the equation for **V** on page 98. For the situation in which a two-level factorial design is being used to fit a first order model, the variance component of **J** becomes

$$\begin{aligned}
\mathbf{V} &= \frac{N}{\sigma_2}\ \mathrm{Trace}\,\{\mathrm{Var}\,[\,\underline{b}_1\,]\ \mu_{11}\} \\
&= \frac{N}{\sigma_2}\ \{\mathrm{Var}\,(b_0) + w_{ii}\sum_{i=1}^{k}\mathrm{Var}\,(b_i)\} \\
&= \frac{N}{\sigma_2}\ \{\mathrm{Var}\,(b_0) + w_{ii}\,k\,\mathrm{Var}\,(b_i)\} \qquad [3.2.19] \\
&= \frac{N}{\sigma_2}\ \mathrm{Var}\,(b_0) + \frac{w_{ii}\,k\,(1-\rho_+)}{[ii]}
\end{aligned}$$

where $\rho_+ = 0$ for the **IR** strategy, and w_{ii} is the pure second order region moment. The variance of the intercept term, $\mathrm{Var}\,(b_0)$, depends on both the least squares estimation technique and the correlation induction strategy. Using the design point notation shown on page 95, the equations for $\mathrm{Var}\,(b_0)$ are given in Appendix F for the following situations:

1. **IR** strategy (same for OLS and WLS)
2. Modified **CR** strategy under OLS

3. Modified **AR** strategy under OLS
4. Modified **CR** strategy under WLS
5. Modified **AR** strategy under WLS
6. *Pure* **CR** strategy (same for OLS and WLS)
7. *Pure* **AR** strategy (same for OLS and WLS).

Incorporating the variance of the intercept terms from Appendix F (pages 297-302) into equation [3.2.19] leads to the following seven equations for **V**:

$$\mathbf{V}_{IR} = \frac{w_{ii}\,k}{[ii]} + 1 \qquad [3.2.20]$$

$$\mathbf{V}_{CR,OLS} = \frac{w_{ii}\,k\,(1-\rho_+)}{[ii]} - \frac{N_1\rho_+}{N} + \frac{N_1^2\rho_+}{N} + 1 \qquad [3.2.21]$$

$$\mathbf{V}_{AR,OLS} = \frac{w_{ii}\,k\,(1-\rho_+)}{[ii]} - \frac{N_1\rho_+}{N} + \frac{N_1^2(\rho_+ - \rho_-)}{2N} + 1 \qquad [3.2.22]$$

$$\mathbf{V}_{CR,WLS} = \frac{w_{ii}\,k\,(1-\rho_+)}{[ii]} + \frac{N - N\rho_+ + NN_1\rho_+}{N - N_2\rho_+ + N_1N_2\rho_+} \qquad [3.2.23]$$

$$\mathbf{V}_{AR,WLS} = \frac{w_{ii}\,k\,(1-\rho_+)}{[ii]} + \frac{N + N\rho_+ + \tfrac{1}{2}NN_1(\rho_+ - \rho_-)}{N - N_2\rho_+ + \tfrac{1}{2}N_1N_2(\rho_+ - \rho_-)} \qquad [3.2.24]$$

$$\mathbf{V}_{pure\ CR} = \frac{w_{ii}\,k\,(1-\rho_+)}{[ii]} - \rho_+ + N\rho_+ + 1 \qquad [3.2.25]$$

$$\mathbf{V}_{pure\ AR} = \frac{w_{ii}\,k\,(1-\rho_+)}{[ii]} - \rho_+ + \tfrac{1}{2}N(\rho_+ - \rho_-) + 1\ . \qquad [3.2.26]$$

The equations for **V** can now be used to find the value of [ii] which results in a Min-**V** design. The optimal value of [ii] is determined by setting the partial derivative of **V** with respect to [ii] equal to zero. Inspection of the equations for **V** indicate that only the variances of the b_i terms involve [ii].

Therefore, regardless of the correlation induction strategy and least squares estimation technique, setting the partial derivative of **V** with respect to [ii] equal to zero yields

$$\frac{\partial \mathbf{V}}{\partial\,[\mathrm{ii}]} = -\frac{w_{ii}\,k\,(1-\rho_+)}{[\mathrm{ii}]^2} = 0\ . \qquad [3.2.27]$$

The second derivative of **V** with respect to [ii], $\frac{\partial^2 \mathbf{V}}{\partial\,[\mathrm{ii}]^2} = \frac{w_{ii}\,k\,(1-\rho_+)}{[\mathrm{ii}]^3}$, is always a positive quantity, and therefore the solution of equation [3.2.27] provides the value of [ii] which minimizes **V**. The **V**-optimal value of the pure second order design moment for a 2^k design requires that [ii] be as large as possible. For a centered and scaled design region, the largest value of [ii] is generally considered to be [ii] = 1. If the levels of the F factorial points are set at $\pm g$, and the design is augmented with N_c center runs, then the value of g which results in a **V**-optimal design is

$$\begin{aligned} g &= \sqrt{\frac{[\mathrm{ii}]\,(F+N_c)}{F}} \\ &= \sqrt{1+\frac{N_c}{F}}\ . \end{aligned} \qquad [3.2.28]$$

Thus, for a **V**-optimal two-level factorial design, the optimal value of [ii] is independent of both the correlation induction strategy and the least squares estimation technique. In the next section, the MSE of response criteria, which calls for minimizing the sum of **B** and **V**, is used to determine the **J**-optimal values of [ii] for each of the correlation induction strategies.

3.2.3 J - Optimum First Order Designs

A design which results in a minimum value for the MSE of response is considered to be a **J**-optimum design. Box and Draper (5) define **J** as the average mean squared error of the predicted responses within the region of interest, normalized with respect to the number of design points and the experimental error variance, and these authors compute **J** as the sum of the bias error, **B**, and the variance error, **V**. For the situation in which a two-level factorial design is being used to fit a first order model, the bias and variance components of **J** have been defined in sections 3.2.1 and 3.2.2. Utilizing equation [3.2.5] on page 93 for **B**, and equation [3.2.19] on page 100 for **V**, the equation for **J** becomes

$$\begin{aligned} \mathbf{J} &= \mathbf{B} + \mathbf{V} \\ &= \theta\left[\phi\,(z\,[ii] - w_{ii})^2 + 2\,y_r\,w_{ii}\right] + \frac{N}{\sigma_2}\,\mathrm{Var}\,(b_0) + \frac{w_{ii}\,k\,(1-\rho_+)}{[ii]} \end{aligned} \qquad [3.2.29]$$

where θ and ϕ are parameters involving the unfitted coefficients in the second order model, z is a scalar term which is equal to one under OLS estimation and a function of the correlation induction strategy in the case of WLS, [ii] is the pure second order design moment whose optimal value is to be solved for, w_{ii} is the pure second order region moment which is equal to $1/3$ for a cuboidal region and equal to $1/(k+2)$ for a spherical region, y_r is a region-dependent parameter which is a function of k and ϕ, $\mathrm{Var}\,(b_0)$ is the variance of the intercept term which depends on both the correlation induction strategy and the least squares estimation technique, and ρ_+ is the magnitude of the induced positive correlation ($\rho_+ = 0$ for the IR strategy).

The equation for **J** can now be used to determine the value of [ii] which results in a Min-**J** design. The optimal value of [ii] is found by setting the partial derivative of **J** with respect to [ii] equal to zero, yielding

$$\frac{\partial\,\mathbf{J}}{\partial\,[ii]} = 2\,\phi\theta\,z\left(z\,[ii] - w_{ii}\right) - \frac{w_{ii}\,k\,(1-\rho_+)}{[ii]^2} = 0\,. \qquad [3.2.30]$$

The second partial derivative of **J** with respect to [ii], $\dfrac{\partial^2\,\mathbf{J}}{\partial\,[ii]^2} = 2\,\phi\theta\,z^2 + \dfrac{w_{ii}\,k\,(1-\rho_+)}{[ii]^3}$, is always a positive quantity and, therefore, the solution of equation [3.2.30] provides the value of [ii] which minimizes **J**.

Because the derivative of **J** involves the unknown bias parameters, ϕ and θ, the optimal value of [ii] is a function of these quantities. There are two ways in which to view the problem at hand. The first way is to view the optimal value of [ii] as a function of the quantity $\phi\theta$. This method is relatively easy to do, especially in the case of ordinary least squares, because one can plot $\phi\theta$ versus the **J**-optimal value of [ii] for various values of k and ρ_+. However, the problem with this method is that it is difficult for experimenters to estimate an intangible quantity such as $\phi\theta$. Recall that $\phi\theta$ is a standardized sum of the unfitted quadratic coefficients in the protection model (shown in equation [3.2.4] on page 93) with a range of $0 \leq \phi\theta \leq \infty$.

Due to difficulties associated with the estimation of $\phi\theta$, an additional method for determining the optimal value of [ii] is used in this research. Rather than estimating $\phi\theta$, this second method requires that the experimenter estimate the relative amounts of bias error and variance error. For example, if an experimenter believes that the predicted responses consist of half as much bias error as variance error, then the experimenter would use the optimal value of [ii] for a **V**-to-**B** ratio of 2. However, in order to determine the optimal value of [ii] for a given **V**-to-**B** ratio, one must also specify a value of the ϕ parameter. The range of possible values for ϕ is limited to $0 \leq \phi \leq k$, which is a significantly smaller range than the infinite range of values for the $\phi\theta$ parameter. Recall that ϕ is parameter which indicates the state of conditioning of the response surface (shown in equation [3.2.3] on page 92). The best state of conditioning corresponds to $\phi = k$, and the worst state of conditioning corresponds to $\phi = 0$. Most response surfaces are likely to fall somewhere between these two extremes. In this research, the value of ϕ that is generally used is k/2; that is, midway between the best and worst states of conditioning. This second method of presenting the **J**-optimal values of [ii] is through plots of **V**-to-**B** versus [ii] for a given value of ϕ, and for various values of k and ρ_+.

The next two sections present a comparison of the **J**-optimal values of [ii] for the three correlation induction strategies. The following two methods are used:

1. Determine the **J**-optimal values of [ii] as a function of the $\phi\theta$ bias parameter.
2. Determine the **J**-optimal values of [ii] as a function of the **V**-to-**B** ratio.

3.2.3.1 J-Optimum Designs as a Function of Bias Parameters

The first method of determining the **J**-optimal value of [ii] is to view $\phi\theta$ as a parameter which must be specified by the experimenter. In order to determine the optimal value of [ii], the partial derivative of **J** with respect to [ii] must be set equal to zero (as shown in equation [3.2.30] on page 103) and solved for [ii]. To facilitate solution of equation [3.2.30], it is convenient to write the equation in the form of a cubic polynomial in [ii], yielding

$$0 = [\mathrm{ii}]^3 - \frac{w_{ii}}{z}[\mathrm{ii}]^2 - \frac{w_{ii}\,k\,(1-\rho_+)}{2\,\phi\,\theta\,z^2}[\mathrm{ii}]^0 . \qquad [3.2.31]$$

Given a specific value of $\phi\theta$, as well as the appropriate values of w_{ii}, k, ρ_+, and z, the optimal value of [ii] can be determined by solving for the one real root of equation [3.2.31]. The **J**-optimal value of the pure second order design moment for a 2^k design becomes

$$[\mathrm{ii}] = \frac{w_{ii}}{3z} + p_1 + q_1 \qquad [3.2.32]$$

where p_1 and q_1 are defined as

$$p_1 = +\sqrt[3]{-\frac{t_1}{2} + \sqrt{\frac{t_1^{\,2}}{4} + \frac{s_1^{\,3}}{27}}}$$

$$q = -\sqrt[3]{+\frac{t_1}{2} + \sqrt{\frac{t_1^{\,2}}{4} + \frac{s_1^{\,3}}{27}}}$$

and s_1 and t_1 are computed as

$$s_1 = -\frac{w_{ii}^2}{3z^2}$$

$$t_1 = -\frac{4\phi\theta w_{ii}^3 - 27 k w_{ii} z(1-\rho_+)}{54\phi\theta z^3} \quad .$$

Inserting the appropriate values of s_1 and t_1 into the equations for p_1 and q_1, and inserting the values of p_1 and q_1 into equation [3.2.32], yields the J-optimal value of [ii], given $\phi\theta$, w_{ii}, k, ρ_+, and z.

The only region-dependent term in equation [3.2.32] is w_{ii}. For a cuboidal region $w_{ii} = 1/3$, and $w_{ii} = 1/(k+2)$ for a spherical region. The only term in equation [3.2.32] which is a function of the least squares estimation technique is z. Under OLS, $z = 1$, but under WLS, $z = 1$ only for the IR strategy. Equations [3.2.11] and [3.2.12] on page 96 give the formulas for the WLS values of z under the CR and AR strategies. The only term in equation [3.2.32] which is solely a function of the correlation induction strategy is ρ_+. For the IR strategy, $\rho_+ = 0$, and for the CR and AR strategies, $0 < \rho_+ < 1$. Because equation [3.2.32] does not involve ρ_-, the optimal value of [ii] is the same for the CR and AR strategies under OLS estimation. Under WLS, however, the optimal values of [ii] are different for the CR and AR strategies due to the differences in the z term.

The J-optimal values of [ii] can best be seen by viewing plots of $\phi\theta$ versus [ii] for a given region shape and least squares analysis. In the case of OLS, the optimal values of [ii] depend only on the values of ρ_+ and k. Therefore, when $\rho_+ = 0$, the optimal values of [ii] correspond to the IR strategy, and when $\rho_+ > 0$, the optimal values of [ii] correspond to both the CR and AR strategies.

J-Optimal* [ii] *versus* $\phi\theta$ *under Ordinary Least Squares

Figure 1 on page 110 illustrates the relationship between the J-optimal value of [ii] and the $\phi\theta$

parameter for a spherical region of interest, and Figure H-1 on page 320 illustrates the same situation for a cuboidal region. Both of these figures correspond to an OLS analysis and are independent of the number of center runs. Figure 1 indicates that the **J**-optimal values of [ii] approach the **B**-optimal values (the horizontal lines) as ρ_+ and $\phi\theta$ increase. Only when $\phi\theta$ is close to zero (little bias error exists) do the **J**-optimal values of [ii] approach the **V**-optimal value of [ii] = 1. The results for a cuboidal region shown in Figure H-1 are similar to those for a spherical region, but the **B**- and **J**-optimal values of [ii] are larger for a cuboidal region due to differences in the values of w_{ii} for the two regions. Both of these figures indicate that the size of the **J**-optimal designs are smaller when correlation is induced ($\rho_+ > 0$) than when no correlation is induced ($\rho_+ = 0$). Thus, when the design criteria is minimization of the MSE of response, it is even more important to reduce the size of the design when correlation is induced among the responses.

J-Optimal [ii] *versus* $\phi\theta$ *under Weighted Least Squares*

For a WLS analysis, equation [3.2.32] on page 105 indicates that **J**-optimal values of [ii] depend on the number of design points utilizing independent, common, and antithetic random number streams through the z term. For the IR strategy, the results under OLS and WLS results are equivalent because $z_{IR} = 1$, but for the CR and AR strategies, the results under WLS depend on the number of center runs (N_c) used in the 2^k designs. Because the **J**-optimal values of [ii] vary with N_c and k in the same way as the **B**-optimal values of [ii] (shown in Appendix G), only a design with $k = 3$ factors and $N_c = 5$ center runs is used to illustrate the results for a WLS analysis. The figures in Appendix G (pages 303-318) indicate that the optimal values of [ii] increase with k for a cuboidal region, but decrease with k for a spherical region. Also, the optimal values of [ii] increase with N_c when $\rho_+ > \rho_- > 0$, do not change with N_c when $\rho_+ = \rho_- = 0$, and decrease with N_c when $\rho_+ = \rho_- > 0$. These results hold for both the **B**- and **J**-optimal values of [ii].

The J-optimal values of [ii] for a WLS analysis are shown in Figures 2 and 3 for the CR and AR strategies, respectively. These figures, which are specific to a 2^k design with $k=3$ factors and $N_c = 5$ center runs, are discussed in the next two sections.

J-*Optimal* [ii] *versus* $\phi\theta$ *for the CR Strategy under WLS*

Figure 2 on page 111 indicates that when the CR strategy is used in a spherical region under WLS estimation, the J-optimal values of [ii] approach the B-optimal values as ρ_+ and $\phi\theta$ increase. Unlike the OLS results, the B-optimal values of [ii] increase with ρ_+, and therefore the J-optimal values increase with ρ_+ also. Figure 2 clearly indicates that the J-optimal values of [ii] closely follow, but are slightly larger than the B-optimal values, and as ρ_+ increases, the differences between the B- and J-optimal values of [ii] become smaller. Only when $\phi\theta$ is near zero (little or no bias error exists) do the J-optimal values approach the V-optimal value of [ii] = 1.

The results for a cuboidal region of interest, as shown in Figure H-2 on page 321, are similar those for a spherical region of interest. Because the optimal values of [ii] are a function of $1/(k+2)$ for a spherical region and $1/3$ for a cuboidal region, the optimal values of [ii] are larger for a cuboidal region. The results for the cuboidal region indicate that the J-optimal values of [ii] for the CR strategy are very close to the B-optimal values, except when $\phi\theta$ is near zero.

J-*Optimal* [ii] *versus* $\phi\theta$ *for the AR Strategy under WLS*

Figure 3 on page 112 illustrates the relationship between ρ_+, ρ_-, and the B- and J-optimal values of [ii] when the AR strategy is used in a spherical region under WLS estimation. The figure indicates that the J-optimal values of [ii] approach the B-optimal values as both ρ_+ and ρ_- increase. Small values of ρ_+ and ρ_- result in larger differences between the J- and B-optimal values of [ii].

However, even for small values of ρ_+ and ρ_-, Figure 3 indicates that the J-optimal values of [ii] closely follow the B-optimal values and are well below the V-optimal value of [ii] = 1.

The results for WLS (Figures 2 and 3) differ from the OLS results (Figure 1) because the B- and J-optimal designs for the CR and AR strategies under WLS are larger (except when $\rho_- = \rho_+$) than the optimal designs for the IR strategy. Under OLS, the J-optimal designs for the CR and AR strategies are smaller than the optimal designs for the IR strategy. However, both the OLS and WLS results indicate that the Min-J values of [ii] are slightly larger than those which minimize B alone, and tend to be much smaller than those which minimize V alone.

The *line* convention used to denote the values of ρ_+ and ρ_- in each of the figures presented in this Chapter is as follows:

	Type of Line	CR Strategy	AR Strategy
1.	————————	$\rho_+ = 0$	$\rho_+ - \rho_- = 0$
2.	----------------------	$\rho_+ = .3$	$\rho_+ - \rho_- = .3$
3.	— — — — — —	$\rho_+ = .6$	$\rho_+ - \rho_- = .6$
4.	—— — —— —	$\rho_+ = .9$	$\rho_+ - \rho_- = .9$.

The next section compares the B-, V-, and J-optimal values of [ii] for the three correlation induction strategies, but instead of viewing the bias error in terms of $\phi\theta$, the bias error is viewed relative to the variance error; that is, as a function of the ratio of V-to-B. This method is presented because the relative amount of variance to bias error is a more tangible quantity than the $\phi\theta$ parameter.

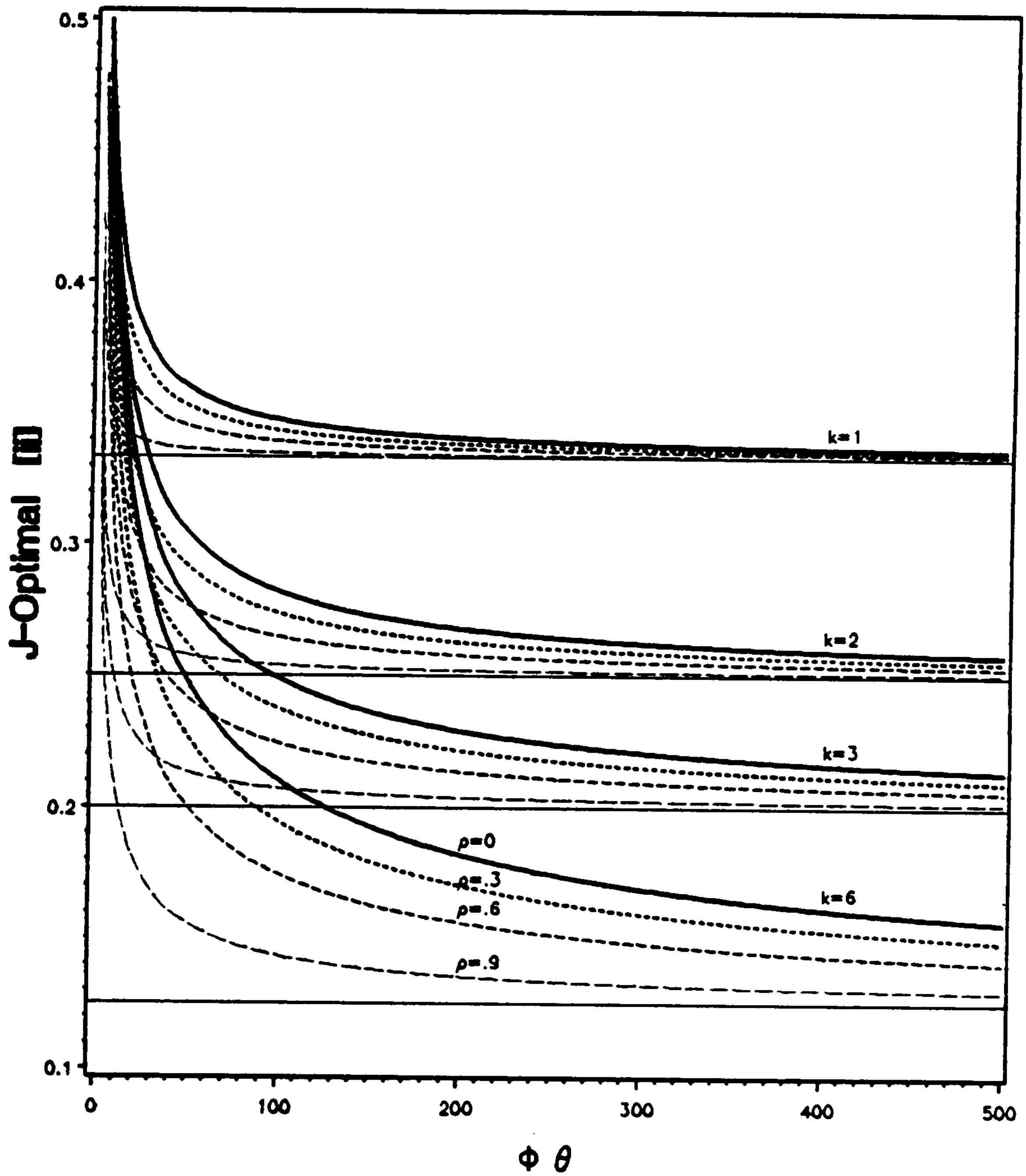


Figure 1. **First Order J-optimal [ii] versus $\phi\theta$ under OLS.**

Region of interest is Spherical.

Values of k = 1, 2, 3, 6 and values of ρ_* = 0, .3, .6, .9.

The V-optimal value of the second moment is [ii] = 1.0.

The horizontal lines at [ii] = 1/(k + 2) indicate the B-optimal values.

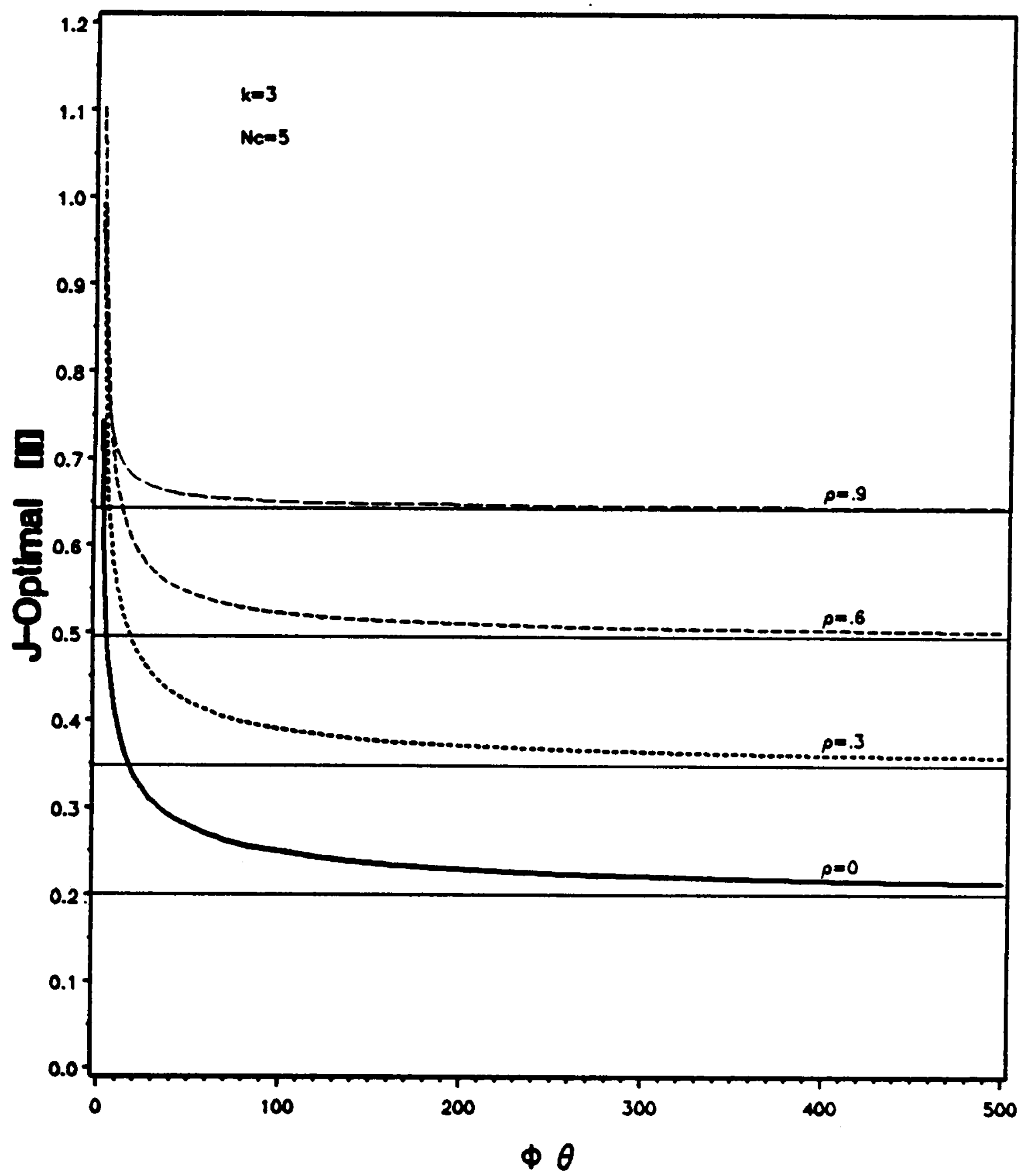


Figure 2. **J-optimal [ii] versus $\phi\theta$ for the Modified CR Strategy under WLS.**
Region of interest is Spherical.
Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.
Values of $\rho_* = 0, .3, .6, .9$.
The V-optimal value of the second moment is $[ii] = 1.0$.
The horizontal lines indicate the B-optimal values of [ii].

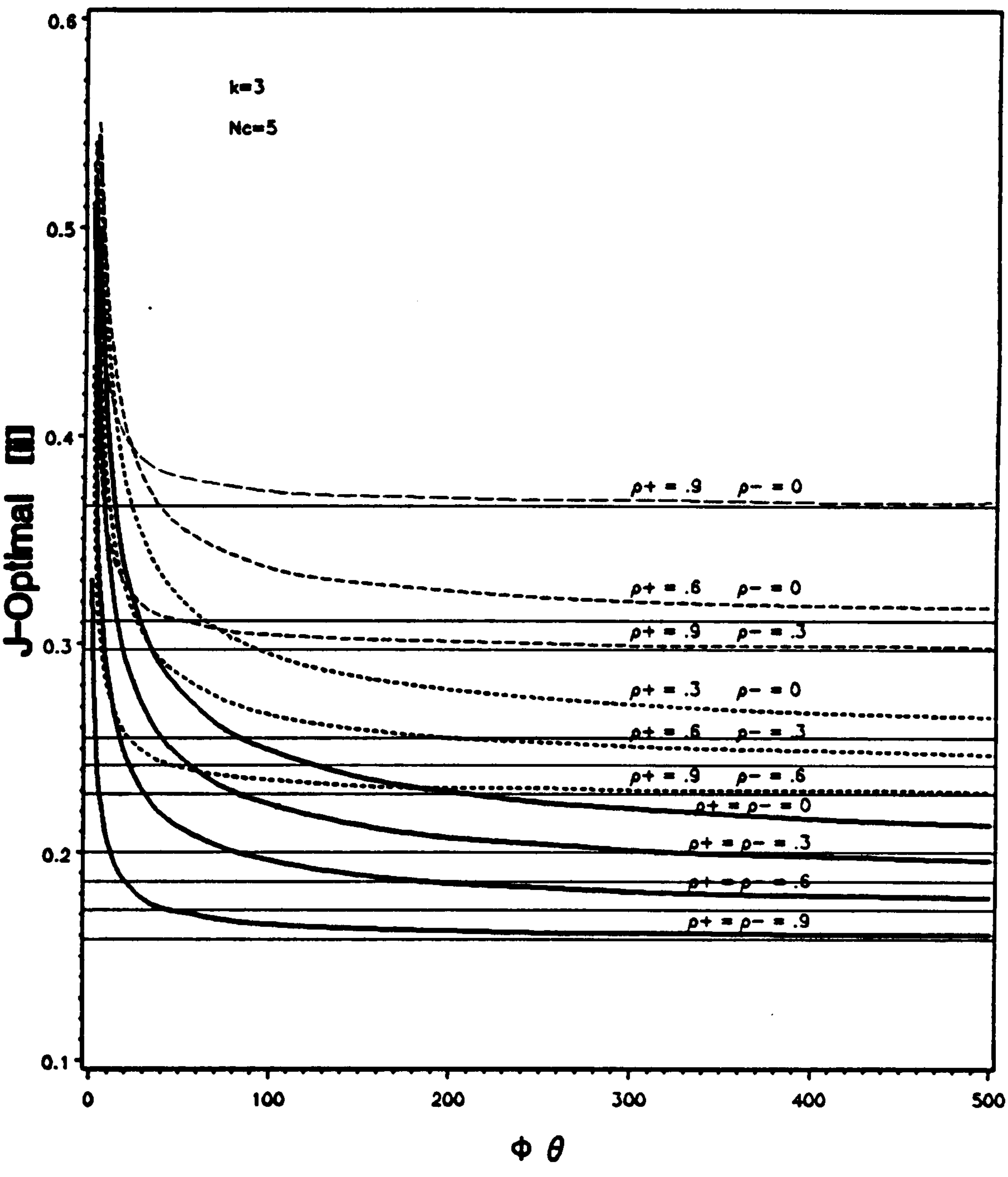


Figure 3. **J-optimal [ii] versus $\phi\theta$ for the Modified AR Strategy under WLS.**

Region of interest is Spherical.

Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.

Values of $\rho_- \leq \rho_+ = 0, .3, .6, .9$.

The V-optimal value of the second moment is $[ii] = 1.0$.

The horizontal lines indicate the B-optimal values of [ii].

3.2.3.2 J-Optimum Designs as a Function of Variance-to-Bias Ratios

The second method of determining the **J**-optimal value of [ii] is to consider the ratio of **V**-to-**B** as a quantity which must be specified by the experimenter. The optimal value of [ii] as a function of **V/B** can be obtained by solving the ratio of **V**-to-**B** for the θ parameter, inserting the equation for θ into the partial derivative of **J** with respect to [ii], setting the partial derivative equal to zero, and solving for [ii]. The main disadvantage of this method is that the **J**-optimal value of [ii] is a function of the ϕ parameter, as well as the **V**-to-**B** ratio. However, the range of possibie values for the ϕ parameter is limited to $0 \leq \phi \leq k$, with $\phi = 0$ representing a response surface in the worst state of conditioning, and $\phi = k$ representing the best state of conditioning. A comparison of the **J**-optimal values of [ii] over the range of ϕ values is presented in Figure 4, but the remainder of the figures in this Chapter utilize $\phi = \mathrm{k}/2$; that is, midway between the worst and best states of conditioning.

The **V**-to-**B** ratio can be defined using equation [3.2.5] on page 93 for the bias error, and using equation [3.2.19] on page 100 for the variance error, yielding

$$\mathbf{V/B} = \frac{\left[\frac{\mathrm{N}}{\sigma_2} \operatorname{Var}(\mathrm{b_0})\right] + \frac{\mathrm{w_{ii}\,k\,(1-\rho_+)}}{[\mathrm{ii}]}}{\theta\left[\phi\,(\mathrm{z\,[ii]} - \mathrm{w_{ii}})^2 + 2\,\mathrm{y_r\,w_{ii}}\right]} \quad . \qquad [3.2.33]$$

Letting $\mathrm{v_{b_0}}$ denote $\left[\frac{\mathrm{N}}{\sigma^2} \operatorname{Var}(\mathrm{b_0})\right]$, and solving equation [3.2.33] for the θ parameter, yields

$$\theta = \frac{\mathrm{v_{b_0}} + \mathrm{w_{ii}\,k\,(1-\rho_+)\,/\,[ii]}}{\mathbf{V/B}\left[\phi\,(\mathrm{z\,[ii]} - \mathrm{w_{ii}})^2 + 2\,\mathrm{y_r\,w_{ii}}\right]} \; . \qquad [3.2.34]$$

In order to minimize **J** with respect to [ii], the partial derivative of **J** with respect to [ii] must be set equal to zero (as shown in equation [3.2.31] on page 105). Inserting the equation for θ into equation [3.2.31] results in the following cubic polynomial in [ii]:

$$0 = [ii]^3 + \left[\frac{w_{ii}\, k\, (1-\rho_+)\, (2 - V/B)}{2\, v_{bo}} - \frac{w_{ii}}{z} \right] [ii]^2$$

$$+ \left[\frac{w_{ii}^2\, k\, (1-\rho_+)\, (V/B - 1)}{z\, v_{bo}} \right] [ii]^1 \qquad [3.2.35]$$

$$- \left[\frac{w_{ii}^2\, k\, (1-\rho_+)\, (w_{ii}\, \phi + 2\, y_r)\, V/B}{2\, \phi\, z^2\, v_{bo}} \right] [ii]^0$$

$$= [ii]^3 + a\,[ii]^2 + b\,[ii]^1 + c\,[ii]^0$$

where a , b , and c (defined for convenience) are the coefficients of $[ii]^2$, $[ii]^1$, and $[ii]^0$, respectively. Given specific values of V/B and ϕ, as well as the appropriate values of w_{ii}, y_r, k, ρ_+, v_{bo}, and z, the J-optimal value of [ii] can be determined by solving for the positive real root of equation [3.2.35], which is one of the following three roots:

$$\#\,1. \quad [ii] = -\frac{a}{3} + p_2 + q_2$$

$$\#\,2. \quad [ii] = -\frac{a}{3} - \frac{p_2 + q_2}{2} + \frac{p_2 - q_2}{2} \sqrt{-3}$$

$$\#\,3. \quad [ii] = -\frac{a}{3} - \frac{p_2 + q_2}{2} - \frac{p_2 - q_2}{2} \sqrt{-3}$$

where p_2 and q_2 are defined as

$$p_2 = + \sqrt[3]{- \frac{t_2}{2} + \sqrt{\frac{t_2^{\,2}}{4} + \frac{s_2^{\,3}}{27}}}$$

$$q_2 = - \sqrt[3]{+ \frac{t_2}{2} + \sqrt{\frac{t_2^{\,2}}{4} + \frac{s_2^{\,3}}{27}}}$$

and s_2 and t_2 are computed as

$$s_2 = b - \frac{a^2}{3}$$

$$t_2 = \frac{2a^3}{27} - \frac{ab}{3} + c \ .$$

If $\frac{s_2^{\,3}}{27} + \frac{t_2^{\,2}}{4} > 0$, then root #1 is the optimal value of [ii], or

if $\frac{s_2^{\,3}}{27} + \frac{t_2^{\,2}}{4} \leq 0$, then the positive real root (#1, #2, or #3) is the optimal value of [ii].

Inserting the appropriate values of a , b , and c into the equations for s_2 and t_2 , and inserting the values for s_2 and t_2 into the equations for p_2 and q_2 , and inserting the values of a, p_2, and q_2 into the three roots (#1, #2, and #3) yields one positive real root which is equal to the J-optimal value of [ii], given **V/B**, ϕ, w_{II} , y_r , k, ρ_+ , z, and v_{b_0} . Under OLS, $z = 1$, and therefore only the value of v_{b_0} depends on the number of design points using independent, common, and antithetic random number streams. Under WLS, both z and v_{b_0} depend on the values of N, N_1, and N_2. Recall the design point notation shown on page 95, in which $N = N_1 + N_2$, where N is the total number of design points in the full or fractional 2^k design, N_1 is the number of design points using common or antithetic random number streams, and N_2 is the number of design points using independent streams. Utilizing this design point notation, the variance of the intercept term, Var (b_0), is defined in Appendix F (pages 297-302) for the following situations:

1. **IR** strategy (same for OLS and WLS)
2. Modified **CR** strategy under OLS
3. Modified **AR** strategy under OLS
4. Modified **CR** strategy under WLS
5. Modified **AR** strategy under WLS
6. *Pure* **CR** strategy (same for OLS and WLS)
7. *Pure* **AR** strategy (same for OLS and WLS).

Multiplying the $\mathrm{Var}(b_0)$ terms shown in Appendix F by N/σ^2, leads to the following seven equations for v_{b_0}, the standardized variance of the intercept:

$$v_{b_0}|_{IR} = 1 \qquad [3.2.36]$$

$$v_{b_0}|_{CR,OLS} = 1 - \frac{N_1\rho_+}{N} + \frac{N_1^{\,2}\rho_+}{N} \qquad [3.2.37]$$

$$v_{b_0}|_{AR,OLS} = 1 - \frac{N_1\rho_+}{N} + \frac{N_1^{\,2}(\rho_+ - \rho_-)}{2N} \qquad [3.2.38]$$

$$v_{b_0}|_{CR,WLS} = \frac{N - N\rho_+ + NN_1\rho_+}{N - N_2\rho_+ + N_1N_2\rho_+} \qquad [3.2.39]$$

$$v_{b_0}|_{AR,WLS} = \frac{N - N\rho_+ + \tfrac{1}{2}NN_1(\rho_+ - \rho_-)}{N - N_2\rho_+ + \tfrac{1}{2}N_1N_2(\rho_+ - \rho_-)} \qquad [3.2.40]$$

$$v_{b_0}|_{pure\ CR} = 1 - \rho_+ + N\rho_+ \qquad [3.2.41]$$

$$v_{b_0}|_{pure\ AR} = 1 - \rho_+ + \tfrac{1}{2}N(\rho_+ - \rho_-)\,. \qquad [3.2.42]$$

The equations for v_{b_0} can now be inserted into equation [3.2.35] on page 114 to find the values of [ii] which result in Min-J designs for each of the seven situations. Figures 4 through 10 at the end of this section illustrate the J-optimal values of [ii] for the two-level factorial designs examined in this research.

The relationship between the ϕ parameter and the J-optimal value of [ii] is shown in Figure 4 on page 120 for a spherical region of interest. Figure H-4 on page 323 illustrates the relationship for a cuboidal region. Both figures indicate that the optimal values of [ii] increase as ϕ decreases and, therefore, J-optimal designs for poorly-conditioned response surfaces (redundant input variables) are larger than those for well-conditioned response surfaces (orthogonal input variables). A well-conditioned response surface is depicted by $\phi = k - .5$ in Figure 4, and a poorly-conditioned response surface is depicted by $\phi = .5$. Figure 4 indicates that the largest difference in the J-optimal values of [ii] occurs at **V/B** ratios in the range of 4 to 5½, and the smallest differences occur at **V/B** ratios larger than 6. The results for a cuboidal region shown in Figure H-4 are similar to those for a spherical region. Throughout the remainder of this Chapter, analyses based on the **V/B** ratio will utilize $\phi = k/2$; that is, midway between the best and worst states of conditioning.

In order to compare the J-optimal values of [ii] for the three correlation induction strategies, one must consider whether the CR and AR strategies are *pure*, and if not, whether the least squares estimation technique is ordinary or weighted. Before examining the *modified* CR and AR strategies, the *pure* AR and CR strategies are compared to the IR technique. Recall that the pure strategies do not involve the use of independent random number streams and the optimal designs are the same under OLS and WLS estimation. Figures 5 and 6 illustrate the relationship between the J-optimal [ii] and the **V/B** ratio for 2^k designs without center runs. Figure 5 corresponds to the pure CR strategy and Figure 6 corresponds to the pure AR strategy, and whenever $\rho_+ = \rho_- = 0$, the results correspond to the IR strategy.

J-*Optimal* [ii] *versus* V/B *for the Pure CR and AR Strategies*

Figure 5 on page 121 indicates that the J-optimal designs for the pure CR strategy (when $\rho_+ > 0$) are smaller than the optimal designs for the IR strategy (when $\rho_+ = 0$), and the J-optimal values of [ii] approach the **B**-optimal value as ρ_+ increases. Even for large **V/B** ratios, the J-optimal values of [ii] are well below the **V**-optimal value of [ii] = 1. Therefore, under the pure CR strategy, the optimal [ii] for a Min-J design is slightly larger than for a Min-**B** design, and the amount larger decreases as the magnitude of ρ_+ increases.

Figure 6 on page 122 illustrates the relationship between the J-optimal [ii] and the V/B ratio for the pure AR strategy. Whenever $\rho_+ = \rho_-$, the J-optimal values of [ii] are the same as those for the IR technique. However, the larger the value of ρ_+ and the larger the difference between ρ_+ and ρ_-, the closer the J-optimal values of [ii] are to the B-optimal value. Again, similar to the results for the pure CR strategy shown in Figure 5, the optimal [ii] for a Min-J design under the pure AR strategy is slightly larger than the B-optimal value, and the amount larger decreases as the magnitudes of ρ_+ and ρ_- increase.

J-Optimal* [ii] *versus* V/B *for the Modified CR and AR Strategies under OLS

In order to compare the J-optimal values of [ii] under the *modified* CR and AR strategies, the least squares estimation technique needs to be specified. Under OLS estimation, the J-optimal values of [ii] are a function of N_c, ρ_+, and ρ_- through v_{b0}, the standardized variance of the intercept, and under WLS, the J-optimal values of [ii] additionally depend on the z term (as shown in equations [3.2.11] and [3.2.12] on page 96).

The relationship between the J-optimal [ii] and the V/B ratio for the modified CR and AR strategies is illustrated for a 2^k design with $k = 3$ factors, augmented with $N_c = 5$ center runs. Under OLS the B-optimal values of [ii] are not a function of ρ_+ and ρ_-, but under WLS the B-optimal values of [ii] depend on the amount of correlation induced. Figure 7 on page 123 illustrates the J-optimal [ii] versus V/B for the CR strategy under OLS, and Figure 9 on page 125 illustrates the same situation for the AR strategy. By comparing these plots for the modified strategies to those for the pure CR and AR strategies (Figures 5 and 6), it appears that the J-optimal values of [ii] are slightly larger for the modified strategies (except when $\rho_+ = \rho_-$).

J-*Optimal* [ii] *versus* V/B *for the Modified CR and AR Strategies under WLS*

The results under WLS are shown in Figures 8 and 10 on pages 124 and 126 for the modified CR and AR strategies, respectively. The horizontal lines correspond to the **B**-optimal values of [ii], which increase as ρ_+ increases. For the AR strategy, the **B**-optimal values of [ii] also increase as the difference between ρ_+ and ρ_- increases. Both of the figures indicate that the **J**-optimal values of [ii] approach the **B**-optimal values of [ii] as ρ_+ increases. Figure 10, depicting the AR strategy under WLS, illustrates that the **J**-optimal [ii] are well below the **V**-optimal value of [ii] = 1.

The results of this section, concerning the relationship between the **J**-optimal [ii] and the **V/B** ratio are shown in Figures 4 through 10 on pages 120-126. The figures indicate that the **J**-optimal design sizes are closer to the **B**-optimal design sizes when correlation is induced than when no correlation is induced. The **J**-optimal values of [ii] approach the **B**-optimal values as ρ_+ and the difference between ρ_+ and ρ_- increase. When no correlation is induced, the **J**-optimal values of [ii] increase rapidly when **V/B** > 4, but when correlation is induced, the **J**-optimal values of [ii] increase slowly with the **V/B** ratio and never reach the **V**-optimal value of [ii] = 1. Therefore, it appears that when correlation is induced, the size of **J**-optimal designs are even closer to the size of **B**-optimal designs, making it even more important to bring the design points closer to the center of the region and away from the **V**-optimal ± 1 extremes.

The next section of this paper deals with a comparison of the correlation induction strategies on the basis of the values of **J**, the MSE of response. The values of **J** are illustrated as a function of the **V/B** ratio for Min-**B** designs, Min-**V** designs, and Min-**J** designs.

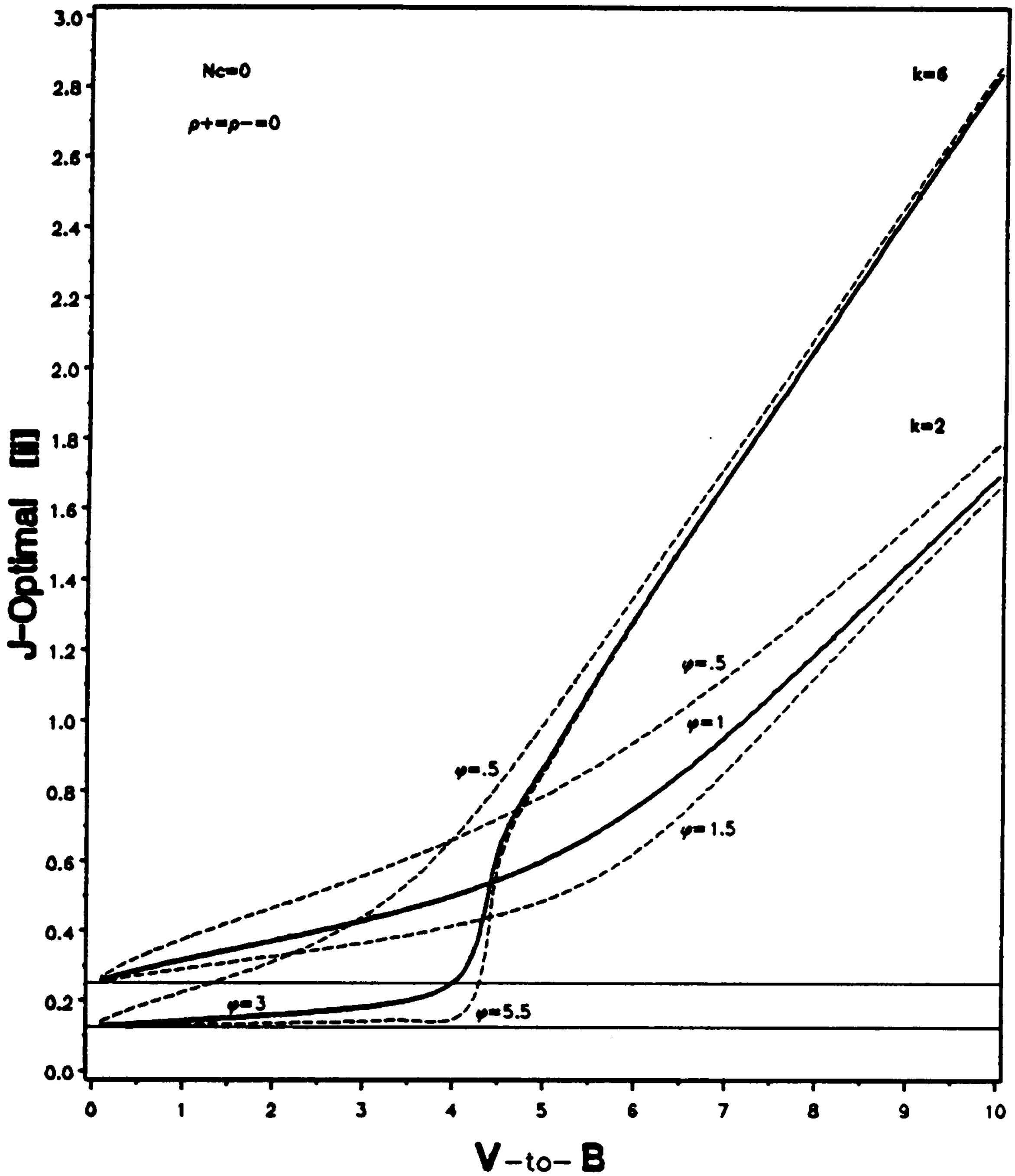


Figure 4. **J-optimal [ii] versus V/B for various ϕ.**

IR strategy in a Spherical region.
Two-level factorial design with $k = 2$ and $k = 6$ factors.
Values of ϕ = .5, k, k - .5 .
The V-optimal value of the second moment is $[ii] = 1.0$.
The horizontal lines at $[ii] = 1/(k+2)$ indicate the B-optimal values.

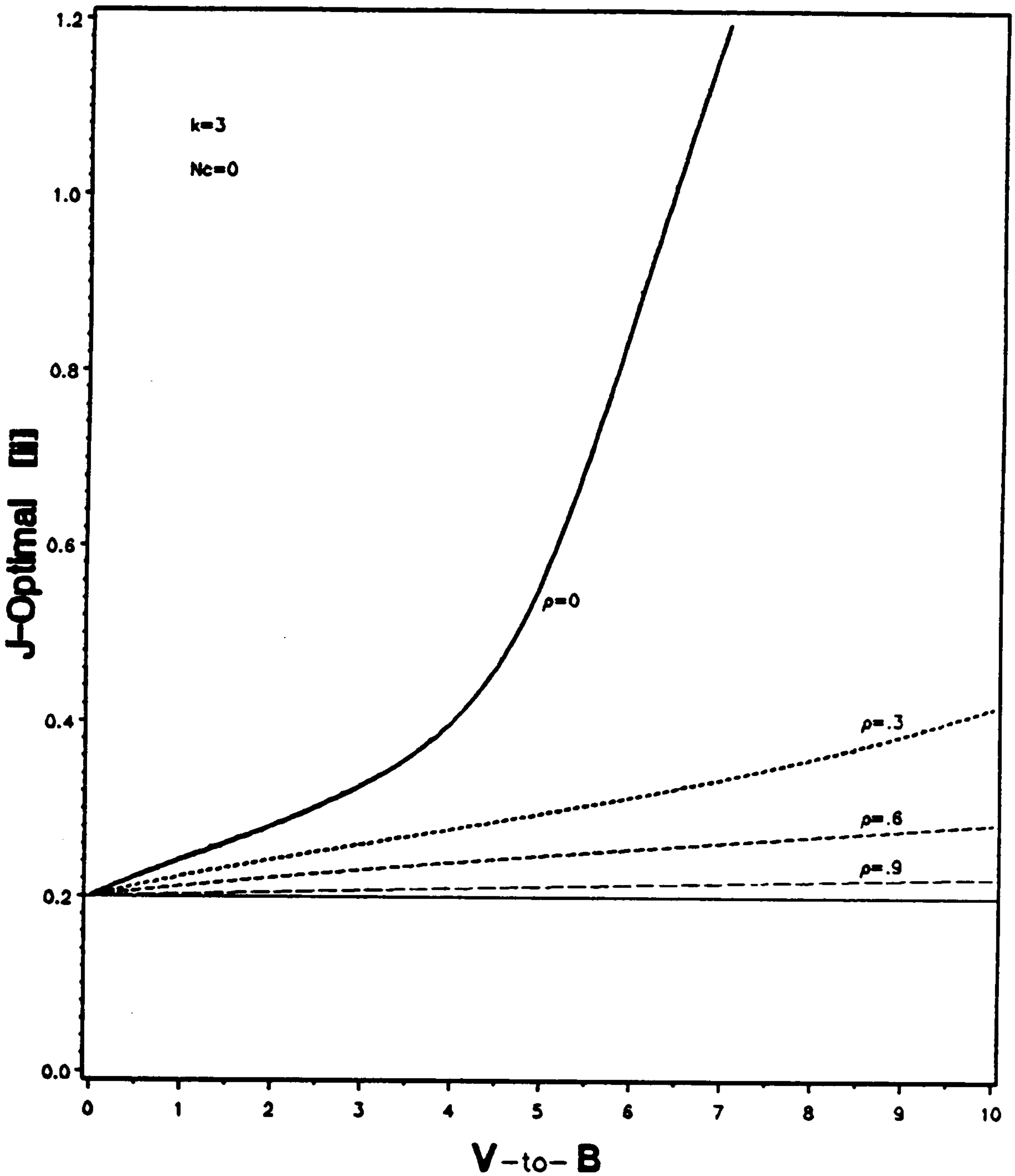


Figure 5. J-optimal [ii] versus V/B for the Pure CR Strategy.

Region of interest is Spherical.

Two-level factorial design with $k = 3$ factors and $N_c = 0$ center runs.

Values of $\rho_* = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The V-optimal value of the second moment is $[ii] = 1.0$.

The horizontal line at $[ii] = 1/(k+2)$ indicates the B-optimal value.

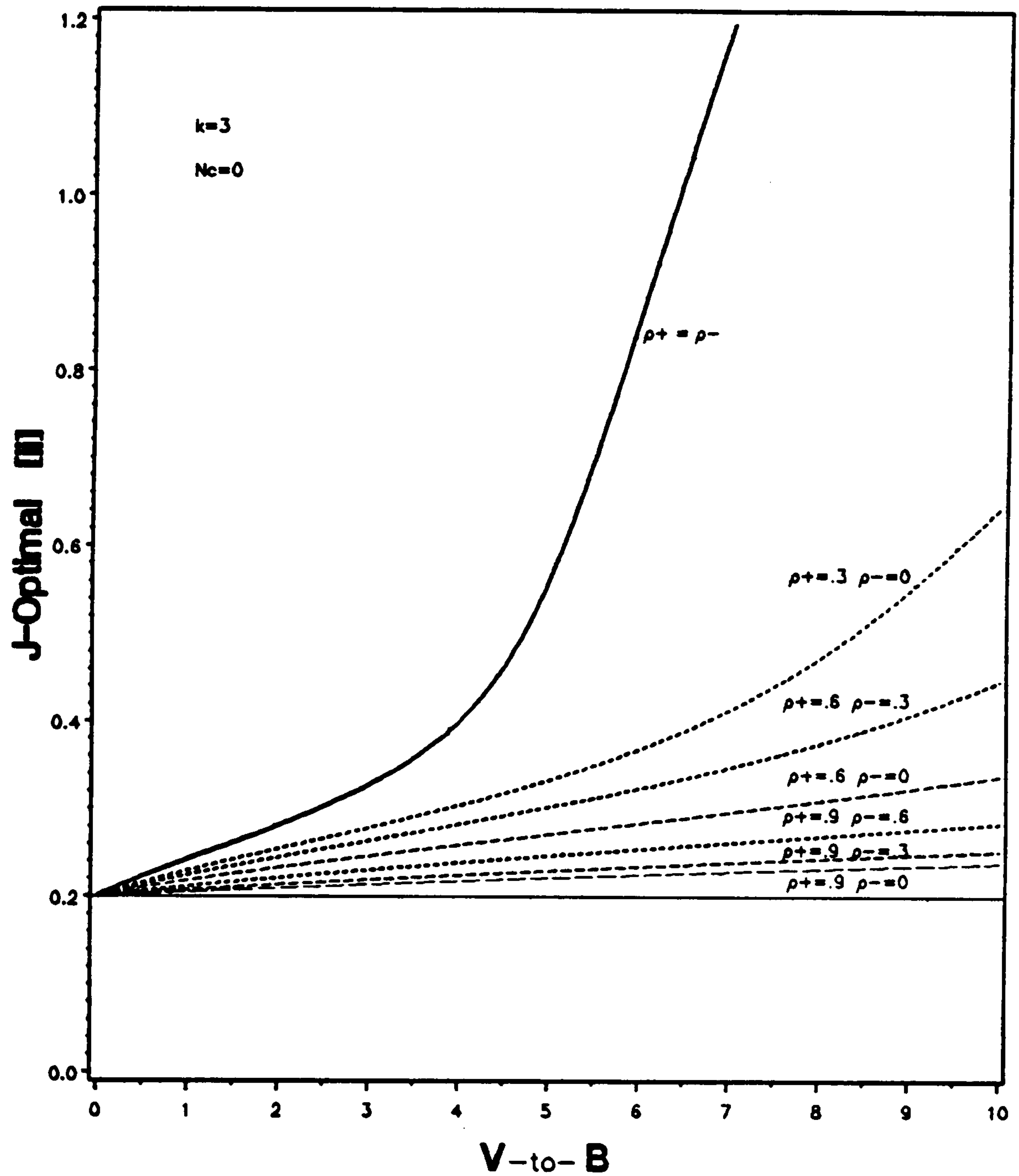


Figure 6. J-optimal [ii] versus V/B for the Pure AR Strategy.

Region of interest is Spherical.

Two-level factorial design with $k = 3$ factors and $N_c = 0$ center runs.

Values of $\rho_- \le \rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The V-optimal value of the second moment is $[ii] = 1.0$.

The horizontal line at $[ii] = 1/(k+2)$ indicates the B-optimal value.

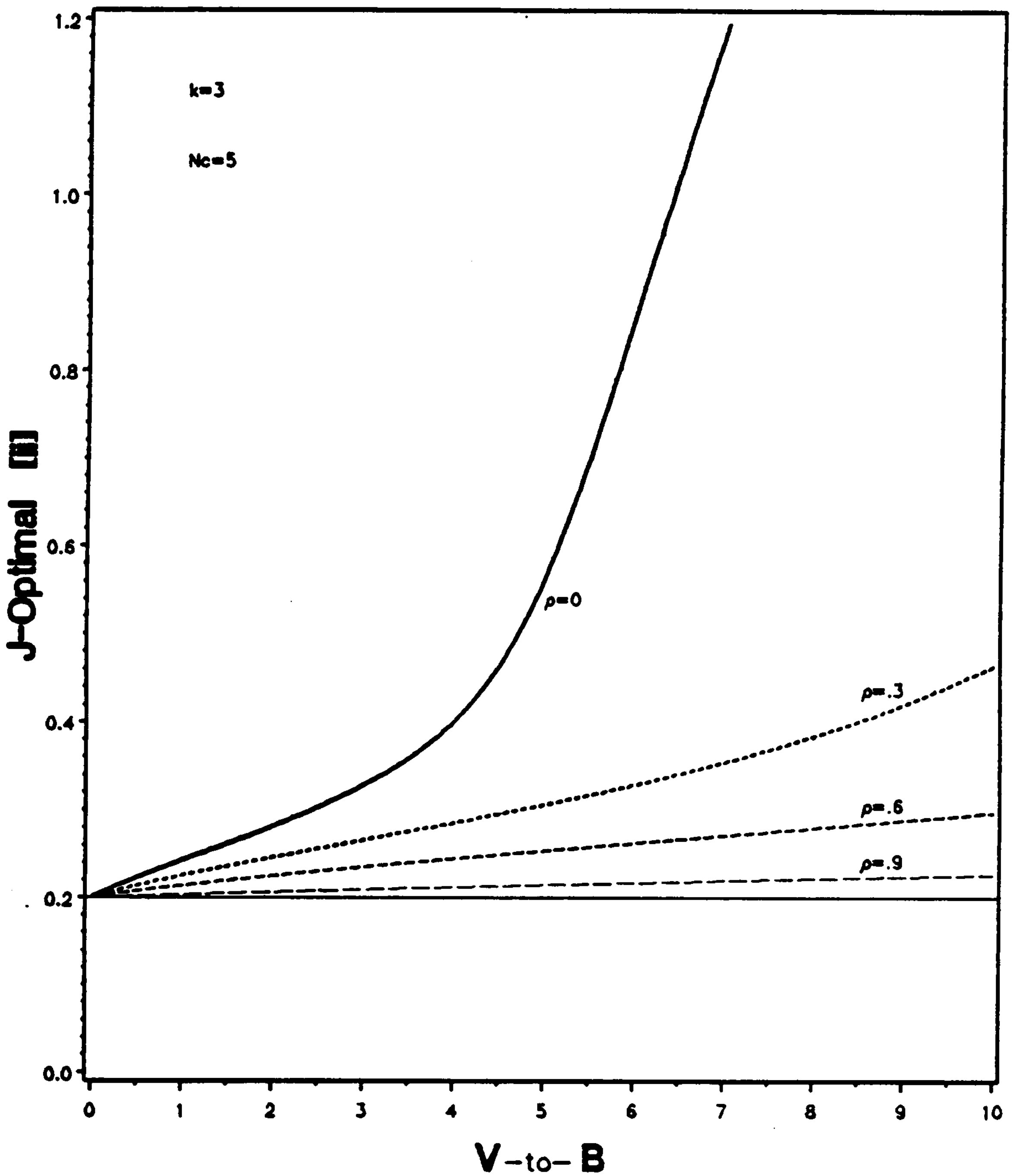


Figure 7. **J-optimal [ii] versus V/B for the** Modified CR Strategy under OLS.

Region of interest is Spherical.

Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.

Values of $\rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The V-optimal value of the second moment is $[ii] = 1.0$.

The horizontal line at $[ii] = 1/(k+2)$ indicates the B-optimal value.

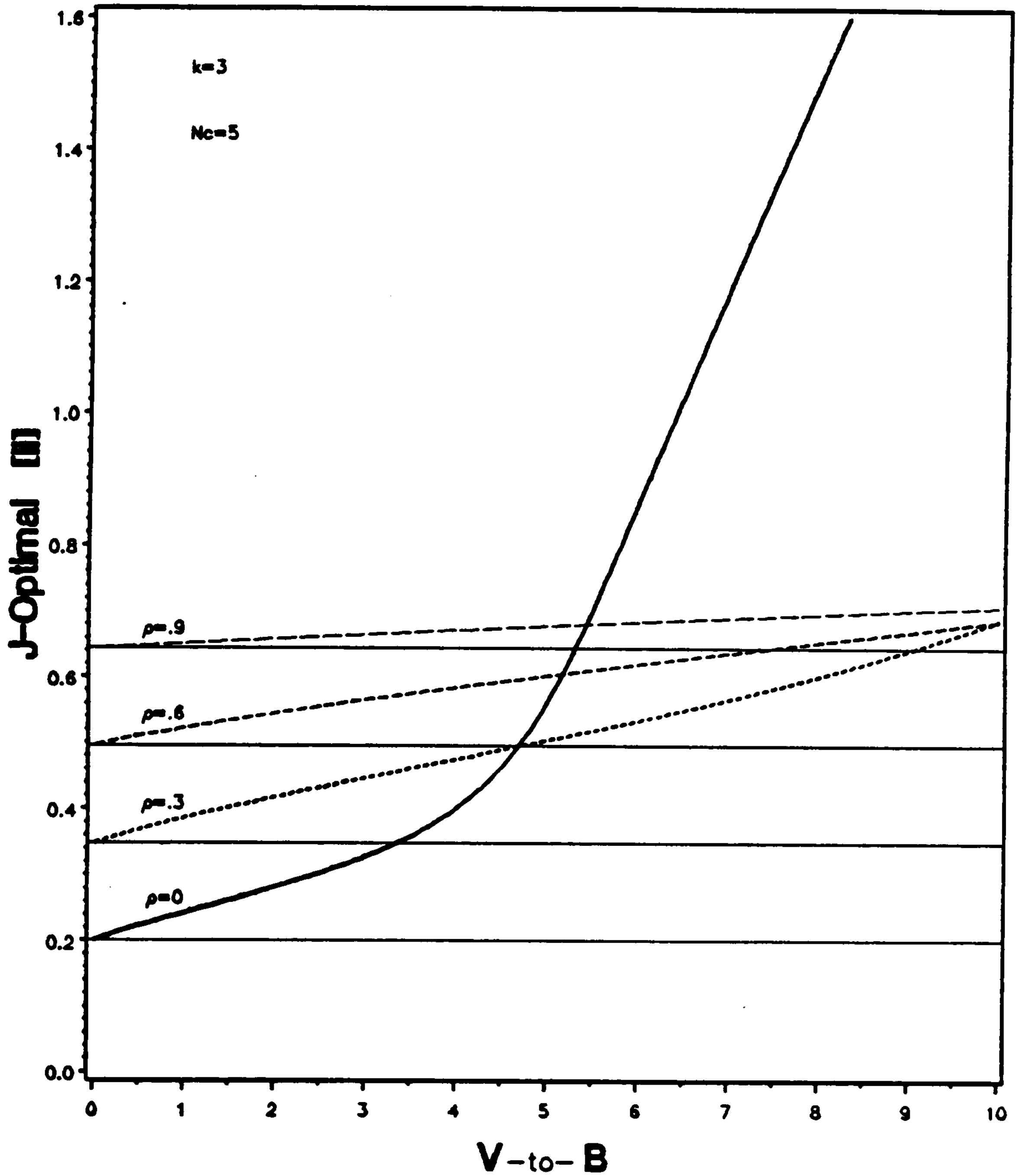


Figure 8. **J-optimal [ii] versus V/B for the Modified CR Strategy under WLS.**
Region of interest is Spherical.
Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.
Values of $\rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.
The V-optimal value of the second moment is $[ii] = 1.0$.
The horizontal lines indicate the B-optimal values.

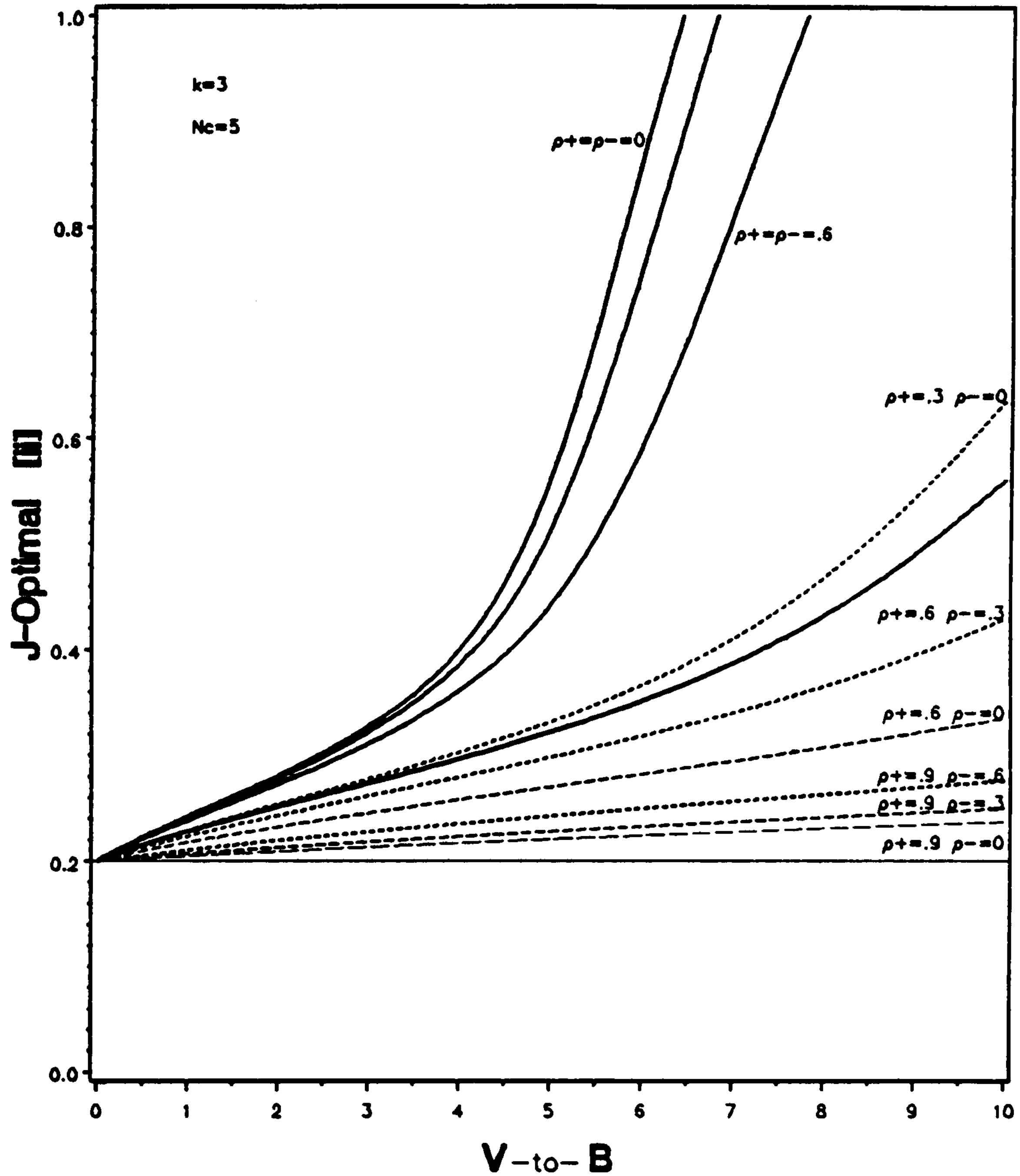


Figure 9. **J-optimal [ii] versus V/B for the Modified AR Strategy under OLS.**
Region of interest is Spherical.
Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.
Values of $\rho_- \le \rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.
The V-optimal value of the second moment is $[ii] = 1.0$.
The horizontal line at $[ii] = 1/(k+2)$ indicates the B-optimal value.

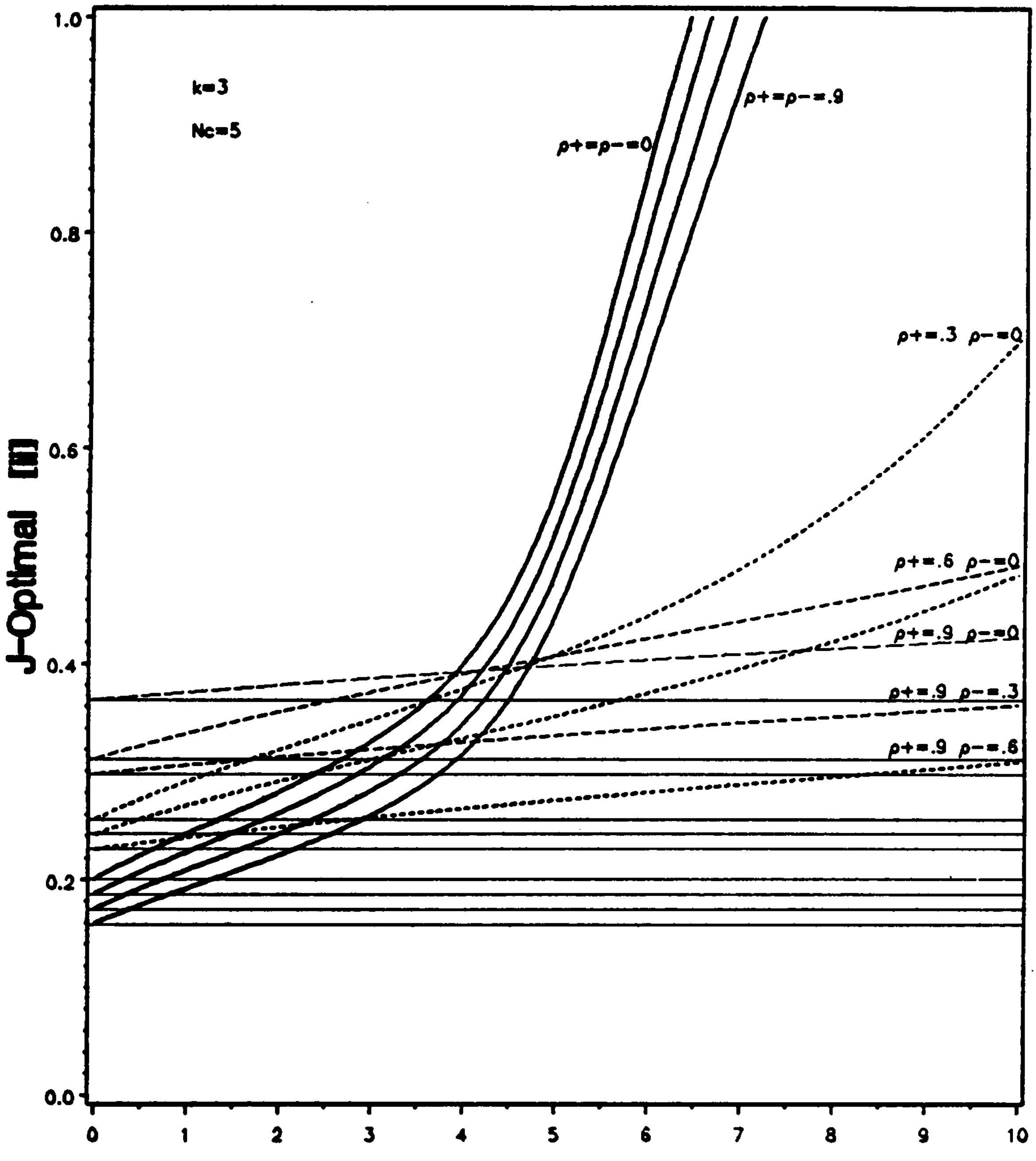


Figure 10. J-optimal [ii] versus V/B for the Modified AR Strategy under WLS.

Region of interest is Spherical.

Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.

Values of $\rho_- \leq \rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The V-optimal value of the second moment is $[ii] = 1.0$.

The horizontal lines indicate the B-optimal values.

3.2.3.3 Values of J For the Correlation Induction Strategies

The MSE of response, **J**, can be computed for designs which minimize **J**, for designs which minimize the **B** component alone, and for designs which minimize the **V** component alone. The **J**-optimal value of [ii] is the positive real root of equation [3.2.35] on page 114, the **B**-optimal value is computed as $[ii] = w_{ii}/z$, and the **V**-optimal value is simply $[ii] = 1$. Utilizing the **J**-, **B**-, and **V**-optimal values of [ii], the three correlation induction schemes are evaluated by comparing the values of, and the differences between, the **J**-, **B**-, and **V**-optimal values of the MSE of response.

The figures presented in this section illustrate the optimal values of **J** as a function of the **V**/**B** ratio for each of the correlation induction strategies using two-level factorial designs with $k = 3$ factors. For the *pure* CR and AR strategies, designs with $N_c = 0$ center runs are used, and for the *modified* CR and AR strategies, designs with $N_c = 5$ center runs are considered (in order to achieve noticeable differences between the pure and modified strategies). The results for the modified strategies are different under OLS and WLS estimation, and therefore separate figures are presented. The results discussed here are illustrated in Figures 11 through 20 (pages 135-144) for a spherical region of interest, and the corresponding figures in Appendix H (pages 330-340) illustrate the results for a cuboidal region of interest.

J-, B-, V-*Optimal Values of* J *versus* V/B *for the Pure CR and AR Strategies*

For the pure CR strategy, the relationship between **J** and **V**/**B** is shown in Figure 11 on page 135. The values of **J** are illustrated for designs which minimize **J**, and for designs which minimize the individual components, **B** and **V**. The results for the IR strategy are indicated whenever $\rho_+ = 0$. When correlation is induced, the **V**-optimal designs result in largest values of **J** and, therefore, are the least desirable in terms of the MSE of response criteria. The **B**-optimal designs perform slightly worse than the **J**-optimal designs, but their performances improve as the magnitude of ρ_+ increases.

However, the values of **J** increase with ρ_+, indicating that the IR strategy performs better than the pure CR strategy in terms of the MSE of response criteria. The results for a cuboidal region, shown in Figure H-11 on page 330, are similar to those for a spherical region.

The results for the pure AR strategy in a spherical region of interest are shown in Figures 12, 13, and 14. (The IR strategy is indicated whenever $\rho_+ = \rho_- = 0$.) Figure 12 on page 136 illustrates the **J**-, **B**-, and **V**-optimal values of the MSE of response for the situation in which $\rho_+ = \rho_-$. The values of **J** decrease as the common value of ρ_+ and ρ_- increases, indicating that the pure AR strategy is preferable to the IR strategy. Similar to the results for the pure CR strategy, the **B**-optimal designs perform well for all values of the **V**/**B** ratio, but the **V**-optimal designs perform poorly when **V**/**B** < 5.

Figure 13 on page 137 illustrates the **J**-, **B**-, and **V**-optimal values of the MSE of response for the pure AR strategy when $\rho_+ > \rho_-$. Similar to Figure 12, in which $\rho_+ = \rho_-$, the **B**-optimal designs perform well for all values of the **V**/**B** ratio, but the **V**-optimal designs perform poorly for small **V**/**B** ratios and become less desirable as the magnitudes of ρ_+ and ρ_- increase. Unlike Figure 12, the AR strategy is not always preferable to the IR strategy. When $\rho_+ > \rho_-$, the AR strategy only performs better than the IR strategy when **V**/**B** < 4 (approximately).

Figure 14 on page 138 illustrates, for various magnitudes of ρ_+ and ρ_-, the **J**-optimal values of the MSE of response for the pure AR strategy. The figure indicates that the AR strategy performs better than the IR strategy whenever $\rho_+ = \rho_- > 0$, but when $\rho_+ > \rho_- > 0$, the AR strategy only performs better over a specific range of **V**/**B** ratios. This range is dependent on the magnitudes of the induced correlations and the number of factors in the model. The larger the magnitudes of ρ_+ and ρ_-, the smaller the difference between ρ_+ and ρ_-, and the smaller the value of k, the larger the **V**/**B** ratio that is necessary for the IR strategy to be preferred. For example, when $k = 3$, $\rho_+ = .9$, and $\rho_- = .6$, the AR strategy performs better than the IR strategy when **V**/**B** < 7, but when

$\rho_- = .3$, the AR strategy only performs better when $\mathbf{V}/\mathbf{B} < 4$. For smaller values of k, the AR strategy performs better than the IR strategy over a slightly wider range of $\mathbf{V}/\mathbf{B}$ ratios. When $\rho_- = 0$, there is little advantage in using the pure AR strategy, because the IR strategy performs better for almost all $\mathbf{V}/\mathbf{B}$ ratios. Thus, the larger the magnitudes of ρ_+ and ρ_-, and the smaller the value of k, the better the performance of the pure AR strategy in terms of the MSE of response criteria.

J-, B-, V-*Optimal Values of* J *versus* V/B *for the Modified CR Strategy*

In order to compare the values of **J** for the IR and modified CR strategies, the least squares estimation technique needs to be specified. Figures 15 and 16 present the **J**-, **V**-, and **B**- optimal values of **J** for the modified CR strategy under OLS and WLS, respectively. The results for OLS, shown in Figure 15 on page 139, are similar to the results for the pure CR strategy shown in Figure 11. Again, the **V**-optimal designs perform poorly when correlation is induced, but the **B**-optimal designs perform well for all values of the $\mathbf{V}/\mathbf{B}$ ratio. The values of **J** are larger for the CR strategy than for the IR strategy, and therefore the IR strategy is always preferable to the modified CR strategy under OLS estimation.

The results for the modified CR strategy under WLS are shown in Figure 16 on page 140. When compared to the OLS results in Figure 15, the values of **J** are smaller and, therefore, a WLS analysis is preferable to an OLS analysis. However, use of a WLS analysis requires knowledge of the magnitudes of the induced correlations, and these magnitudes are generally unknown. Under WLS, the modified CR strategy is preferable to the IR strategy when the $\mathbf{V}/\mathbf{B}$ ratio is less than a particular value of $\mathbf{V}/\mathbf{B}$ which is dependent on the magnitude of ρ_+ and the number of design points. The modified CR strategy performs better over a wider range of $\mathbf{V}/\mathbf{B}$ ratios when the values of N_c and ρ_+ are large and the value of k is small. For example, when $k = 3$, $N_c = 5$, and $\rho_+ = .9$, the modified CR strategy performs better than the IR strategy when $\mathbf{V}/\mathbf{B} < 4$, but when $\rho_+ = .3$, the CR strategy only performs better when $V/B < 3$. Similar to the OLS results, the B-optimal designs perform well for all values of the V/B ratio, and their performance improves as the magnitude of ρ_+ increases.

J-*Optimal Values of* J *versus* V/B *for the Modified AR Strategy*

For the modified AR strategy, the optimal values of J are different under OLS and WLS. Figures 17 and 18 present the J-optimal values of J for the modified AR strategy under OLS and WLS estimation, respectively. The B- and V-optimal values of J are not shown in these figures because the results are similar to those previously discussed for the pure AR strategy. Figure 17 on page 141 indicates that the performance of the modified AR strategy improves as the magnitudes of ρ_+ and ρ_- increase. Whenever $\rho_+ = \rho_- > 0$, the modified AR strategy is preferable to the IR strategy. However, when $\rho_+ > \rho_- > 0$, the modified AR strategy performs better than the IR strategy only when the V/B ratio is less than a particular value of V/B which is dependent on the values of N_c, ρ_+, ρ_-, and k. The smaller the value of k and the larger the values of N_c, ρ_+, and ρ_-, the better the performance of the modified AR strategy under OLS estimation. For example, when $k = 3$, $N_c = 5$, $\rho_+ = .9$, and $\rho_- = .6$, the AR strategy is preferable to the IR strategy when $V/B < 6$, but when $\rho_+ = .6$ and $\rho_- = .3$, the AR strategy only performs better when $V/B < 4$.

The results for the modified AR strategy under WLS are shown in Figure 18 on page 142. The WLS values of J are smaller than the OLS values and, therefore, WLS is the preferred estimation technique. When $\rho_+ = \rho_- > 0$, the AR strategy performs better than the IR strategy, and when $\rho_+ > \rho_- > 0$, the AR strategy performs better when $V/B < 4$. Unlike the OLS results, the modified AR strategy tends to perform better than the IR strategy when $\rho_- = 0$, provided $V/B < 3$. When compared to the OLS results, the modified AR strategy under WLS estimation performs better than the IR strategy over a wider range of V/B ratios, and for smaller magnitudes of the induced correlations.

J-Optimal Values of J versus V/B for Various Values of k and N_c

Figures 19 and 20 illustrate the effects of k, N_c, and the least squares estimation technique on the J-optimal values of the MSE of response. Figure 19 on page 143 shows the optimal values of J for 2^k designs (or fractions) with $k = 2, ..., 7$ factors, and Figure 20 on page 144 illustrates the optimal values of J under OLS and WLS for various values of N_c. These figures correspond to a spherical region of interest and Figures H-19 and H-20 in the Appendix correspond to a cuboidal of interest. For a given V/B ratio, the values of J increase with k, decrease with N_c, and are larger under OLS than WLS estimation. Therefore, the use of multiple center runs and WLS estimation improves the performance of a given design in terms of the MSE of response criteria.

3.2.4 Comparison of the Correlation Induction Strategies

The mathematical developments needed to compare the three correlation induction strategies in terms of the optimal values of [ii] and J were presented in section 3.2.3. The problem considered was that of fitting a first order response model using a two-level factorial design, while protecting against second order bias in the fitted model coefficients. Figures 1 through 20 were developed to illustrate the optimal values of [ii] and J when correlation is induced between simulated responses. Recall that whenever $\rho_+ = 0$, the CR and IR strategies are equivalent, and whenever $\rho_+ = \rho_- = 0$, the AR and IR strategies are equivalent. Therefore, increasing the magnitude of ρ_+ is indicative of the relationship of the CR strategy relative to the IR strategy, and increasing the magnitudes of ρ_+ and ρ_- is indicative of the relationship of the AR strategy relative to the IR strategy. The list on the following two pages summarizes the results for the MSE of response criteria in the first order case.

1. Pure CR Strategy (Figures 1, 5, 11):

- The J-optimal values of [ii] decrease as the magnitude of ρ_+ increases. Therefore, the B-optimal values of [ii] are closer to the J-optimal values for the CR strategy than for the IR strategy.

- For the CR strategy, the J-optimal values of [ii] remain fairly constant for all values of the V/B ratio, but the optimal values increase rapidly for the IR strategy when V/B > 5. Therefore, when compared to the IR strategy, the J-optimal design sizes for the CR strategy are less dependent on the relative amounts of variance and bias errors.

- The B-optimal designs perform well for all values of the V/B ratio, but the V-optimal designs perform poorly when V/B < 5 and their performances deteriorate as the magnitude of ρ_+ increases.

- The optimal values of J increase as the magnitude of ρ_+ increases. Therefore, the IR strategy is always preferable to the pure CR strategy in terms of the MSE of response criteria.

2. Modified CR Strategy under OLS (Figures 1, 7, 15):

- The findings for the modified CR strategy are similar to those for the pure CR strategy. However, for a given V/B ratio, the optimal values of [ii], B, V, and J decrease as the number of center runs increases. Therefore, in terms of the MSE of response criteria, the modified CR strategy generally performs better than the pure CR strategy, but not as well as the IR strategy.

3. Modified CR Strategy under WLS (Figures 2, 8, 16):

- The optimal values of J for the modified CR strategy under WLS are smaller than those under OLS, and therefore WLS is the preferred estimation technique. However, the disadvantage of WLS is that the magnitude of ρ_+ is needed for estimation of the model coefficients.

- Unlike the OLS results, the B- and J-optimal values of [ii] increase as the magnitude of ρ_+ increases. Therefore, the size of the B- and J-optimal designs under WLS are larger for the CR strategy than for the IR strategy.

- Unlike the OLS results, the optimal values of J do not always increase as the magnitude of ρ_+ increases. When V/B < 3 (approximately), the modified CR strategy under WLS performs better than the IR strategy.

- The performance of the modified CR strategy relative to the IR strategy improves as ρ_+ and N_c increase, and as k decreases.

4. Pure AR Strategy (Figures 1, 6, 12, 13, 14):

- For a given **V/B** ratio, the **J**-optimal values of [ii] decrease as the magnitude of ρ_+ increases and as the difference between ρ_+ and ρ_- decreases. Therefore, when compared to the IR strategy, the size of the **B**-optimal designs for the AR strategy are closer to the size of the **J**-optimal designs.

- The **J**-optimal values of [ii] remain fairly constant for all values of the **V/B** ratio when $\rho_+ > \rho_- > 0$, but increase rapidly when $\mathbf{V/B} > 5$ and $\rho_+ = \rho_- = 0$. Therefore, when compared to the IR strategy, the size of a **J**-optimal design for the AR strategy is less dependent on the relative amounts of variance and bias errors.

- When $\rho_+ = \rho_- > 0$, the optimal values of **J** decrease as the common value of ρ_+ and ρ_- increases. Therefore, the pure AR strategy is always preferable to the IR strategy when the magnitudes of ρ_+ and ρ_- are equal.

- When $\rho_+ > \rho_- > 0$, the pure AR strategy performs better than the IR strategy when $\mathbf{V/B} < 3$ (approximately).

- When $\rho_+ > \rho_- = 0$, the optimal values of **J** increase as the magnitude of ρ_+ increases, and therefore the pure AR strategy performs worse than the IR strategy when $\rho_- = 0$.

- The pure AR strategy becomes preferable to IR strategy over a wider range of **V/B** ratios as the magnitudes of ρ_+ and ρ_- increase, the difference between ρ_+ and ρ_- decreases, and the value of k decreases.

5. Modified AR Strategy under OLS (Figures 1, 9, 17):

- The findings for the modified AR strategy are similar to those for the pure AR strategy, except that the number of center runs additionally affects the relative performance of the IR and AR strategies. As N_c increases, the modified AR strategy performs better than the IR strategy over a wider range of **V/B** ratios.

6. Modified AR Strategy under WLS (Figures 3, 10, 18):

- Under WLS estimation, the size of the **B**- and **J**-optimal designs increase as the difference between ρ_+ and ρ_- increases. Only when $\rho_+ = \rho_- > 0$ are the **B**- and **J**- optimal designs smaller for the modified AR strategy than for the IR strategy.

- Unlike the OLS results, the IR strategy is not always preferable to the AR strategy under WLS estimation when $\rho_+ > \rho_- = 0$; the modified AR strategy generally performs better than the IR strategy when $\mathbf{V/B} < 3$.

- The performance of the modified AR strategy under WLS improves as ρ_+, ρ_-, and N_c increase, and as k decreases.

These results indicate that the AR strategy performs better than the IR strategy whenever $\rho_+ = \rho_- > 0$, but when $\rho_+ > \rho_- > 0$, the AR strategy only performs better than the IR strategy when $\mathbf{V}/\mathbf{B} < 3$ (approximately). The IR strategy performs better than the CR strategy under OLS estimation, but under WLS estimation, the CR strategy performs better than the IR strategy when $\mathbf{V}/\mathbf{B} < 3$. For both the CR and AR strategies, the size of the J-optimal designs become smaller and approach the size of the B-optimal designs as the magnitudes of the induced correlations increase. Therefore, when correlation is induced among simulated responses, is it even more important that the size of the design be reduced by moving the design points away from the ± 1 extremes and closer to the center of the design.

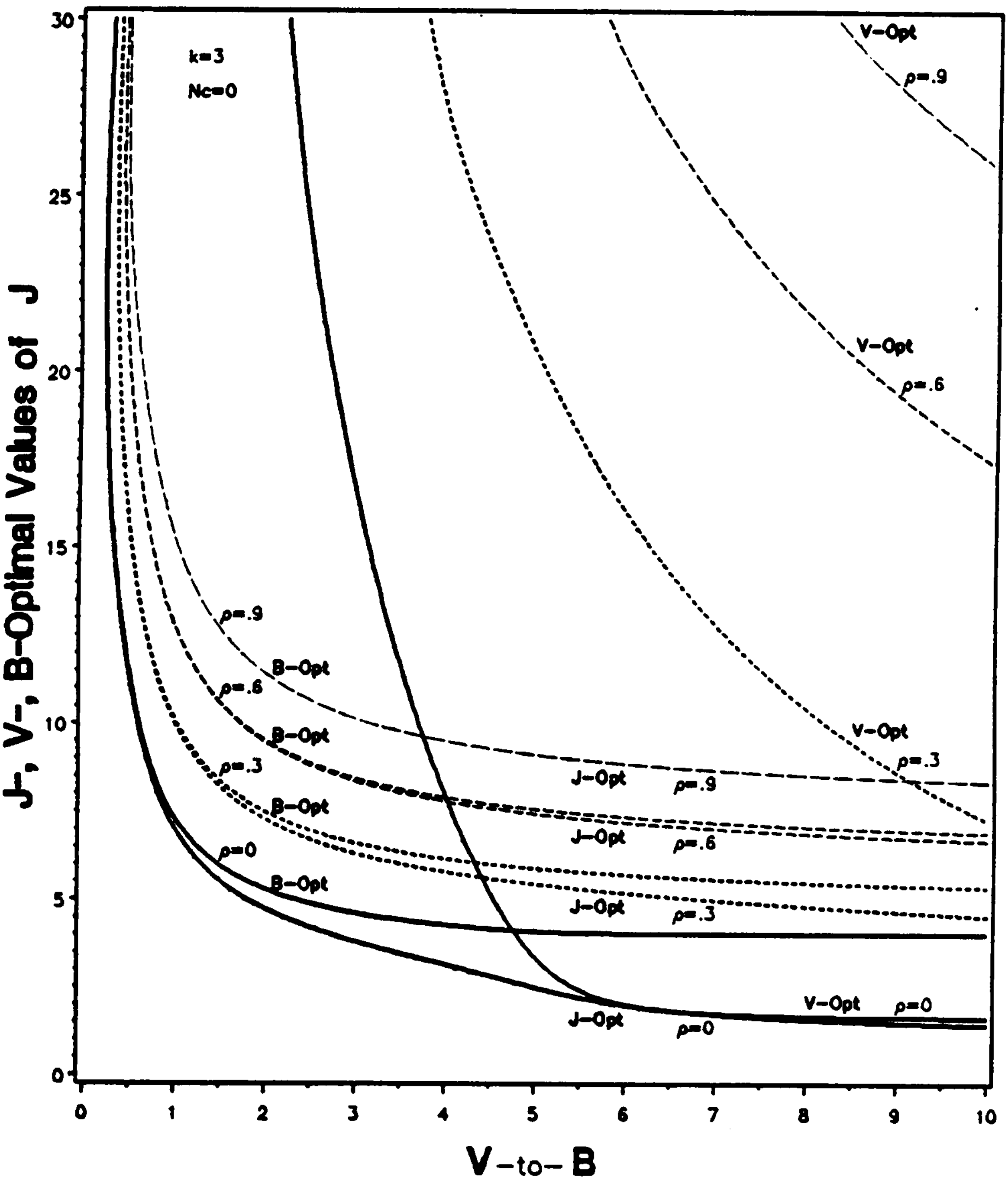


Figure 11. J-, V-, B-optimal values of J versus V/B for the Pure CR Strategy.

Region of interest is Spherical.

Two-level factorial design with $k = 3$ factors and $N_c = 0$ center runs.

Values of $\rho_* = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

The B-optimal values of J are computed using $[ii] = w_{ii} = 1/(k+2)$.

The V-optimal values of J are computed using $[ii] = 1$.

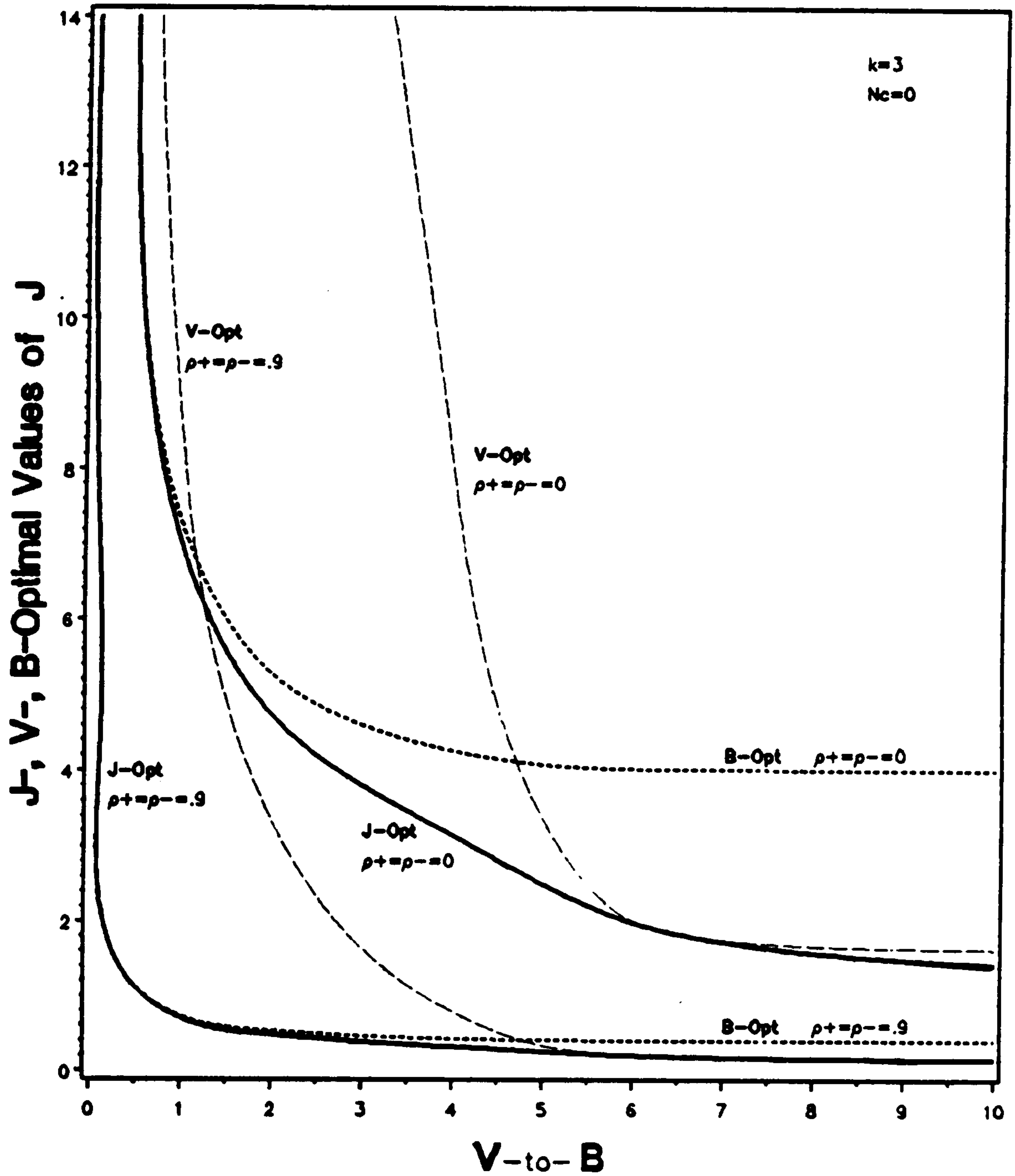


Figure 12. J-, V-, B-optimal values of J versus V/B for the Pure AR Strategy.

Region of interest is Spherical.

Two-level factorial design with $k = 3$ factors and $N_c = 0$ center runs.

Value of $\rho_- = \rho_+ = .9$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

The B-optimal values of J are computed using $[ii] = w_{ii} = 1/(k+2)$.

The V-optimal values of J are computed using $[ii] = 1$.

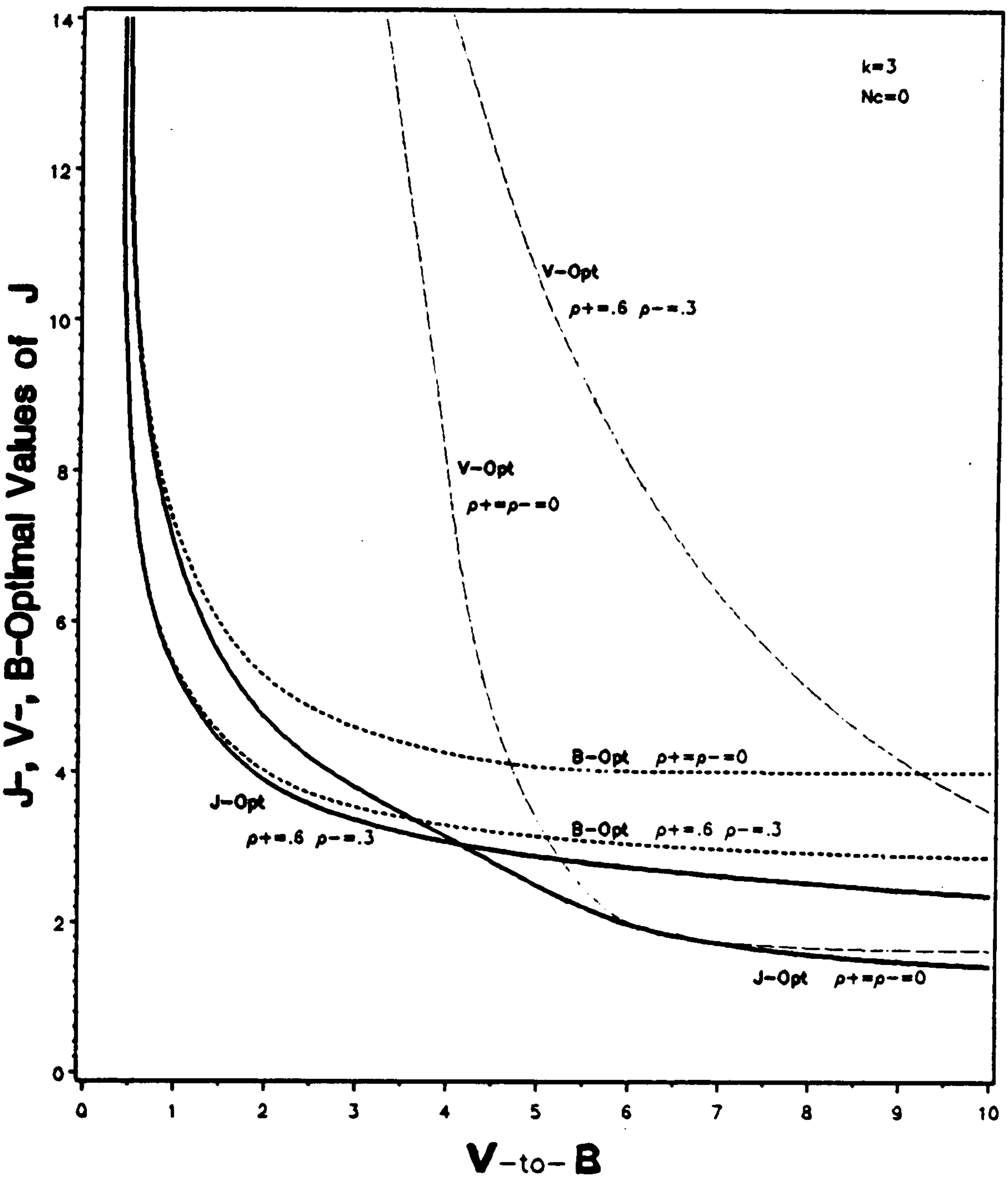


Figure 13. J-, V-, B-optimal values of J versus V/B for the Pure AR Strategy.

Region of interest is Spherical.

Two-level factorial design with $k = 3$ factors and $N_c = 0$ center runs.

Value of $\rho_+ = .6$ and $\rho_- = .9$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V B ratio.

The B-optimal values of J are computed using $[ii] = w_{ii} = 1/(k+2)$.

The V-optimal values of J are computed using $[ii] = 1$.

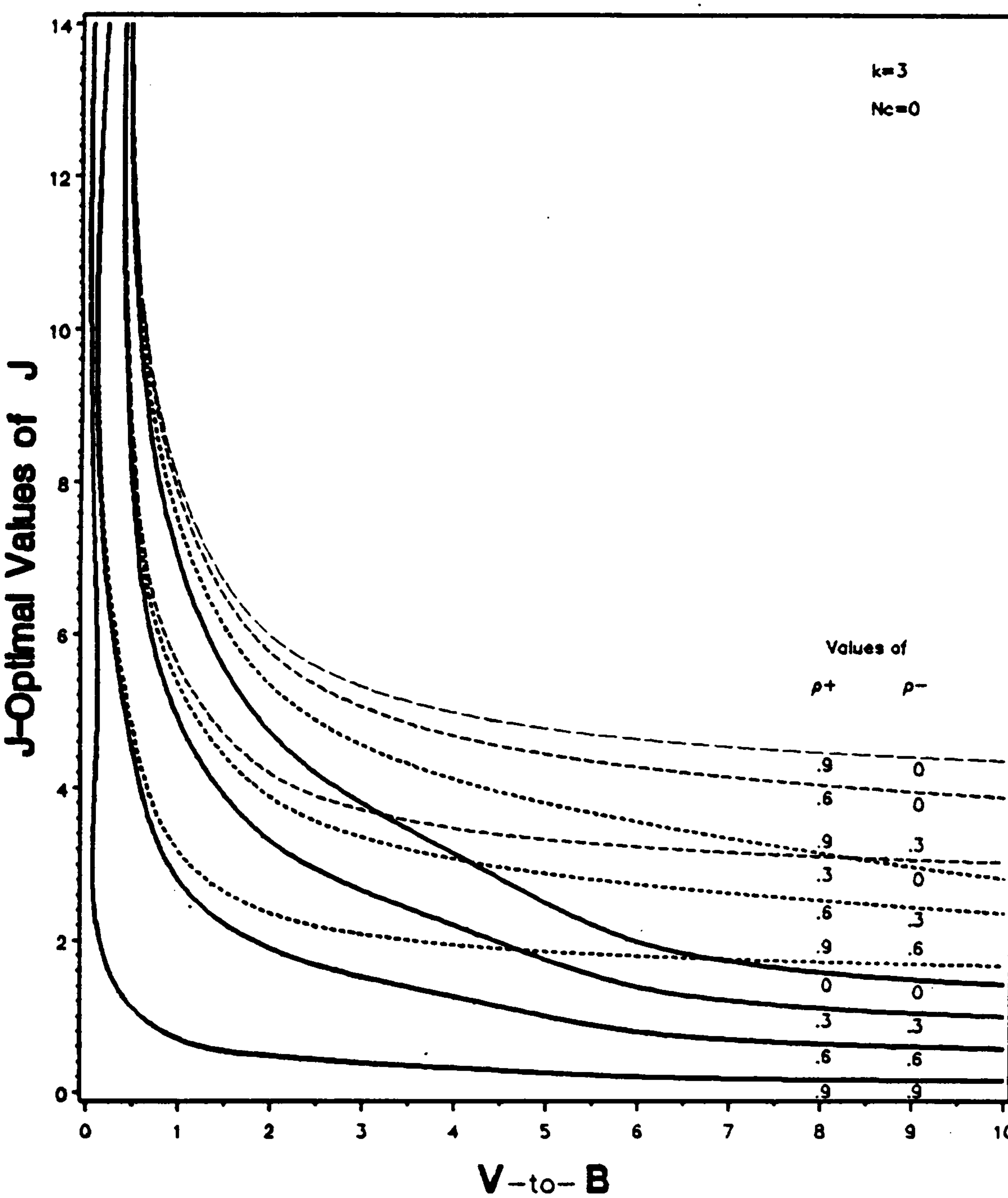


Figure 14. J-optimal values of J versus V/B for the Pure AR Strategy.

Region of interest is Spherical.

Two-level factorial design with $k = 3$ factors and $N_c = 0$ center runs.

Value of $\rho_- \leq \rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

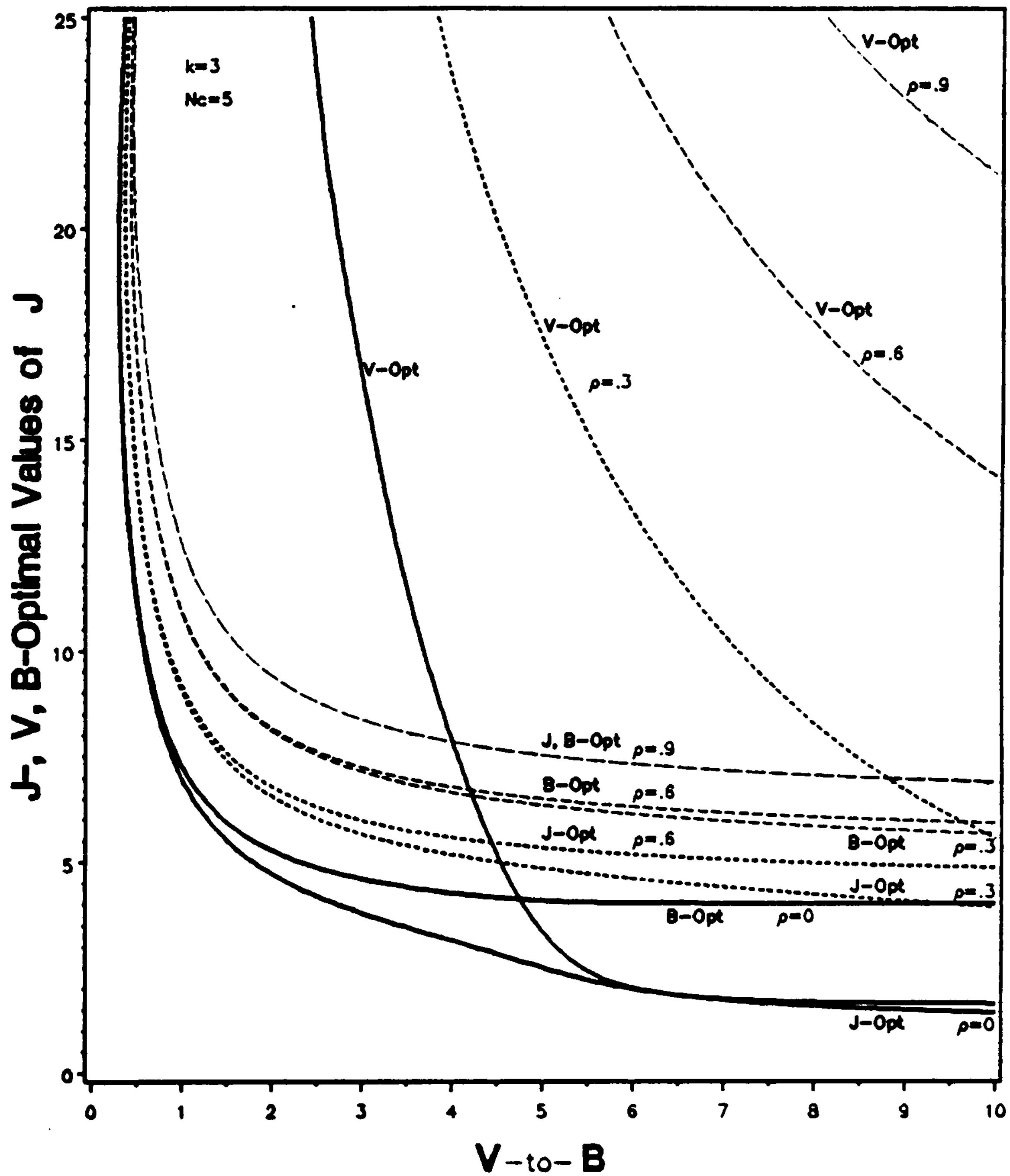


Figure 15. J-, V-, B-optimal values of J versus V/B for the Modified CR Strategy under OLS.

Region of interest is Spherical.

Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.

Values of $\rho_* = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

The B-optimal values of J are computed using $[ii] = w_{ii} = 1/(k+2)$.

The V-optimal values of J are computed using $[ii] = 1$.

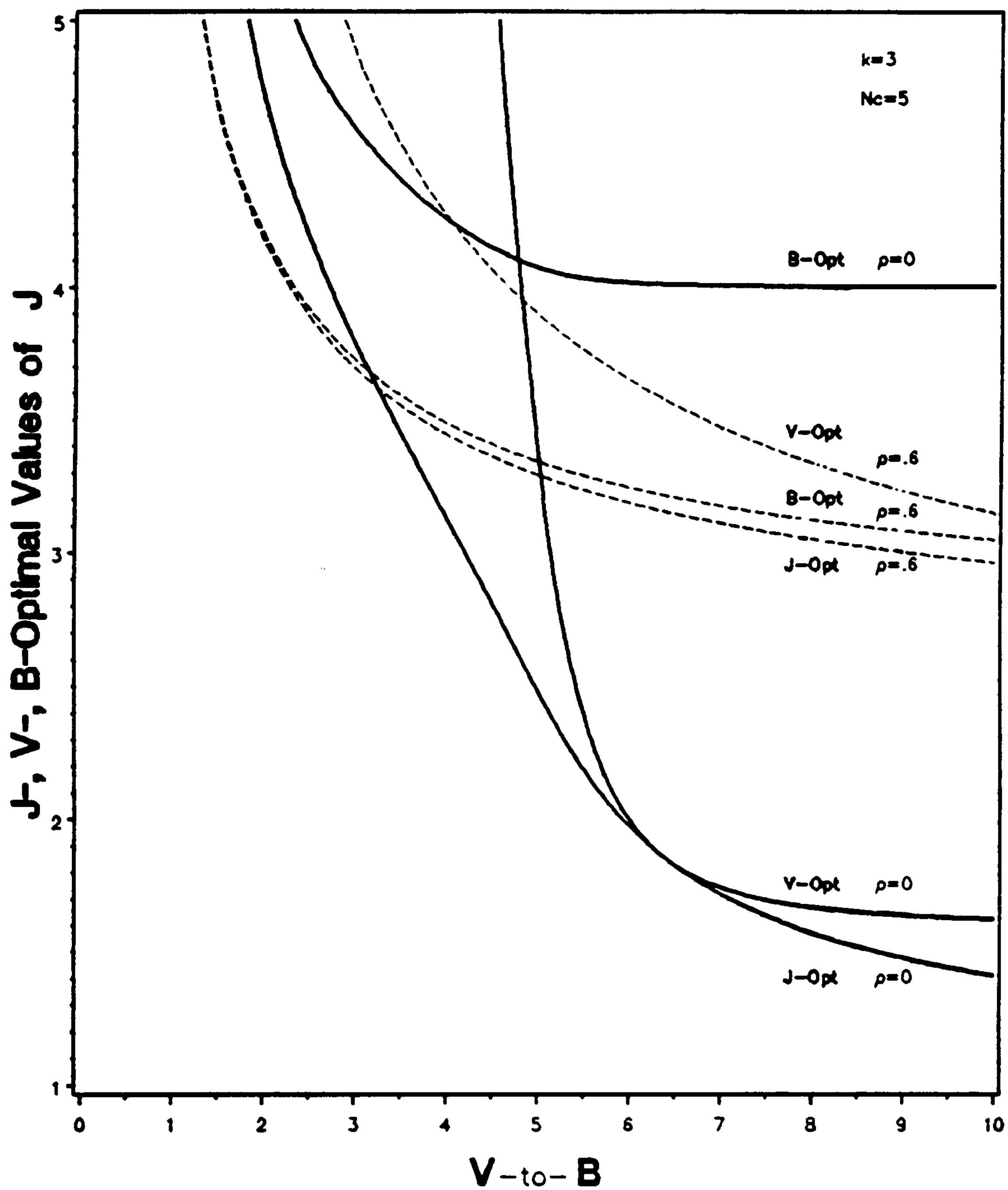


Figure 16. J-, V-, B-optimal values of J versus V/B for the Modified CR Strategy under WLS.

Region of interest is Spherical.

Two-level factorial design with $k=3$ factors and $N_c = 5$ center runs.

Values of $\rho_{+} = 0, .6$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

The B-optimal values of J are computed using $[ii] = w_{ii}/z = 1/z \; (k+2)$.

The V-optimal values of J are computed using $[ii] = 1$.

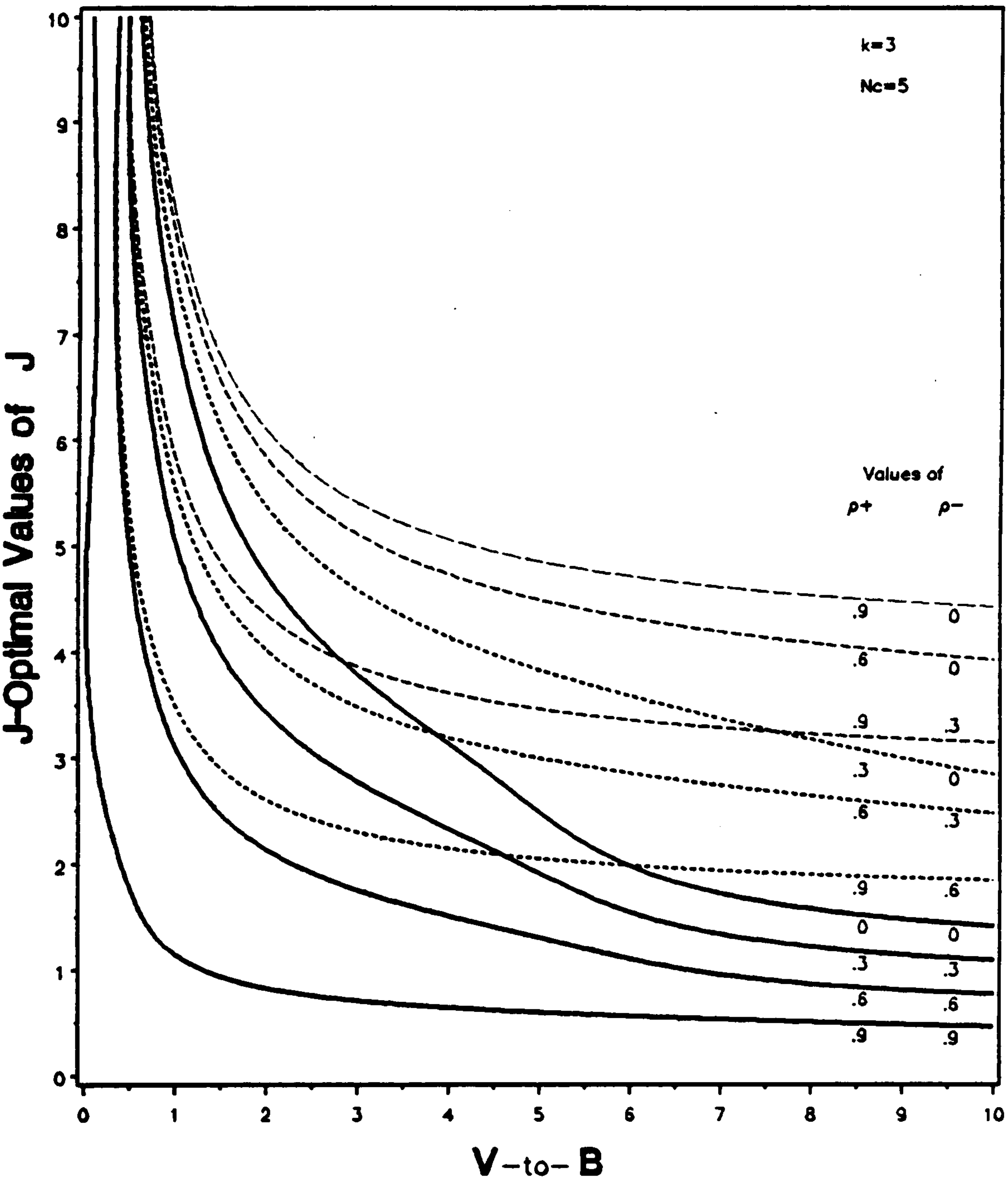


Figure 17. J-optimal values of J versus V/B for the Modified AR Strategy under OLS.

Region of interest is Spherical.

Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.

Value of $\rho_- \leq \rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

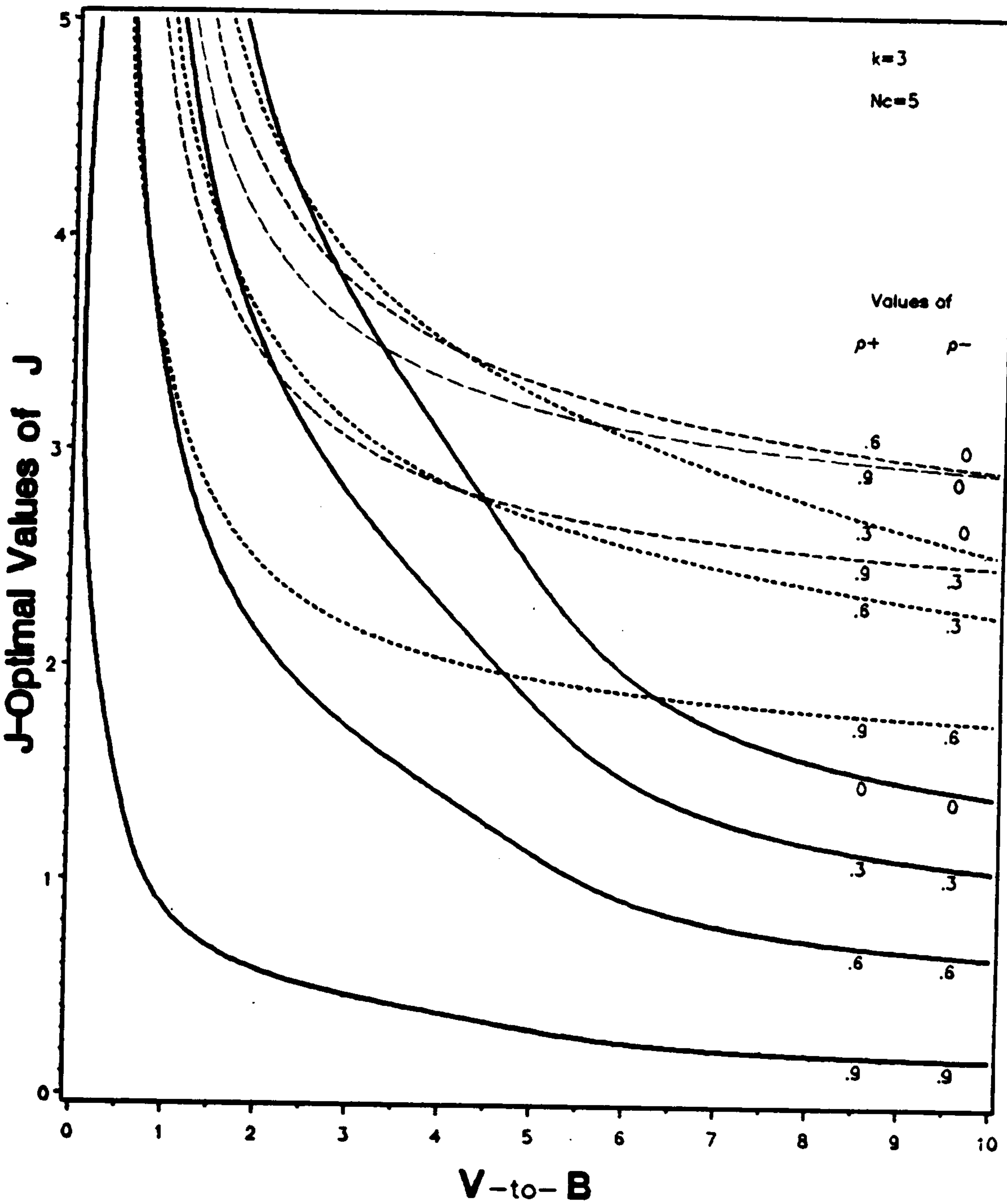


Figure 18. J-optimal values of J versus V/B for the Modified AR Strategy under WLS.

Region of interest is Spherical.

Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.

Value of $\rho_- \le \rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

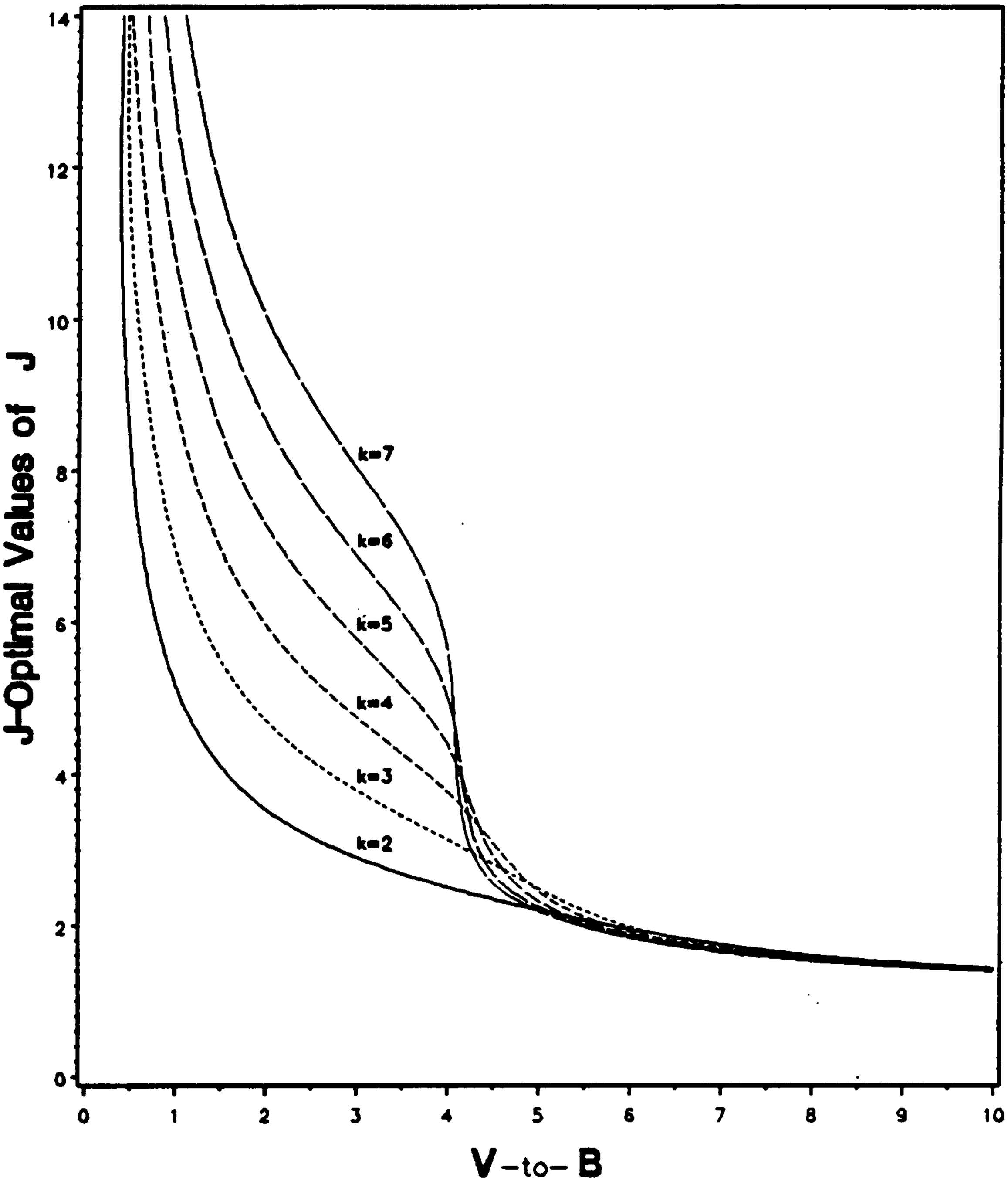


Figure 19. **J-optimal values of J versus V/B for various values of k.**

IR Strategy .

Two-level factorial design with $k = 2, ..., 7$ factors.

Value of $N_c = 0$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

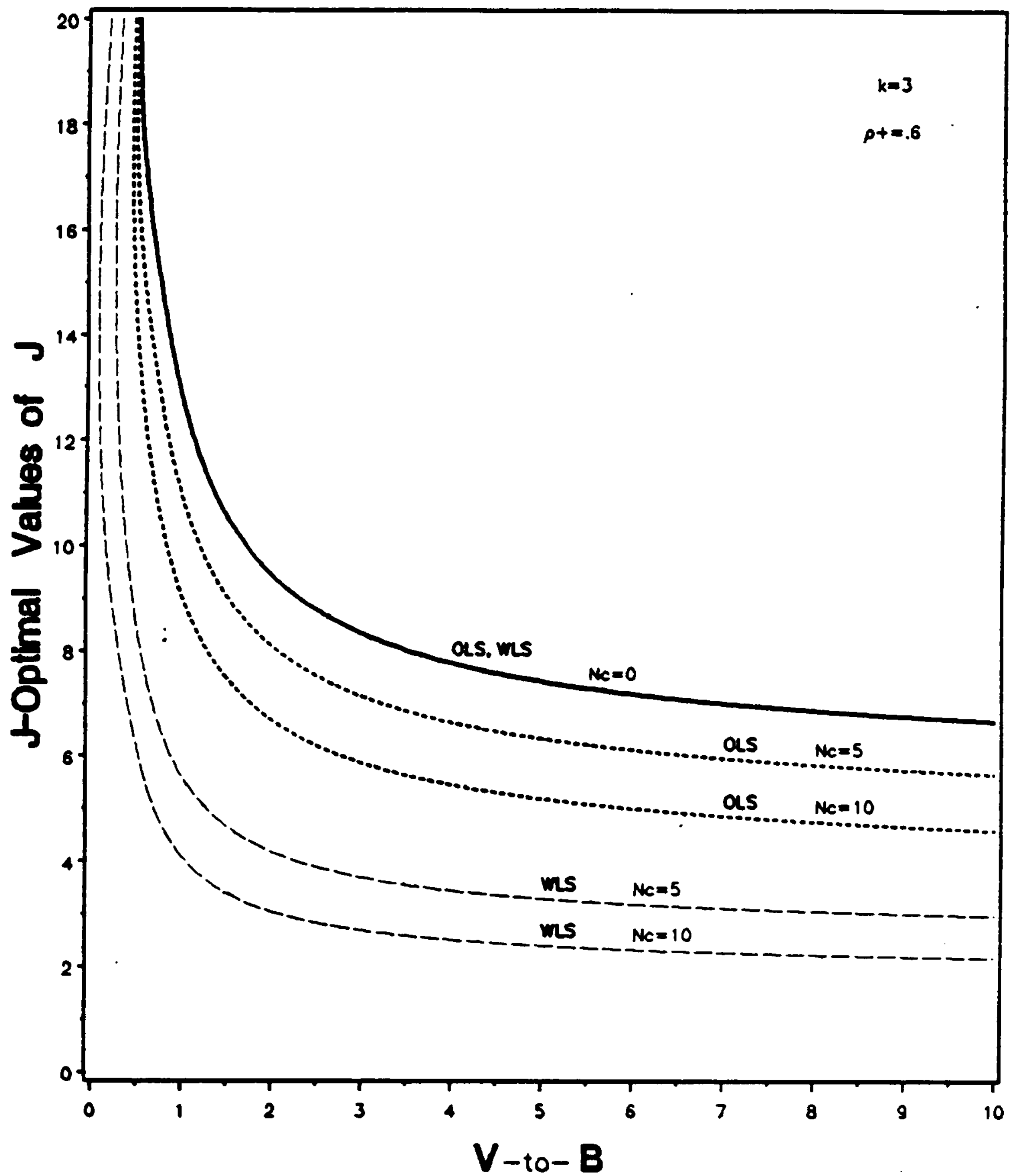


Figure 20. J-optimal values of J versus V/B for designs with various values of N.

CR Strategy with $\rho_{*} = .6$.
Least squares estimation techniques are OLS and WLS.
Two-level factorial design with $k = 3$ factors.
Values of N_c = 0, 5, 10 and the value of $\phi = k/2$.
The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

3.3 *MSE of Slope Criteria in the First Order Case*

The focus of this section is the evaluation of the **IR**, **CR**, and **AR** correlation induction strategies in terms of the MSE of slope criteria. The design plans used to evaluate the correlation induction strategies are two-level factorial designs with $k = 2, \ldots, 7$ factors, with the $k = 5, 6$ designs being ½-fractions, and the $k = 7$ design being a ¼-fraction. It is assumed that a first order polynomial model is being used to fit the simulated response data and protection against second order bias in the fitted model coefficients is desired.

The MSE of slope criteria, discussed on pages 73-78, is a generalization of the MSE of response criteria in which the objective is to minimize the average, normalized mean squared error of the partial derivatives (or slopes) of the fitted response function. This criteria is useful during the steepest ascent and ridge analysis stages of RSM, when the objective is determination of the direction in which the change in the estimated response is the greatest. The equations for the partial derivatives of the fitted and true response functions, as shown in equations [2.2.50] and [2.2.49] on pages 75-76, can be written as

$$\hat{\underline{\gamma}}_{(\underline{x})} = \frac{\partial \hat{\underline{y}}}{\partial \underline{x}} = \Lambda'_{1\,(\underline{x})}\, \underline{b}_1$$

$$\underline{\gamma}_{(\underline{x})} = \frac{\partial \underline{y}}{\partial \underline{x}} = \Lambda'_{1\,(\underline{x})}\, \underline{\beta}_1 + \Lambda'_{2\,(\underline{x})}\, \underline{\beta}_2$$

where the $\Lambda'_{1(\underline{x})}$ matrix contains the partial derivatives of the fitted first order regressor terms and the $\Lambda'_{2(\underline{x})}$ matrix contains the partial derivatives of the unfitted second order and two-way interaction terms. Both of these matrices are illustrated in Appendix C (pages 283-286). The MSE of slope, $\mathbf{J}^*$, is the average, normalized mean squared error of $\hat{\underline{\gamma}}_{(\underline{x})}$, computed as the sum of $\mathbf{B}^*$ and $\mathbf{V}^*$, the bias and variance components of $\mathbf{J}^*$. The next two sections present the mathematical developments of the $\mathbf{B}^*$ and $\mathbf{V}^*$ components, and section 3.3.3 compares the correlation induction strategies in terms of $\mathbf{J}^*$.

3.3.1 Bias Component of $\mathbf{J}^*$ in the First Order Case

The bias component of $\mathbf{J}^*$ is computed as the average squared bias of the slopes of the response function, normalized with respect to N and σ^2. Myers and Lahoda (48) present a mathematical development of $\mathbf{B}^*$ for the situation in which the responses are *uncorrelated*; that is, for the **IR** correlation induction strategy. This research extends their work by considering situations in which the responses are *correlated* and develops the equations for $\mathbf{B}^*$ under the **CR** and **AR** correlation induction strategies.

Myers and Lahoda (48) define the bias component of $\mathbf{J}^*$, as shown in equation [2.2.56] on page 78, as

$$\begin{aligned}
\mathbf{B}^* &= \frac{N\Omega_r}{\sigma^2} \int_R \text{Bias}^2 \left[\hat{\underline{y}}_{(\underline{x})} \right] d\underline{x} \\
&= \frac{N\Omega_r}{\sigma^2} \int_R \left\{ \left[E(\hat{\underline{y}}_{(\underline{x})}) - \underline{y}_{(\underline{x})} \right]' \left[E(\hat{\underline{y}}_{(\underline{x})}) - \underline{y}_{(\underline{x})} \right] \right\} d\underline{x} \\
&= \frac{N}{\sigma^2}\, \underline{\beta}'_2 \left\{ A' \mu^*_{11} A - 2 \mu^*_{21} A + \mu^*_{22} \right\} \underline{\beta}_2
\end{aligned}$$

where Ω_r is the inverse of the volume of the centered and scaled region of interest (defined on page 71), A is the alias matrix (defined on page 43), and the μ^* terms are the region matrices of the partial derivatives of the response function (defined on page 77).

The equation for $\mathbf{B}^*$ can be simplified when the fitted model is first order and the experimental design is orthogonal. Appendix C (pages 283-286) illustrates the μ^*_{11}, μ^*_{21}, and μ^*_{22} matrices for the $d_1 = 1$, $d_2 = 2$ fit-protection situation, in which the number of parameters in the fitted model is $p_1 = k+1$ and the number of unfitted second order parameters is $p_2 = \frac{1}{2}k(k+1)$. Because the elements of the region matrices which involve odd powers of x_i are equal to zero, μ_{21}^* reduces to a null matrix, and μ_{11}^* and μ_{22}^* reduce to diagonal matrices; that is

$$\mu_{11}^* = \text{Diag}\ (0, 1, 1, \ldots\ldots, 1) \qquad [3.3.1]$$

$$\mu_{21}^* = \mathbf{0} \qquad [3.3.2]$$

$$\mu_{22}^* = \text{Diag}\ (4\,w_{ii}, \ldots\ldots, 4\,w_{ii},\ \ 2\,w_{ii}, \ldots\ldots, 2\,w_{ii}) \qquad [3.3.3]$$

where μ_{11}^* is a $(p_1 \times p_1)$ diagonal matrix, μ_{21}^* is a $(p_2 \times p_1)$ null matrix, and μ_{22}^* is a $(p_2 \times p_2)$ diagonal matrix whose first *k* diagonal elements are $4\,w_{ii}$, and the remaining $\frac{1}{2}k(k-1)$ diagonal elements are $2\,w_{ii}$. The w_{ii} scalar term is the pure second order region moment which is defined in equations [3.2.7] and [3.2.9] on page 94.

Appendix E (pages 292-296) illustrates the alias matrices for first order orthogonal designs under both OLS and WLS estimation. Both of these matrices consist of nonzero elements in the first *k* columns of the first row, and all other elements of the alias matrices are zero. Therefore, the first two component matrices in the equation for $\mathbf{B}^*$ become

$$A' \mu_{11}^* A = \mathbf{0}$$

$$2 \mu_{21}^* A = \mathbf{0}$$

where the $\mathbf{0}$'s are $(p_2 \times p_2)$ null matrices. Thus, the equation for $\mathbf{B}^*$ simplifies to

$$\mathbf{B}^* = \frac{N}{\sigma^2}\, \underline{\beta}'_2\, \mu^*_{22}\, \underline{\beta}_2$$

$$= \frac{N}{\sigma^2}\left[\sum_{i=1}^{k}\beta_{ii}^2 + \tfrac{1}{2}\sum_{i<j}\sum\beta_{ij}^2\right]4\,w_{ii} \qquad [3.3.4]$$

$$= 4\,w_{ii}\,\theta$$

where θ is a standardized measure of the quadratic tendency of the response surface (defined in equation [3.2.2] on page 92).

Equation [3.3.4] indicates that the bias component of $\mathbf{J}^*$ is a function of the unknown β parameters in the protection model through θ, and depends on the shape of the region of interest through w_{ii}. However, the $\mathbf{B}^*$ component is independent of the correlation induction strategy and the value of [ii]. Therefore, all values of [ii] yield the same value of $\mathbf{B}^*$, and there is no single value which is considered $\mathbf{B}^*$-optimal.

3.3.2 Variance Component of $\mathbf{J}^*$ in the First Order Case

The variance component of $\mathbf{J}^*$ is computed as the average variance of the partial derivatives of the response function, normalized with respect to N and σ^2. Myers and Lahoda (48) present a mathematical development of $\mathbf{V}^*$ for the situation in which the responses are uncorrelated, and Hussey, Myers, and Houck (28) expand on their work by allowing for correlated responses and develop the equations for $\mathbf{V}^*$ under the pure CR and AR correlation induction strategies. This research expands on the previous work by developing the equations for $\mathbf{V}^*$ under the modified CR and AR correlation induction strategies.

Myers and Lahoda (48) define the variance component of $\mathbf{J}^*$, as shown in equation [2.2.55] on page 77, as

$$
\begin{aligned}
\mathbf{V}^* &= \frac{N\Omega_r}{\sigma^2} \int_R \mathrm{Var}\,[\,\hat{\underline{y}}_{(\underline{x})}\,]\, d\underline{x} \\
&= \frac{N\Omega_r}{\sigma^2} \int_R E\,\{\,[\,\hat{\underline{y}}_{(\underline{x})} - E\,(\hat{\underline{y}}_{(\underline{x})})]'\ [\,\hat{\underline{y}}_{(\underline{x})} - E\,(\hat{\underline{y}}_{(\underline{x})})]\,\}\, d\underline{x} \\
&= \frac{N}{\sigma_2}\ \mathrm{Trace}\left\{ \mathrm{Var}\,[\,\underline{b}_1\,]\ \mu_{11}^{*} \right\}
\end{aligned}
$$

where Ω_r is the inverse of the volume of the centered and scaled region of interest, $\mathrm{Var}\,[\,\underline{b}_1\,]$ is the variance-covariance matrix of the least squares estimators of the fitted model coefficients, and μ_{11}^{*} is the region matrix of the slopes of the fitted response function. When a two-level factorial design is used to fit a first order model, the μ_{11}^{*} matrix (shown in equation [3.3.1] on page 147) and the $\mathrm{Var}\,[\,\underline{b}_1\,]$ matrix (shown in equation [3.2.17] on page 99) are both diagonal matrices, and the equation for $\mathbf{V}^*$ becomes

$$
\begin{aligned}
\mathbf{V}^* &= \frac{N}{\sigma_2}\ \mathrm{Trace}\,\{\, \mathrm{Var}\,[\,\underline{b}_1\,]\ \mu_{11}^{*} \,\} \\
&= \frac{N}{\sigma_2}\ \{\sum_{i=1}^{k} \mathrm{Var}\,(\,b_i\,)\,\} \\
&= \frac{N}{\sigma_2}\ \{\,k\,\mathrm{Var}\,(\,b_i\,)\,\} \\
&= \frac{k\,(1-\rho_+\,)}{[ii]}\ .
\end{aligned}
\qquad [3.3.5]
$$

The equation for $\mathbf{V}^*$ can now be used to find the value of [ii] which results in a Min-$\mathbf{V}^*$ design. The optimal value of [ii] is found by setting the partial derivative of $\mathbf{V}^*$ with respect to [ii] equal to zero, yielding

$$
\frac{\partial\,\mathbf{V}^*}{\partial\,[ii]} = -\frac{k\,(1-\rho_+\,)}{[ii]^2} = 0\ . \qquad [3.3.6]
$$

The second partial derivative of $\mathbf{V}^*$ with respect to [ii], $\frac{\partial^2 \mathbf{V}^*}{\partial\,[\mathrm{ii}]^2} = \frac{k\,(1-\rho_+)}{[\mathrm{ii}]^3}$, is always a positive quantity and, therefore, the solution of equation [3.3.6] provides the value of [ii] which minimizes $\mathbf{V}^*$. The optimal value of [ii] for a Min-$\mathbf{V}^*$ design is to make [ii] as large as possible, implying that the $\mathbf{V}^*$-optimal value of [ii] for a coded design region is [ii] = 1.

Thus, the $\mathbf{V}^*$-optimal value of [ii] for a two-level factorial design is independent of the shape of the region of interest, the correlation induction strategy, and the least squares estimation technique. In addition, the optimal values of $\mathbf{V}^*$ are equivalent under OLS and WLS estimation because the variances of the b_i coefficients are equivalent and the variance of the b_0 intercept term (which is reduced under WLS) does not affect $\mathbf{V}^*$. In the next section the MSE of slope criteria, which calls for minimizing the sum of $\mathbf{B}^*$ and $\mathbf{V}^*$, is used to compare the correlation induction strategies in terms of the optimal values of $\mathbf{J}^*$.

3.3.3 J* - Optimum Designs

A design which results in a minimum value for the MSE of slope is considered to be a $\mathbf{J}^*$-optimum design. Myers and Lahoda (48) define $\mathbf{J}^*$ as the average mean squared error of the partial derivatives (or slopes) of the fitted response function and compute $\mathbf{J}^*$ as the sum of the bias and variance components, $\mathbf{B}^*$ and $\mathbf{V}^*$. For the situation in which a two-level factorial design is being used to fit a first order model, the bias and variance components of $\mathbf{J}^*$ were defined in sections 3.3.1 and 3.3.2. Utilizing equation [3.3.4] on page 148 for $\mathbf{B}^*$, and equation [3.3.5] on page 149 for $\mathbf{V}^*$, the equation for $\mathbf{J}^*$ becomes

$$\begin{aligned} \mathbf{J}^* &= \mathbf{B}^* + \mathbf{V}^* \\ &= 4\,w_{ii}\,\theta + \frac{k\,(1-\rho_+)}{[\mathrm{ii}]} . \end{aligned} \qquad [3.3.7]$$

Equation [3.3.7] does not involve the magnitude of ρ_- nor the number of center runs in the design. Therefore, the value of **J*** is the same for the pure CR, pure AR, modified CR, and modified AR correlation induction strategies. Also, since $\rho_+ = 0$ for the IR strategy, the value of **J*** is higher for the IR strategy than for the CR and AR strategies.

The equation for **J*** can be used to find the value of [ii] which results in a Min-**J*** design. The optimal value of [ii] could be found by setting the partial derivative of **J*** with respect to [ii] equal to zero. However, since **B*** is not a function of [ii], the **J***-optimal value of [ii] is the same as the **V***-optimal value; that is, as large as possible. Thus, for a coded design region, the **J***-optimal designs have pure second order design moments of $[ii] = 1$.

In order to compare the correlation induction strategies in terms of the MSE of slope criteria, the optimal values of **J*** can be viewed as a function of the θ bias parameter or as function of the **V*/B*** ratio. Figure 21 on page 152 illustrates the optimal values of **J*** as a function of θ for various values of ρ_+ in a spherical region of interest. The figure indicates that **J*** increases linearly as θ increases, but decreases as ρ_+ increases. Therefore, since $\rho_+ = 0$ for the IR strategy, the CR and AR strategies perform better than the IR strategy in terms of the MSE of slope criteria. The results for a cuboidal region, shown in Figure H-21 of the Appendix, are similar to those for a spherical region. Again, the CR and AR strategies perform better than the IR strategy, and their performances improve as the magnitude of ρ_+ increases.

Figure 22 on page 153 illustrates the optimal values of **J*** as a function of the **V*/B*** ratio for various values of ρ_+ in both spherical and cuboidal regions of interest. The figure indicates that **J*** decreases rapidly when **V*/B*** < 1, but remains fairly stable when **V*/B*** > 1. Similar to the results in shown in Figure 21, the CR and AR strategies always perform better than the IR strategy in terms of the MSE of slope criteria, and their performances improve as the magnitude of ρ_+ increases.

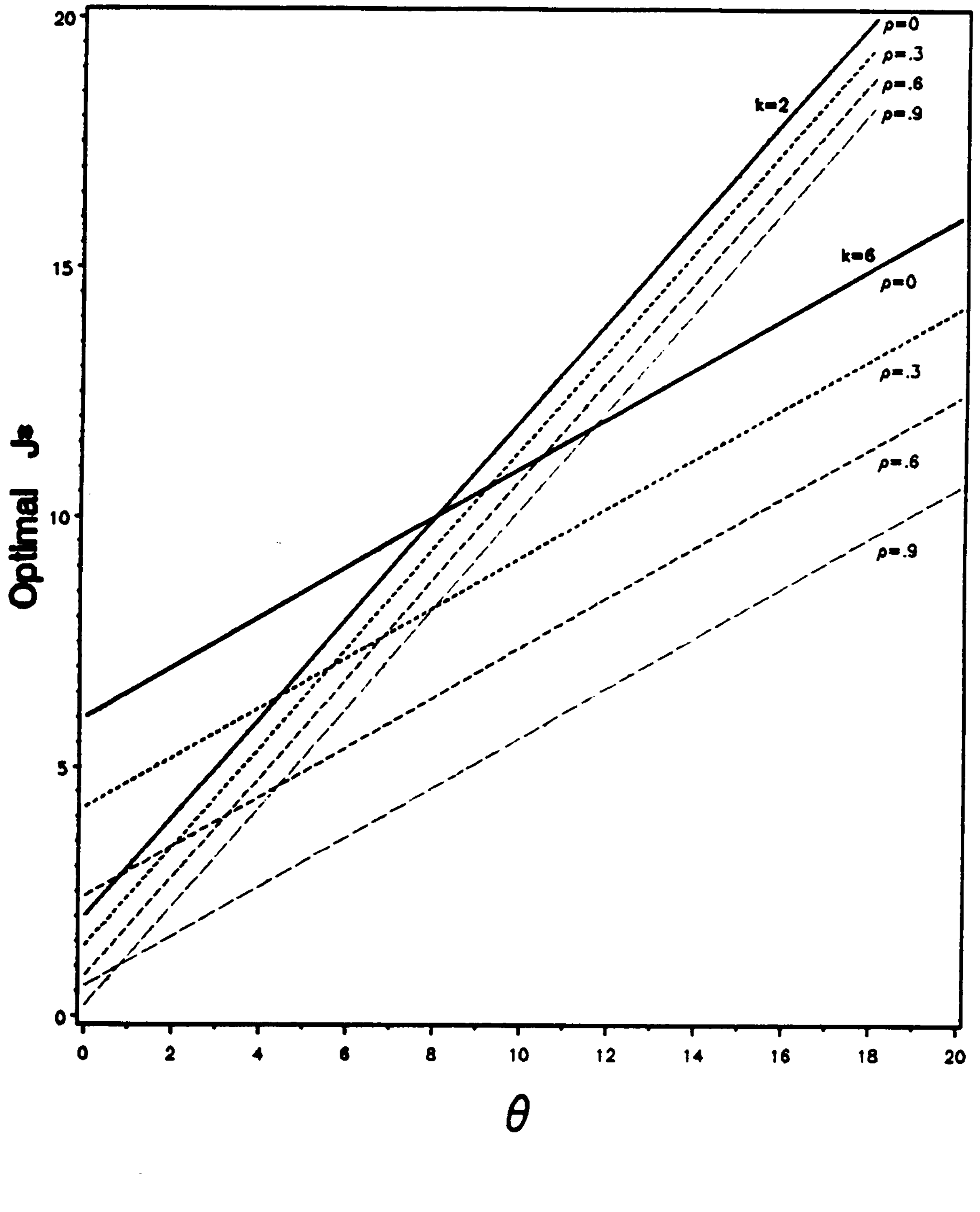


Figure 21. First order optimal values of J^* versus θ.

Region of interest is Spherical.

Two-level factorial design with $k = 2$ and $k = 6$ factors.

Values of $\rho_* = 0, .3, .6, .9$.

The optimal values of J are computed using $[ii] = 1$.

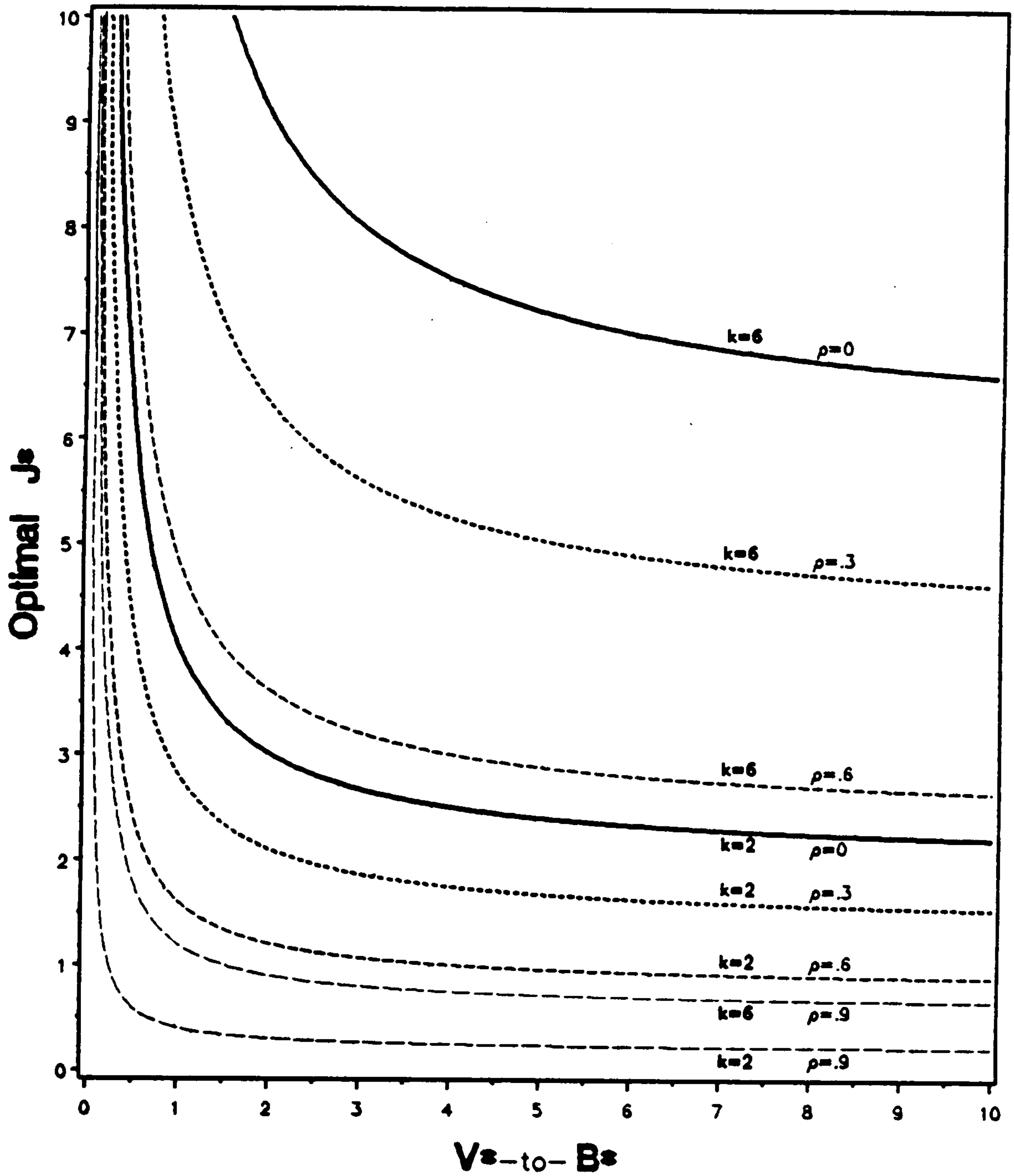


Figure 22. First order optimal values of J* versus V*/B*.

Region of interest is Spherical.
Two-level factorial design with $k = 2$ and $k = 6$ factors.
Values of $\rho_{+} = 0, .3, .6, .9$.
The optimal values of J are computed using $[ii] = 1$.

3.4 Summary of Results in the First Order Case

In the previous two sections, three correlation induction strategies were evaluated in terms of two mean squared error design criteria for the case of fitting a first order model using a two-level factorial design. It was assumed that protection against second order bias was desirable. Section 3.2 evaluated the strategies in terms of the MSE of response criteria and section 3.3 evaluated the strategies in terms of the MSE of slope criteria.

The results for the MSE of response criteria indicate that the size of the **J**-optimal designs (as measured by [ii], the pure second order design moment) are slightly larger than the size of the **B**-optimal designs, and the amount larger decreases as the magnitudes of the induced correlations increase. Therefore, the use of the CR and AR strategies makes it even more important to reduce the size of the design by bringing the design points closer to the center of the experimental region and/or by augmenting the design with center runs.

The AR strategy generally performs better than the IR strategy in terms of the MSE of response criteria when $\mathbf{V}/\mathbf{B} < 3$ (approximately). The range of $\mathbf{V}/\mathbf{B}$ ratios over which the AR strategy is preferable to the IR strategy becomes wider as the magnitudes of ρ_+ and ρ_- increase. Under OLS estimation, the CR strategy performs worse than the IR strategy, but under WLS estimation, the CR strategy performs better than the IR strategy when $\mathbf{V}/\mathbf{B} < 3$.

The results for the MSE of slope criteria indicate that the $\mathbf{J}^*$-optimal design size is equal to the $\mathbf{V}^*$-optimal design size of $[ii] = 1$ because $\mathbf{B}^*$ is not a function of [ii]. Thus, the MSE of slope criteria suggests that the design points be placed at the ± 1 extremes of the coded design region. The CR and AR correlation induction strategies are always preferable to the IR strategy in terms of the MSE of slope criteria, and their performances improve as the magnitude of ρ_+ increases.

The first order research findings indicate that planned correlation induction can affect reductions in the MSE of the predicted responses (or slopes). In particular, the **CR** and **AR** correlation induction strategies are useful techniques during the steepest ascent and ridge analysis stages of RSM optimization studies, when estimation of the response function gradient is important. The **AR** strategy is also useful during the canonical and ridge analysis stages of the optimization process, when the objectives are prediction of the response and estimation of the optimum operating conditions.

The **CR** and **AR** strategies achieve reductions in the MSE of response and MSE of slope through reductions in the variances of the estimated model parameters. In the case of fitting a first order model, the estimated model parameters are the b_0 intercept term and the b_i linear coefficients. For a two-level factorial design, the variances of the k linear coefficients are equivalent, and therefore insight into the variance reductions can be provided through comparisons of $\mathrm{Var}(b_0)$ and $\mathrm{Var}(b_i)$ for the three correlation induction strategies.

An optimal two-level factorial design with $k=3$ factors has been chosen to facilitate comparison of the coefficient variances for each strategy. In order to develop the optimal experimental design plan, the optimal design size (the value of [ii] which minimizes **J** or **J***) must be determined. The optimal value of [ii] is a function of the MSE criteria, the shape of the region of interest, the ratio of variance error to bias error, the ϕ parameter, the number of center runs, and the magnitudes of ρ_+ and ρ_-. Each of these quantities are determined, or selected, by the experimenter. For the 2^3 design under consideration, it is assumed that the experimenter is interested in predicting response, and therefore has decided to use a design which minimizes **J**, the MSE of response. Additionally, the experimenter has selected a spherical region of interest; believes that the error involved with prediction of response consists of about twice as much variance error as bias error ($\mathbf{V}/\mathbf{B}=2$); suspects that the response surface is in an average state of conditioning ($\phi=k/2$); can afford five center runs ($N_c=5$); and has estimated the magnitudes of the induced positive and negative correlations to be $(\rho_+, \rho_-) = (.7, .2)$.

The J-optimal 2^3 design plan which the experimenter uses to fit the first order response model is completely determined by the value of the pure second order design moment. The optimal values of [ii], however, are different for each correlation induction strategy and, for the **CR** and **AR** strategies, the J-optimal [ii] are also different for the OLS and WLS parameter estimation techniques. The optimal values of [ii] are computed as the positive real root of the cubic polynomial shown in equation [3.2.35] on page 114, and are achieved through appropriate choices for the value of the scaling factor (g), computed as

$$g = \sqrt{N\,[ii]\,/\,F} = \sqrt{1.625\,[ii]}$$

for a $k = 3$ two-level factorial design with $N_c = 5$ center runs. The J-optimal values of [ii] and g for this design are shown in the table below.

Table 3. J-Optimal Values of [ii] and g for a 2^3 Design

The design is a $k = 3$ two-level factorial with $Nc = 5$ center runs.
The region of interest is spherical, with $V/B = 2$ and $\phi = k/2$.
The induced correlation magnitudes are: $\rho_+ = .7$ and $\rho_- = .2$.

Correlation Induction Strategy	Least Squares Technique	J-Optimal Design Size ([ii])	Required Scaling Factor (g)
IR	-	.281	.675
CR	OLS	.218	.595
AR	OLS	.230	.611
CR	WLS	.583	.973
AR	WLS	.321	.722

Table 3 indicates that slightly more scaling is required for the **CR** and **AR** strategies than for the **IR** strategy under OLS estimation, but these strategies require less scaling than the **IR** strategy under WLS estimation. For the **IR** strategy, the optimal placement of the design points requires a 33% reduction in the ± 1 design variable levels, and for the **CR** and **AR** strategies under OLS estimation, the levels of the design variables should be reduced about 40%. Under WLS estimation, the CR strategy requires only a 3% reduction in the coded design variable levels, whereas the AR strategy requires a 28% reduction, to coded levels of $g = .72$. Thus, the *size* of the J-optimal designs depend on the correlation strategy and least squares estimation technique chosen by the experimenter. In addition, the variances of the fitted model coefficients (b_0 and b_i) are affected by these strategy alternatives.

Utilizing the $x_1 x_2 x_3$ term as the blocking contrast, the design matrix for the 2^3 design becomes:

$$
\begin{array}{c}
\begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\
D = \left[\begin{array}{ccc}
-1 & -1 & -1 \\
-1 & 1 & 1 \\
1 & -1 & 1 \\
1 & 1 & -1 \\
0 & 0 & 0 \\
\hline
-1 & -1 & 1 \\
-1 & 1 & -1 \\
1 & -1 & -1 \\
1 & 1 & 1 \\
0 & 0 & 0 \\
\hline
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{array} \right] \times g
\end{array}
$$

Random Number Stream Sets

IR Strategy	CR Strategy	AR Strategy
R_1	R_1	R_1
R_2	R_1	R_1
R_3	R_1	R_1
R_4	R_1	R_1
R_5	R_1	R_1
R_6	R_1	$\overline{R}_1$
R_7	R_1	$\overline{R}_1$
R_8	R_1	$\overline{R}_1$
R_9	R_1	$\overline{R}_1$
R_{10}	R_2	$\overline{R}_1$
R_{11}	R_3	R_2
R_{12}	R_4	R_3
R_{13}	R_5	R_4

where g is the scaling factor applied to the design matrix (the J-optimal values were given in Table 3 on page 156), R_i denotes the i^{th} set of pseudorandom number streams, and $\overline{R}_i$ denotes the i^{th} antithetic set of streams. For the IR strategy, a different set of random number streams is used for each design point; for the CR strategy, the same set of streams is assigned to each non-replicated design point, and replicated design points are assigned different stream sets; and for the AR strategy, the design points in the first and second blocks are assigned common and antithetic stream sets, respectively, and replicated center runs (beyond two) are assigned independent stream sets.

The variances of the first order model coefficients for this 2^3 design can be computed for each correlation induction strategy, under OLS and WLS estimation, using the **J**-optimal values of [ii] shown in Table 3. The variance of the intercept term, $\mathrm{Var}(b_0)$, is computed using equations [3.2.36]-[3.2.40] on page 116, and the variances of the linear coefficients, $\mathrm{Var}(b_i)$, are computed using equation [3.2.18] on page 100. The variances of the estimated model coefficients under the **IR**, **CR**, and **AR** strategies are shown in Table 4 on page 159, with the percentage decreases $\downarrow$, or increases $\uparrow$, relative to the **IR** strategy indicated in parentheses. For the **IR** strategy, $\rho_+ = \rho_- = 0$; for the **CR** strategy, induced positive correlations of $\rho_+ = .7$ and $\rho_+ = .3$ are used; and for the **AR** strategy, induced positive and negative correlations of $(\rho_+, \rho_-) = (.7, .2)$ and $(\rho_+, \rho_-) = (.3, .1)$ are considered. In addition to the variances of the b_0 and b_i coefficients, Table 4 presents the sum of the variances of the four coefficient estimators (b_0, b_1, b_2, and b_3); that is, the *trace* of the variance-covariance matrix of the fitted model coefficients.

Table 4 indicates that the variance of the intercept term (b_0) is larger for the **CR** and **AR** strategies than for the **IR** strategy, but the variances of the linear coefficients achieve reductions of approximately $100\,\rho_+$%. The variances of b_0 and b_i are larger under OLS estimation (as compared to WLS estimation) and the increases in the variances of b_0 are smaller for the **AR** strategy than for the **CR** strategy under both OLS and WLS. The largest reduction in the trace of the variance-covariance matrix is achieved by the **CR** strategy under WLS estimation, and smallest reduction is achieved by the **CR** strategy under OLS estimation. However, the **AR** strategy achieves substantial reductions in the trace of this matrix under both OLS and WLS, and the percentage reductions increase as the magnitudes of the induced correlations increase. The **IR** strategy achieves the least variable estimates of the intercept, but the most variable estimates of the k linear coefficients, resulting in the largest "total" coefficient variability of the three correlation induction strategies.

Table 4. Variances of the Fitted Model Coefficients for Optimal k = 3 First Order Designs

The optimal design is a Min-J two-level factorial design with five center runs.
The region of interest is spherical, with V/B = 2 and $\phi = k/2$.
b_0 and b_i are the intercept and linear coefficients of the fitted model.

Correlation Induction Strategy	Variances of the Fitted Model Coefficients	Estimation Technique	
		OLS	WLS
IR	$Var(b_0)/\sigma^2$	.077	.077
	$Var(b_i)/\sigma^2$	.274	.274
	$\{Var(b_0) + 3\,Var(b_i)\}/\sigma^2$	.899	.899
CR	$Var(b_0)/\sigma^2$	.375 (390% ↑)	.186 (140% ↑)
$\rho_+ = .7$	$Var(b_i)/\sigma^2$	.106 (61% ↓)	.040 (86% ↓)
	$\{Var(b_0) + 3\,Var(b_i)\}/\sigma^2$	.693 (23% ↓)	.306 (66% ↓)
AR	$Var(b_0)/\sigma^2$	.183 (140% ↑)	.152 (98% ↑)
$\rho_+ = .7$	$Var(b_i)/\sigma^2$	.100 (63% ↓)	.072 (74% ↓)
$\rho_- = .2$	$\{Var(b_0) + 3\,Var(b_i)\}/\sigma^2$	.483 (46% ↓)	.368 (59% ↓)
CR	$Var(b_0)/\sigma^2$	.205 (160% ↑)	.150 (95% ↑)
$\rho_+ = .3$	$Var(b_i)/\sigma^2$	.219 (20% ↓)	.129 (53% ↓)
	$\{Var(b_0) + 3\,Var(b_i)\}/\sigma^2$	.862 (4% ↓)	.537 (40% ↓)
AR	$Var(b_0)/\sigma^2$	.118 (54% ↑)	.102 (33% ↑)
$\rho_+ = .3$	$Var(b_i)/\sigma^2$	.207 (24% ↓)	.180 (34% ↓)
$\rho_- = .1$	$\{Var(b_0) + 3\,Var(b_i)\}/\sigma^2$	.739 (18% ↓)	.642 (29% ↓)

Under the MSE of slope criteria, the variance of the intercept term (b_0) does not affect the relative performance of the three correlation induction strategies because the intercept is eliminated upon taking partial derivative of the response function. Therefore, the **CR** and **AR** strategies are always preferable to the **IR** strategy in terms of the MSE of slope criteria. For the MSE of response criteria, however, the increased variances of the intercept adversely affect the performances of both the CR and AR strategies. When the magnitudes of the induced correlations are large, the reductions in the variances of the linear coefficients are large, thereby improving the performances of the **CR** and **AR** strategies relative to the **IR** strategy in terms of the MSE of response criteria.

In summary, the first order results presented here suggest that the preferred correlation induction strategy is the assignment rule blocking strategy. The common streams strategy also performs well when estimation of the response function gradient is important and the **IR** strategy also performs well when the fitted response model is adequate (little bias error exists) and prediction of response is important. The next Chapter focuses on the evaluation of the correlation induction strategies for the $d_1 = 2$, $d_2 = 3$ fit-protection situation. A variety of second order response surface designs are used to examine the **IR**, **CR**, and **AR** correlation induction strategies in terms of the MSE of slope and MSE of response criteria.

Chapter 4

SECOND ORDER DESIGNS

Second order response surface designs are experimental plans useful in the estimation of the parameters of second order polynomial models. Relative to first order designs, these designs require a large number of design points but generally yield more precise estimates of the optimum response. RSM optimization studies typically utilize second order designs in conjunction with methods such as canonical and ridge analysis for estimation of the optimum response. The research of this Chapter focuses on the evaluation of the IR, CR, and AR correlation induction strategies using four classes of second order response surface designs.

The properties and specific classes of second order designs were discussed in section 2.2.2 (pages 50-64). The four classes of second order designs, each of which permits orthogonal blocking, include: central composite designs; Box-Behnken designs; three-level factorial designs; and small composite designs. Section 4.1 describes each design class and discusses the ability of these designs to block orthogonally, accommodate center runs, and be fractionally replicated. The moment properties of the designs are also discussed in this section. Section 4.2 evaluates the three correlation induction strategies in terms of the MSE of response criteria, and section 4.3 evaluates the strategies in terms of the MSE of slope criteria. The final section of this Chapter presents a summary of the research findings in the second order case.

4.1 Second Order Design Classes

Second order designs are predominantly used in RSM for estimation of the optimum response *after* the region containing the optimum has been tentatively located. In order to estimate the linear, quadratic, and two-way interaction coefficients of a second order polynomial model, each factor must be present at three or more levels and there must be at least one design point for each estimated model coefficient. Despite the larger size and greater number of factor levels required (when compared to first order designs), the ability of second order designs to fit curvature is particularly useful during the final stage of RSM for obtaining a precise estimate of the optimum response.

In this Chapter, the following four second order design classes are used to evaluate the IR, CR, and AR correlation induction strategies:

1. CCDs — Central composite designs
2. BBDs — Box-Behnken designs
3. FACs — Three-level factorial designs
4. SCDs — Small composite designs of Hartley and Draper.

The first class of second order designs examined in this research is the CCDs, introduced by Box and Wilson (10). These designs consist of a 2^k factorial portion, augmented with $2k$ axial points and a chosen number of center runs, thereby requiring five levels of each factor (± 1, 0, and $\pm\alpha$ levels). The second class of designs utilized in this research is the BBDs, developed by Box and Behnken (4). The BBDs are three-level second order designs formed by combining 2^2 or 2^3 two-level factorial designs with incomplete block designs in a particular manner. The third design class used to evaluate the correlation induction strategies is the FACs, or 3^k designs. These designs utilize all possible combinations of k factors at each of three levels. The final class of second order designs examined in this research is the SCDs, consisting of a factorial portion and the axial portion of a CCD. The factorial portion of the Hartley SCDs (24) are fractional replications of 2^k designs and the factorial portion of the Draper SCDs (14) are incomplete Plackett-Burman designs.

The input variables of a second order response surface design are generally coded for the region of experimentation such that $-1, 0$, and $+1$ represent the low, intermediate, and high levels of each variable, respectively. The levels of the coded design variables are determined from the uncoded, or natural, variables using the transformation shown in equation [2.2.8] on page 34. Depending on the design criteria chosen by the experimenter, the ± 1 high/low levels of the design variables may be scaled up or down through the use of a scaling factor, g. The scaling factor is applied to the entire design matrix (multiplied by each level of each design variable) and its value is function of the shape of the region of interest and the design criteria chosen by the experimenter.

The remainder of this section discusses the properties of the four second order design classes examined in this research. The properties discussed are those relevant to the evaluation of the correlation induction strategies using the MSE design criteria (orthogonal blocking, addition of center runs, fractional replication, and design moment properties).

4.1.1 Orthogonal Blocking in the Second Order Designs

Second order orthogonally blockable designs are those designs which can be partitioned into two or more blocks in such a manner that the block effects do not alter the usual estimates of the linear, quadratic, and two-way interaction coefficients. For a second order design partitioned into two orthogonal blocks, with N_{b1} design points in the first block and N_{b2} design points in the second block, the conditions for orthogonal blocking become

1. Each block itself must be a first order orthogonal design,

$$\sum_{u=1}^{N_{b1}} x_{iu}\, x_{ju} = 0 \quad \text{and} \quad \sum_{u=1}^{N_{b2}} x_{iu}\, x_{ju} = 0 \quad \text{for all} \quad i \neq j = 1, \ldots, k.$$

2. The sum of squares for each design variable within each block must be proportional to the block size,

$$\frac{\sum_{u=1}^{N_{b1}} x_{iu}^{2}}{\sum_{u=1}^{N_{b2}} x_{iu}^{2}} = \frac{N_{b1}}{N_{b2}} \quad \text{for all} \quad i = 1, \ldots\ldots, k.$$

Orthogonal blocking is possible in each of the second order design classes examined here. The CCDs and SCDs can be blocked orthogonally with appropriate choices for the levels of the axial design points and each of the BBDs and FACs permit orthogonal blocking with the exception of the $k=3$ BBD. The *pure* AR correlation induction strategy requires the use of a design which partitions into an even number of orthogonal blocks, such as the CCDs, SCDs, and $k=5, 7$ BBDs. A *modified* AR strategy is required for the FACs and $k=4$ BBD because these designs partition into three orthogonal blocks. The modified *3-block* strategy, suggested by Hussey, Myers, and Houck (29), involves the use of the pure AR strategy in two blocks and the IR strategy in the third block.

The blocking strategies for each of the second order designs examined in this research are shown in Appendix L on pages 363-380. The CCDs and SCDs utilize the factorial portion as the first block and the axial portion as the second block, the blocking schemes for the BBDs are those suggested by Box and Behnken (4), and the FACs are partitioned into three orthogonal blocks using the highest order interaction term (k-way) as the defining contrast. The reader is referred to Montgomery (40) for the techniques required to partition the FACs into three orthogonal blocks.

4.1.2 Addition of Center Runs to Second Order Designs

A center run is a design point taken at the intermediate level of each factor, or more specifically, at the center of the experimental region. In terms of the coded design variables, a center run has $x_i = 0$ for each of the $i = 1, \ldots, k$ design variables. In this research, the number of center runs (N_c) added to the second order designs depends on two design alternatives chosen by the experimenter. Given a design class and the number of design variables, the value of N_c required for an optimal design is a function of the following:

1. Region of interest (spherical or cuboidal)
2. Design criteria (MSE of response or MSE of slope).

Center runs are typically used with the CCDs, BBDs, and SCDs to avoid singularity of the $X'X$ matrix (when each design point lies on a common sphere about the center of the experimental region). Unlike the first order case, the addition of center runs to second order designs affects the ability of the designs to block orthogonally and, in order to maintain orthogonal blocking in the CCDs and SCDs, the levels of the axial design points must be altered. For the BBDs and FACs, an equal number of center runs must be added to each block to maintain the orthogonal blocking property.

In addition to affecting the orthogonal blocking property of second order designs, the addition of center runs affects whether or not the **CR** and **AR** correlation induction strategies are *pure* or *modified*. As noted in section 2.1.4.2, the addition of more than one center run necessitates use of a modified **CR** strategy and the addition of more than two center runs necessitates use of a modified **AR** strategy. The modified **CR** strategy utilizes common streams for each design point and one center run, and this strategy uses independent streams for any additional center runs. The modified **AR** strategy utilizes common streams for each design point and one center run in the 1st block, antithetic streams for each design point and one center run in the 2nd block, and independent streams for any additional center runs. For the designs which partition into three blocks, the modified **AR** strategy utilizes independent streams for each center run and for each design point in the 3rd block. The following notation is used to designate the number of design points in the CCDs, BBDs, FACs, and SCDs examined in this research:

$$\begin{aligned} N &= \text{total number of design points} \\ &= F + n_a + N_c \\ &= N_1 + N_2 \end{aligned}$$

where F = number of factorial points

n_a = number of axial points

($n_a = 2k$ for the CCDs and SCDs)

($n_a = 0$ for the BBDs and FACs)

N_c = number of center points

N_1 = number of points utilizing *common* or *antithetic* random number streams

N_2 = number of points utilizing *independent* random number streams.

Under the **IR** and **CR** correlation induction strategies, the number of design points using independent random number streams (N_2) becomes

IR Strategy: $N_2 = N$

CR Strategy: $N_2 = N_c - 1$ if $N_c > 1$

$= 0$ if $N_c \leq 1$.

Under the **AR** strategy, the number of design points using independent streams depends on the number of blocks into which the design is partitioned. For the designs which partition into two orthogonal blocks (CCDs, SCDs, and $k = 5, 7$ BBDs), the number of design points assigned independent streams becomes

2-block **AR** Strategy: $N_2 = N_c - 2$ if $N_c > 1$

$= 0$ if $N_c \leq 1$.

For the designs which partition into three orthogonal blocks ($k = 4$ BBD and FACs), the number of design points using independent streams becomes

3-block AR Strategy: $N_2 = N_c + \frac{F}{3}$.

The **IR** strategy uses independent random number streams for each design point, and therefore the value of N_2 is equal to the total number of design points. The **CR** strategy uses common streams for each design point and one center run, and independent streams for replications of center runs. Therefore, N_2 is equal to $N_c - 1$ (or zero if $N_c = 0$) under the **CR** strategy. The 2-block **AR** strategy uses common streams for the design points in the 1st block and antithetic streams for the design points in the 2nd block. If the number of center runs is two or more, then one center run is placed in each block. If the number of center runs is one, then the center run is placed in the smaller of the two blocks (making the two blocks closer in size). Therefore, N_2 is equal to $N_c - 2$ (or zero if $N_c = 0$ or 1) under the 2-block **AR** strategy. The 3-block **AR** strategy uses common, antithetic, and independent streams for the design points in the 1st, 2nd, and 3rd blocks, respectively, and independent streams are used for each center run. Therefore, N_2 is equal to $N_c + F/3$ under the 3-block **AR** strategy.

4.1.3 Fractional Replication of Second Order Designs

In the first order case, ½-fractions of the $k = 5, 6$ two-level factorial designs and a ¼-fraction of the $k = 7$ design were used for evaluation of the correlation induction strategies. In the second order case, fractional replication is also useful in reducing the number of experimental runs when the number of design variables is large. However, due to the nature of the second order designs examined in this research, fractional replication is not applicable to all of the design classes. Fractional replication of the CCDs refers to the factorial portions of the designs but not the axial portions, the FACs lend themselves to ⅓ fractional replications, and the BBDs and SCDs cannot be fractionally replicated.

The factorial portion of a CCD is a 2^k design, and therefore the methods used to fractionally replicate the 2^k designs (described on page 88) are applicable to the CCDs. For the FACs, the procedures for developing fractional replicates are similar to those used in the development of orthogonal blocking arrangements. A defining contrast is used to partition the 3^k design into three orthogonal blocks in such a manner that the linear, quadratic, and two-way interaction terms are not aliased with each other. The use of a three-way (or higher) interaction term is sufficient to prevent aliasing of the second order model coefficients. Once a 3^k design has been partitioned into three orthogonal blocks, use of the design points in *one* of the three blocks represents a $^1/_3$ fractional replication of the design.

This Chapter presents an evaluation of the **IR**, **CR**, and **AR** correlation induction strategies using the following full and fractionally replicated second order designs:

1. Central Composite Designs (CCDs)
   - a) Full factorial replications $k = 2, 3, 4, 5$ factors
   - b) ½-fractional factorials $k = 6, 7$ factors
2. Box-Behnken Designs (BBDs) $k = 4, 5, 7$ factors
3. Three-level Factorial Designs (FACs)
   - a) Full factorial replications $k = 3, 4, 5$ factors
   - b) $^1/_3$-fractional factorials $k = 6, 7$ factors
4. Small Composite Designs (SCDs)
   - a) Hartley Designs $k = 3, 4, 6$ factors
   - b) Draper Designs $k = 5, 7$ factors.

The $k=3$ BBD and $k=2$ FAC are not considered in this research because these designs do not block orthogonally, and the $k=6$ BBD is not examined because [iijj] is not constant for all pairs of design variables. The $k=5, 7$ Hartley designs are not pursued because the Draper SCDs are closer to saturation (require fewer design points).

4.1.4 Design Moment Properties

The MSE of response and MSE of slope criteria, which are used for evaluation of the correlation induction strategies, take both variance error and bias error into account. Variance error represents the variability in the predicted responses (or slopes) when the fitted polynomial model is the *correct* model. Bias error represents the difference between the true and predicted responses (or slopes) when the fitted model is *incorrect*. In this research, bias error is assumed to result from the use of a polynomial model which is under-specified by one degree. In the second order case, the fitted model includes linear, quadratic, and two-way interaction terms, and the protection model additionally includes cubic and third order interaction terms. Thus, the fit-protection situation considered in this Chapter is $d_1 = 2$, $d_2 = 3$.

The design moment matrices (M_{11}, M_{21}, and M_{22}) and the region moment matrices (μ_{11}, μ_{21}, and μ_{22}) are useful in formulating the variance and bias components of the MSE design criteria. The design and region moment matrices are defined as

$$M_{11} = N^{-1}\ X'_1 X_1 \qquad\qquad \mu_{11} = \Omega_r \int_R \mathbf{x}_1'\ \mathbf{x}_1\, d\mathbf{x}$$

$$M_{21} = N^{-1}\ X'_2 X_1 \qquad\qquad \mu_{21} = \Omega_r \int_R \mathbf{x}_2'\ \mathbf{x}_1\, d\mathbf{x}$$

$$M_{22} = N^{-1}\ X'_2 X_2 \qquad\qquad \mu_{22} = \Omega_r \int_R \mathbf{x}_2'\ \mathbf{x}_2\, d\mathbf{x}$$

where the $\mathbf{x}_1$ vector contains the p_1 regressor terms in the X_1 matrix and the $\mathbf{x}_2$ vector contains the p_2 regressor terms in the X_2 matrix. For the $d_1 = 2$, $d_2 = 3$ fit-protection situation, the X_1 and X_2 matrices are shown on pages 45-46 and the M_{11}, M_{21}, and M_{22} design moment matrices are shown in Appendix J (pages 348-352). The μ_{11}, μ_{21}, and μ_{22} matrices are identical in form to the design moment matrices, except that the elements are region moments rather than design moments (for example, [ii] is replaced with w_{ii}).

For the case of a second order fitted model with protection against a true third order model, the M_{11} and μ_{11} matrices contain moments through order $2d_1 = 4$, the M_{21} and μ_{21} matrices contain moments through order $d_1 + d_2 = 5$, and the M_{22} and μ_{22} matrices contain moments through order $2d_2 = 6$. The variance error is a function of μ_{11} and M_{11}^{-1}, and the bias error is additionally a function of μ_{21}, μ_{22}, and M_{12}. Therefore, the mathematical developments of **B** and **V** are simplified when these moment matrices are sparse. Box and Draper (6) and Draper and Lawrence (6) utilize three design moment conditions (to ensure rotatability) in their developments of **B** and **V**. This research utilizes two of the three rotatability conditions, as follows:

1. *Odd* design moments through order five are equal to zero,

   $0 = [i] = [ij] = [iij] = [ijk] = [iiij] = [iijk] = [iiiij] = \ldots = [ijklm]$ for all i, j, k, l, m.

2. *Even* design moments through order four are the same for all pairs of design variables,

   $[ii] = [jj]$ and $[iijj] = [iikk]$ for all i, j, k.

Box and Draper additionally require that $[iiii]/[iijj] = 3$; that is, the authors restrict their developments to spherically-rotatable designs, and Draper and Lawrence require that $[iiii]/[iijj] = {}^9/_5$, thereby restricting their work to cuboidally-rotatable designs. This research places no restriction on the value of $[iiii]/[iijj]$, and therefore does not require the use of a rotatable design. However, utilization of the first two rotatability conditions enables common forms of the M_{11}^{-1} and M_{21} matrices to be written for the CCDs, BBDs, and FACs, thereby simplifying the mathematical developments of **B** and **V**. These conditions also enable the equation for **B** to be written in terms of three unknown third order parameters regardless of the value of k. With the exception of the SCDs, each of the second order design classes examined in this research adhere to these moment conditions. The SCDs meet the second condition, but do not necessarily have odd design moments through order five equal to zero. The following design moments are not necessarily equal to zero for the SCDs examined in this research:

1. $k = 3, 4, 6$ Hartley SCDs — $[ijk]$ and $[iiijk]$
2. $k = 5$ Draper SCD — $[ijk]$, $[iiijk]$, and $[iijkl]$
3. $k = 7$ Draper SCD — $[ij]$, $[ijk]$, $[iiijk]$, $[iijkl]$, and $[ijklm]$.

Because the odd design moments of the SCDs are not all equal to zero, the mathematical developments of this Chapter can only be used as approximations for the SCDs. Exact computation of **V** and **B** for the SCDs requires matrix algebra on the specific design matrices.

The remainder of the research in this Chapter is an evaluation of the correlation strategies in the second order case. The mathematical developments pertain to the CCDs, BBDs, and FACs design classes but, in general, do not hold for the SCDs.

4.2 MSE of Response Criteria in the Second Order Case

The focus of this and the next section is the evaluation of three correlation induction strategies for the situation in which an experimenter is fitting a second order polynomial model using a second order response surface design. The three correlation induction strategies are the **IR**, **CR**, and **AR** strategies (discussed on pages 26-30) and the second designs used for evaluation of the strategies are the CCDs, BBDs, FACs, and SCDs (described on pages 55-64). The MSE of response criteria is used to evaluate the correlation induction strategies in this section, and the MSE of slope serves as the performance criteria in the following section.

The MSE of response criteria (discussed on pages 68-73) calls for minimizing **J**, the average, normalized, mean squared error of the predicted responses. The optimal designs in the first order case were Min-**J** designs, achieved through an appropriate choice for the value of [ii]. In the second order case, **J** is a function of [ii], [iiii], and [iijj], and therefore **J** cannot be minimized with respect to [ii] alone. The optimal second order designs evaluated in this research are Min-**V** | Min-**B** designs rather than Min-**J** designs. These designs utilize the value of [iiii]/[ii] which minimizes **B**, achieved through an appropriate choice for the value of the scaling factor, g. Additionally, these designs utilize the value of [ii] which minimizes **V** (given the Min-**B** value of [iiii]/[ii]), achieved through an appropriate choice for the number of center runs, N_c. Because the optimal number of center runs must be rounded to the nearest integer value, the optimal second order designs are only *near* Min-**V** | Min-**B** designs.

Here it is assumed that a second order response surface design is being used to model the relationship between a set of input variables and a response variable. The equation which the experimenter uses to fit the data is a second order polynomial regression model ($d_1 = 2$), and the equation which the experimenter desires protection against being the true model is a cubic polynomial model ($d_2 = 3$). The equations for the predicted and true response variables at the u^{th} setting of the factors, as shown in equations [2.2.30] and [2.2.31] on pages 69-70, respectively, are

$$\hat{y}_{(\underline{x}_u)} = \underline{x}_{1u}' \underline{b}_1$$

$$y_{(\underline{x}_u)} = \underline{x}_{1u}' \underline{\beta}_1 + \underline{x}_{2u}' \underline{\beta}_2 + \varepsilon_u$$

where the subscript $_1$ indicates that the vectors contain terms from the fitted second order model, and the subscript $_2$ indicates that the vectors contain the cubic and third order interaction terms in the protection model. The number of fitted model coefficients is $p_1 = \frac{1}{2}(k+1)(k+2)$, of which $k+1$ are linear, k are quadratic, and $\binom{k}{2}$ are two-way interaction coefficients. The number of unfitted coefficients in the protection model is $p_2 = \frac{1}{6}k(k+1)(k+2)$, of which k are cubic, $k(k-1)$ are linear × quadratic interactions, and $\binom{k}{3}$ are three-way interaction coefficients.

Sections 4.2.1 and 4.2.2 present the bias and variance components of the mean squared error of $\hat{y}_{(\underline{x}_u)}$ for the second order designs examined here.

4.2.1 Bias Component of **J** in the Second Order Case

The bias component of the MSE of response, **B**, is computed as the average squared bias of the predicted responses, normalized with respect to the number of design points and the experimental error variance. Box and Draper (6) present a mathematical development of **B** in a spherical region of interest, and Draper and Lawrence (16) derive the equation for **B** in a cuboidal region of interest. For a spherical region, Box and Draper require that the designs be rotatable ($[iiii]/[iijj] = 3$), and for a cuboidal region Draper and Lawrence require that the designs be cuboidally-rotatable ($[iiii]/[iijj] = 9/5$). These authors also restrict their developments to designs with odd design moments through order five equal to zero, thereby fixing the number of unknown third order parameters at two (regardless of the value of k) for spherically and cuboidally rotatable designs.

The research presented herein extends the previous developments of Box, Draper, and Lawrence by allowing for the use of second order designs which are not rotatable. In this research, no restriction is placed on the value of $[iiii]/[iijj]$, but similar to the previous work, $[iiii]$ and $[iijj]$ are assumed to be constant and odd design moments through order five are assumed to be equal to zero (the two design moment conditions presented on page 170). For the non-rotatable second order designs examined in this research, the number of unknown third order parameters is fixed at three.

For practical and computational reasons, the developments presented in this Chapter are restricted to OLS estimation of the model parameters. For the **IR** and **CR** correlation induction strategies, the equations for **B** are the same under OLS and WLS estimation. For the **AR** strategy, however, the equations for **B** are the same under WLS and OLS only for designs which partition into two orthogonal blocks (CCDs and $k = 5, 7$ BBDs).

Box and Draper (5) define the bias error (as shown in equation [3.2.1] on page 92) as

$$\mathbf{B} = \frac{N\Omega_r}{\sigma^2} \int_R \text{Bias}^2 [\hat{y}_{(\underline{x}_u)}] \, d\underline{x}$$

$$= \frac{N}{\sigma^2} \underline{\beta}'_2 \{ A' \mu_{11} A - 2\mu_{21} A + \mu_{22} \} \underline{\beta}_2$$

where Ω_r is the inverse of the volume of the centered and scaled region of interest (defined on page 71), A is the alias matrix (defined on page 43), and the μ terms are the region moment matrices of the design (defined on page 71).

The equation for **B** indicates that the bias error is a function of the unknown β parameters in the protection model. Therefore, the bias error cannot be computed using the above equation unless the β coefficients of the unfitted cubic and interaction terms can be estimated. However, for the CCDs, BBDs, and FACs, which have the two design moment properties shown on page 170, **B** can be simplified to a scalar equation involving only three unknown third order parameters.

The $A'\mu_{11} A$, $2\mu_{21} A$, and μ_{22} component matrices of **B** must be formulated in order to derive the scalar equation for **B**. Similar to the notation of Box and Draper (6), the following two ratios are introduced for this purpose:

$$\theta = \frac{[iiii]}{[ii]} \qquad [4.2.1]$$

$$r = \frac{[iiii]}{[iijj]} \qquad [4.2.2]$$

The $A'\mu_{11} A$, $2\mu_{21} A$, and μ_{22} matrices are shown in Appendix O (pages 396-407). Upon summing these matrices, the **B** component of **J** for the second order designs examined in this research becomes

$$\mathbf{B} = \frac{N}{\sigma^2}\,\underline{\beta}_2' \left[\begin{array}{c|c} G_B & \mathbf{0} \\ \hline \mathbf{0} & w_{iijjkk}\,\mathbf{I}_{\binom{k}{3}} \end{array}\right] \underline{\beta}_2 \qquad [4.2.3]$$

where the $\mathbf{0}$'s indicate null matrices and G_B is a $(k^2 \times k^2)$ block diagonal matrix, with k identical blocks of the $(k \times k)$ H_B matrix

$$H_B = \left[\begin{array}{c|c} h_1 & h_2\,\mathbf{1}'_{k-1} \\ \hline h_3\,\mathbf{1}_{k-1} & (h_4 - h_5)\,\mathbf{I}_{k-1} + h_5\,\mathbf{1}_{k-1}\mathbf{1}_{k-1}' \end{array}\right]$$

where $\mathbf{1}_{k-1}$ is a $k-1$ column vector of ones, $\mathbf{I}_{k-1}$ is a $(k-1)\times(k-1)$ identity matrix, and the h_i terms are scalar quantities defined as

$$\begin{aligned}
h_1 &= \theta^2 w_{ii} - 2\theta\, w_{iiii} + w_{iiiiii} \\
h_2 &= \theta^2 w_{ii}/r - 2\theta\, w_{iiii}/r + w_{iiiijj} \\
h_3 &= \theta^2 w_{ii}/r - 2\theta\, w_{iijj} + w_{iiiijj} \\
h_4 &= \theta^2 w_{ii}/r^2 - 2\theta\, w_{iijj}/r + w_{iiiiii} \\
h_5 &= \theta^2 w_{ii}/r - 2\theta\, w_{iijj}/r^2 + w_{iiiijj} \,.
\end{aligned}$$

Upon pre- and post-multiplication of the matrix in equation [4.2.3] by $\underline{\beta}_2 \sqrt{N/\sigma^2}$, the **B** component of **J** can be written in equation form as

$$\mathbf{B} = \frac{N}{\sigma^2}\left[\{h_1+(k-1)h_3\}\sum_{i=1}^{k}\beta_{iii}^2 + \{h_2+h_4+(k-2)h_5\}\sum_{i\neq}\sum_{j}\beta_{ijj}^2 + w_{iijjkk}\sum_{i\neq}\sum_{j\neq}\sum_{k}\beta_{ijk}^2\right]$$

$$= \phi_1\Theta_1 + \phi_2\Theta_2 + \phi_3\Theta_3 \qquad [4.2.4]$$

where $\phi_1 = h_1 + (k-1)h_3$ $\qquad \Theta_1 = \frac{N}{\sigma^2}\sum_{i=1}^{k}\beta_{iii}^2$

$\phi_2 = h_2 + h_4 + (k-2)h_5$ $\qquad \Theta_2 = \frac{N}{\sigma^2}\sum_{i\neq}\sum_{j}\beta_{ijj}^2$

$\phi_3 = w_{iijjkk}$ $\qquad \Theta_3 = \frac{N}{\sigma^2}\sum_{i\neq}\sum_{j\neq}\sum_{k}\beta_{ijk}^2$.

Equation [4.2.4] indicates that $\mathbf{B}$ is a function of the following parameters and constants:

$$\begin{aligned}\mathbf{B} &= f(\phi_1, \phi_2, \phi_3, \Theta_1, \Theta_2, \Theta_3)\\ &= f(h_1, h_2, h_3, h_4, h_5, k, w_{iijjkk}, \Theta_1, \Theta_2, \Theta_3)\\ &= f(\theta, r, k, w_{ii}, w_{iiii}, w_{iijj}, w_{iiiiii}, w_{iiiijj}, w_{iijjkk}, \Theta_1, \Theta_2, \Theta_3)\ .\end{aligned}$$

The values of r and k are fixed for a given design plan, the $w_{ii} \ldots w_{iijjkk}$ region moments are fixed for a given region shape, and Θ_1, Θ_2, and Θ_3 are unknown parameters representing standardized measures of cubic curvature in the response surface. Therefore, θ is the only term in the equation for $\mathbf{B}$ which can be altered to minimize $\mathbf{B}$. The value of θ which results in a Min-$\mathbf{B}$ design is determined by setting the partial derivative of $\mathbf{B}$ with respect to θ equal to zero, yielding

$$\frac{\partial \mathbf{B}}{\partial \theta} = \frac{\partial}{\partial \theta}\left[\frac{w_{ii}\,\theta^2\,(r+k-1)\,\Theta_1}{r} - 2\theta\,(w_{iiii} + (k-1)\,w_{iijj})\,\Theta_1 + (w_{iiiiii} + (k-1)\,w_{iiiijj})\,\Theta_1\right]$$

$$+ \frac{\partial}{\partial \theta}\left[\frac{w_{ii}\,\theta^2\,(r+k-1)\,\Theta_2}{r^2} - \frac{2\theta\,(w_{iiii} + (k-1)\,w_{iijj})\,\Theta_2}{r} + (w_{iiiiii} + (k-1)\,w_{iiiijj})\,\Theta_2\right]$$

$$+ \frac{\partial}{\partial \theta}\left[w_{iijjkk}\,\Theta_3\right] = 0$$

[4.2.5]

$$0 = \frac{2\,w_{ii}\,\theta\,(r+k-1)}{r}\left[\Theta_1 + \frac{\Theta_2}{r}\right] - 2\,(w_{iiii} + (k-1)\,w_{iijj})\left[\Theta_1 + \frac{\Theta_2}{r}\right].$$

Solving equation [4.2.5] for θ, the Min-**B** optimal value of θ becomes

$$\theta = \frac{r\,(w_{iiii} + (k-1)\,w_{iijj})}{w_{ii}\,(r+k-1)} \qquad [4.2.6]$$

The second partial derivative of **B** with respect to θ is a positive quantity and, therefore, the value of θ in equation [4.2.6] minimizes **B**. The Min-**B** value of θ can be obtained with an appropriate choice of the scaling factor, g. For the CCDs, BBDs, and FACs, the Min-**B** value of g (as shown in Appendix P on pages 411-412) becomes

$$g = \sqrt{\frac{\theta\,(F + 2\alpha^2)}{(F + 2\alpha^4)}} \qquad \text{for the CCDs}$$

[4.2.7]

$$g = \sqrt{\theta} \qquad \text{for the BBDs and FACs}$$

where θ is the value of [iiii]/[ii] which minimizes **B**, F is the number of factorial design points, and α is the level of the axial design points. Equation [4.2.7] indicates that the Min-**B** values of g for the CCDs are a function of α. In this research, the CCDs utilize the value of α which permits partitioning of the designs into two orthogonal blocks (one center run in each block). The use of more than two center runs does not change the value of α required for orthogonal blocking because the additional center runs use independent random number streams (rather than common or antithetic streams). Therefore, the blocking remains orthogonal for the purposes of the AR strategy. The value of α which permits orthogonal blocking in the CCDs becomes

$$\alpha = \sqrt{\frac{F(2k+1)}{2(F+1)}} \quad . \qquad [4.2.8]$$

The α level needed for orthogonal blocking in the SCDs depends on the number of center runs in the design. Unlike the CCDs, some of the optimal SCDs require only one center run, and therefore the value of α cannot be computed using equation [4.2.8]. The single center run of these designs is placed in the smaller of the two blocks in order to make the block sizes similar. The single center run of the $k=3$ SCD is placed in the factorial block and the value of α which permits orthogonal blocking becomes $\alpha = \sqrt{kF/(F+1)}$. The single center run of the $k=4, \ldots, 7$ SCDs is placed in the axial block and the required α level becomes $\alpha = \sqrt{k + ½}$. Table 5 indicates the values of α needed for orthogonal blocking in the CCDs and SCDs examined in this research. Each of the CCDs utilizes two or more center runs, but the SCDs utilize one or more center run and therefore two different α levels are indicated for the SCDs.

Table 5. Levels of the Axial Design Points in the CCDs and SCDs. Values of α required for orthogonal blocking.

DESIGN CLASS	Values of α in the CCDs					
	k=2	k=3	k=4	k=5	k=6	k=7
CCD ($N_c \geq 2$)	1.414	1.764	2.058	2.309	2.511	2.717
SCD ($N_c = 1$)	-	1.549	2.121	2.345	2.549	2.739
SCD ($N_c \geq 2$)	-	1.673	2.000	2.253	2.473	2.691

The Min-B values of g for the second order designs examined in this research are shown in Tables 6 and 7 for spherical and cuboidal regions of interest, respectively. (More precise values of g can be found in Appendix P on page 419.) Figure 23 on page 181 graphically displays the Min-B values of the scaling factor as a function of the number of factors in the model, k. The figure indicates that the levels of the factorial design points should be reduced from the ± 1 extremes which are typically used in practice, with a spherical region requiring more of a scale reduction than a cuboidal region. The Min-B values of the scaling factor appear to be similar for the CCDs and SCDs, slightly larger for the FACs, and the largest for the BBDs.

Table 6. Min-B Values of the Scaling Factor in a Spherical Region of Interest. Optimal values of g for the CCDs, BBDs, FACs, and SCDs.

DESIGN CLASS	Min-B Values of g					
	k=2	k=3	k=4	k=5	k=6	k=7
CCD	.58	.48	.42	.39	.36	.34
BBD	-	-	.61	.62	-	.52
FAC	-	.55	.50	.46	.43	.40
SCD	-	.52	.42	.39	.36	.34

Table 7. Min-B Values of the Scaling Factor in a Cuboidal Region of Interest. Optimal values of g for the CCDs, BBDs, FACs, and SCDs.

DESIGN CLASS	Min-B Values of g					
	k=2	k=3	k=4	k=5	k=6	k=7
CCD	.68	.64	.63	.62	.61	.61
BBD	-	-	.89	.98	-	.93
FAC	-	.74	.73	.73	.72	.72
SCD	-	.66	.63	.62	.61	.61

The bias component of **J** derived in this section applies to the CCD, BBD, and FAC design classes. Because the SCDs do not necessarily have odd design moments through order five equal to zero, the Min-**B** values of g for the SCDs are only approximations. The **B**-optimal second order designs presented in this section are the same for the **IR**, **CR**, and **AR** strategies because the equations for **B** are independent of the correlation induction strategy under ordinary least squares estimation.

The **V** component of **J** is derived for the second order designs in the next section. This component is a function of the correlation induction strategy and, therefore, separate forms of **V** are presented for the **IR**, **CR**, and **AR** strategies.

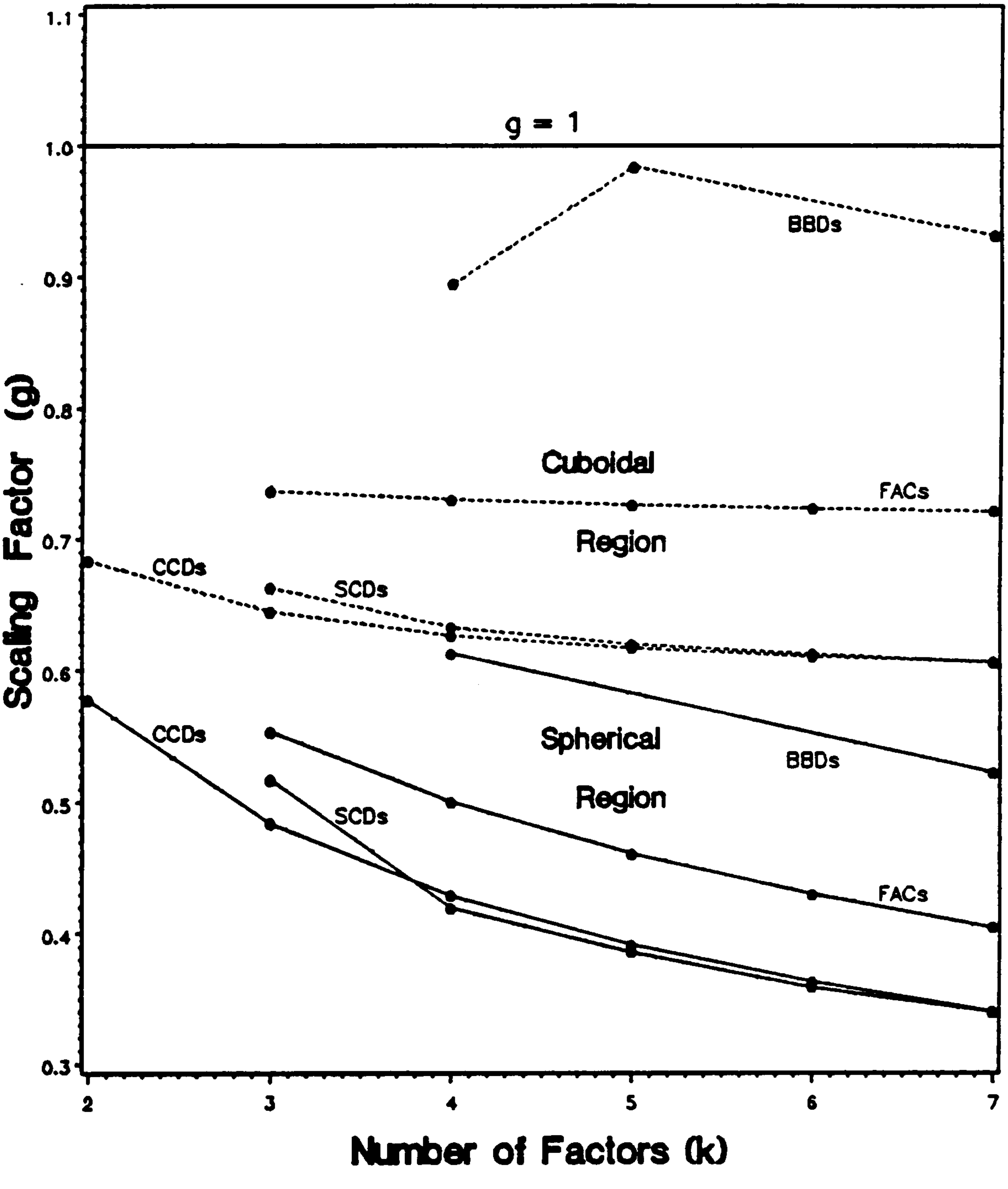


Figure 23. Optimal Values of the Scaling Factor for Min-B Second Order Designs.
The horizontal line at $g = 1$ indicates the value of the scaling factor which is typically used in practice.
Dotted lines connect the values of g for a Cuboidal region of interest.
Solid lines connect the values of g for a Spherical region of interest.

4.2.2 Variance Component of J in the Second Order Case

The variance component of the MSE of response, **V**, is computed as the average variance of the predicted responses, normalized with respect to N and σ^2. Box and Draper (5) present a mathematical development of **V** in a spherical region, and Draper and Lawrence (16) extend their work to a cuboidal region. These authors assume uncorrelated responses and restrict their developments to rotatable and cuboidally-rotatable designs. Hussey, Myers, and Houck (29) expand on the previous work by allowing for correlated responses and the use of non-rotatable designs. These authors develop the variance component for the **IR**, pure **CR**, and pure **AR** correlation induction strategies for the CCD and BBD design classes. This research extends the work of Hussey et. al. in the following ways:

1. Allowing for the use of replicated center runs.
2. Permitting the factorial design points to be placed at levels other than ± 1.
3. Extending the developments to the FAC design class.
4. Extending the evaluations to the SCD design class.
5. Constructing optimal second order design plans.

Box and Draper (5) define the variance error (as shown in equation [3.2.15] on page 98) as

$$\mathbf{V} = \frac{N\Omega_T}{\sigma^2} \int_R \mathrm{Var}\,[\,\hat{y}_{(\underline{x}_u)}\,]\, d\underline{x}$$

$$= \frac{N}{\sigma_2}\ \mathrm{Trace}\,\{\,\mathrm{Var}\,[\,\underline{b}_1\,]\ \mu_{11}\}\ .$$

where Ω_r is the inverse of the volume of the centered and scaled region of interest (defined on page 71), $\mathrm{Var}\,[\,\underline{b}_1]$ is the variance-covariance matrix of the coefficient estimators (defined on page 72) and μ_{11} is a region moment matrix of the design (defined on page 71).

The variance error, **V**, depends on the shape of the region of interest through μ_{11}. For the second order designs considered here, the $(p_1 \times p_1)$ μ_{11} matrix becomes

$$\mu_{11} = \left[\begin{array}{c|c|c|c} 1 & \mathbf{0} & w_{ii}\,\mathbf{1}_k' & \mathbf{0} \\ \hline \mathbf{0} & w_{ii}\,\mathbf{I}_k & \mathbf{0} & \mathbf{0} \\ \hline w_{ii}\,\mathbf{1}_k & \mathbf{0} & \begin{array}{c} w_{iiii}\,\mathbf{I}_k \\ +\,w_{iijj}\,\mathbf{1}_k\mathbf{1}_k' \\ -\,w_{iijj}\,\mathbf{I}_k \end{array} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & w_{iijj}\,\mathbf{I}_{\binom{k}{2}} \end{array}\right]$$

where $\mathbf{1}$ is a column vector of ones, $\mathbf{I}$ is an identity matrix, $\mathbf{0}$ is a null matrix, and the w_{ii}, w_{iiii}, and w_{iijj} terms are region moments of the design, defined on pages 289-291 for spherical and cuboidal regions of interest.

The variance error is a function of the correlation induction strategy through the variance-covariance matrix of $\underline{b}_1$. Under ordinary least squares estimation, the $(p_1 \times p_1)$ $\mathrm{Var}[\,\underline{b}_1\,]$ matrix (as shown in equation [2.2.41] on page 72) becomes

$$\mathrm{Var}[\,\underline{b}_1] = (X'_1\,X_1)^{-1}\,X'_1\;\; V\,X_1\,(X'_1\,X_1)^{-1}\,\sigma^2$$

where V is the $(N \times N)$ variance-covariance matrix of the response (shown in equations [2.1.11], [2.1.12], and [2.1.13] on pages 28-29, for the **IR**, **CR**, and **AR** correlation induction strategies, respectively) and X_1 is the $(N \times p_1)$ matrix of fitted model terms. For the CCD, BBD, and FAC design classes, the $(X'_1\,X_1)^{-1}$ matrix can be written as

$$(X'_1 X_1)^{-1} = \left[\begin{array}{c|c|c|c} \dfrac{a}{N D} & 0 & \dfrac{-[ii]}{N D}\, \mathbf{1}_k' & 0 \\ \hline 0 & \dfrac{1}{N[ii]}\, \mathbf{I}_k & 0 & 0 \\ \hline \dfrac{-[ii]}{N D}\, \mathbf{1}_k & 0 & \begin{array}{c} \dfrac{(b-c)}{N D_2}\, \mathbf{I}_k \\ + \dfrac{c}{N D_2}\, \mathbf{1}_k \mathbf{1}_k' \end{array} & 0 \\ \hline 0 & 0 & 0 & \dfrac{1}{N[iijj]}\, \mathbf{I}_{\binom{k}{2}} \end{array}\right]$$

where
$$a = \frac{\theta\,[ii]\,(r+k-1)}{r}$$

$$b = \frac{\theta\,[ii]\,(r+k-2)}{r} - (k-1)\,[ii]^2$$

$$c = [ii]\left[\frac{r\,[ii]-\theta}{r}\right]$$

$$D = a - k\,[ii]^2$$

$$D_2 = \frac{D\,\theta\,[ii]\,(r-1)}{r}.$$

For the second order designs under consideration, $X'_1 X_1$ would be a *singular* matrix if the value of D were equal to zero, or equivalently, if the value of $[iijj]$ were equal to

$$[iijj] = \frac{[ii]}{k-1}\,(k\,[ii] - \theta)\ .$$

Each of the optimal second order designs are checked for singularity of the $X'_1 X_1$ matrix by comparing the actual values of [iijj] to the singular values. (See the computer output in Appendices Q and R for verification of non-singular $X'_1 X_1$ matrices.)

Before presenting the equations for **V** under the **IR**, **CR**, and **AR** correlation induction strategies, it is useful to define the following notation:

N_1 = number of design points generated with *common* or *antithetic* streams

N_{1a} = number of design points generated with *common* streams (block 1)

N_{1b} = number of design points generated with *antithetic* streams (block 2)

N_2 = number of design points generated with *independent* streams

$\Delta = N_{1a}\,[ii]_{1a} - N_{1b}\,[ii]_{1b}$ = Block 1 sum of squares $-$ Block 2 sum of squares

$$e = \frac{(1-h)\,b}{N D_2} + \frac{h\,N_c\,[ii]^2}{(N D)^2}$$

$$f = \frac{N_c\,[ii]^2}{(N D)^2} - \frac{4}{(N\,[ii])^2}$$

$$g = w_{iiii} + (k-1)\,w_{iijj}$$

$$h = \begin{cases} 2\,(\text{number of blocks})^{-1} & \text{for the AR strategy} \\ 1 & \text{for the CR strategy} \end{cases}$$

$$m = \frac{(N_{1b}-N_{1a})\,[ii] + \Delta}{N D}$$

$$p = \frac{(N_{1a}-N_{1b})\,a - k\,[ii]\,\Delta}{N D}$$

$$q = \frac{\sqrt{N_2}\;a}{N D}$$

$$r = -\frac{\sqrt{N_2}\;[ii]}{N D}$$

$$s = \frac{(N-N_2)\,a - h\,k\,N\,[ii]^2}{N D}$$

$$t = \frac{N_2\,[ii] + (h-1)\,N\,[ii]}{N D}.$$

The **V** component of **J** is obtained by taking the *trace* of the product of μ_{11} and Var $[\underline{b}_1]$. The component parts of the Var $[\underline{b}_1]$ matrices are shown in Appendix M (pages 381-388) and the equations for **V** under the **IR**, **CR**, and **AR** correlation induction strategies are derived in Appendix

N (pages 389-395). Only the results of the derivations in the Appendix are presented here. Under the **IR** strategy, V is equal to the identity matrix, and the **V** component of **J** for the CCDs, BBDs, and FACs under the **IR** strategy becomes

$$\mathbf{V}_{IR} = N \;\; \text{Trace}\,\{ (X'_1 X_1)^{-1} \; \mu_{11} \}$$

$$= \frac{a - 2k\,w_{ii}\,[ii]}{D} + \frac{k\,w_{ii}}{[ii]} + \frac{b\,k\,w_{iiii} + c\,k\,(k-1)\,w_{iijj}}{D_2} + \frac{k\,(k-1)\,\theta\,[ii]\,w_{iijj}}{2r} \quad [4.2.9]$$

Under the **CR** strategy, two forms of the variance error are relevant. For the designs utilizing less than two center runs, the equation for **V** under the *pure* **CR** strategy becomes

$$\mathbf{V}_{pure\,CR} = (1-\rho_+)\,\mathbf{V}_{IR} + N\,\rho_+ \quad [4.2.10]$$

and for the designs utilizing two or more center runs, the equation for **V** under the *modified* **CR** strategy becomes

$$\mathbf{V}_{mod.\,CR} = (1-\rho_+)\,\mathbf{V}_{IR} + N\,\rho_+\,(s^2 + 2k\,w_{ii}\,s\,t + k\,t^2\,g)$$
$$+ N\,\rho_+\,(q^2 + 2k\,w_{ii}\,r\,q + k\,r^2\,g)\;. \quad [4.2.11]$$

Under the **AR** strategy, four forms of the variance error are relevant. For the designs which partition into two orthogonal blocks and utilize less than three center runs, the equation for **V** under the *pure* **AR** strategy becomes

$$\mathbf{V}_{pure\,AR} = (1-\rho_+)\,\mathbf{V}_{IR} + \tfrac{1}{2}\,N\,(\rho_+ - \rho_-) + \tfrac{1}{2}\,(\rho_+ + \rho_-)\,(N_{1a} - N_{1b})^2 / N \quad [4.2.12]$$

and for the designs which partition into two orthogonal blocks and utilize more than two center runs, the equation for **V** under the *modified* **AR** strategy becomes

$$\mathbf{V}_{mod.\,AR} = (1-\rho_+)\,\mathbf{V}_{IR} + \tfrac{1}{2}\,N\,(\rho_+ - \rho_-)\,(s^2 + 2\,k\,w_{ii}\,s\,t + k\,t^2\,g)$$

$$+ \; N\,\rho_+\,(q^2 + 2\,k\,w_{ii}\,r\,q + k\,r^2\,g) \qquad [4.2.13]$$

$$+ \; \tfrac{1}{2}\,N\,(\rho_+ + \rho_-)\,(p^2 + 2\,k\,w_{ii}\,m\,p + k\,m^2\,g)\ .$$

The third form of **V** under the **AR** strategy applies to the FACs. These designs partition into three orthogonal blocks and utilize no additional center runs. The equation for $\mathbf{V}_{AR}$ for the 3^k designs becomes

$$\mathbf{V}_{AR,\,3fac} = (\,1 - h\,\rho_+)\,\mathbf{V}_{IR} + \tfrac{1}{2}\,N\,s^2\,(\rho_+ - \rho_-)\ . \qquad [4.2.14]$$

The fourth form of **V** under the **AR** strategy applies to the $k=4$ BBD. This design partitions into three orthogonal blocks and all of the center runs are placed in the 3rd block. The equation for $\mathbf{V}_{AR}$ for the $k=4$ BBD becomes

$$\mathbf{V}_{AR,\,bbd4} = (\,1 - \rho_+)\,\mathbf{V}_{IR} + \tfrac{1}{2}\,N\,k\,t^2\,g\,(\rho_+ - \rho_-) \qquad [4.2.15]$$

$$+ \; \rho_+ \left[\frac{a - 8\,w_{ii}\,[ii]}{D} + \frac{4\,w_{ii}}{3\,[ii]} + \frac{2\,\theta\,[ii]\,w_{iijj}}{r} + 4\,N\,(\,e\,w_{iiii} + (\,2\,f + e\,)\,w_{iijj}\,) \right] .$$

The variance component of **J** under the **IR**, **CR**, and **AR** strategies, as shown in equations [4.2.9]-[4.2.15], indicate that **V** is a function of the following parameters and constants:

$$\mathbf{V} = f\,(\,\theta\,,[ii]\,,r\,,k\,,w_{ii}\,,w_{iiii}\,,w_{iijj}\,,\rho_+\,,\rho_-\,,N\,,N_{1a}\,,N_{1b}\,)\ .$$

The values of r, k, N, N_{1a}, and N_{1b} are fixed for a given design plan, the w_{ii}, w_{iiii}, and w_{iijj} region moments are fixed for a given region shape, the ρ_+ and ρ_- induced correlations are fixed for a given simulation study, and θ is fixed for a Min-**B** design. Therefore, [ii] is the only term in the equations for **V** which can be altered for the minimization of **V** (given the Min-**B** value of θ). The value of [ii] which results in a Min-**V** | Min-**B** design is determined by setting the partial derivative of **V** with respect to [ii] equal to zero. The partial derivatives of $\mathbf{V}_{pure\,CR}$ and $\mathbf{V}_{pure\,AR}$ result in the same optimal values of [ii] as $\mathbf{V}_{IR}$ because the $(1-\rho_+)$ terms cancel out when the partial derivatives are set equal to zero and the additional terms in equations [4.2.10] and [4.2.12] are not functions of [ii]. The additional terms in $\mathbf{V}_{mod.\,CR}$, $\mathbf{V}_{mod.\,AR}$, $\mathbf{V}_{AR,\,3fac}$, and $\mathbf{V}_{AR,\,blk}$ are functions of [ii] through the Δ, e, f, m, p, q, Γ, s, and t terms. However, the $\mathbf{V}_{IR}$ optimal values of [ii] are used for the modified CR and AR strategies for the following two reasons:

1. The optimal values of [ii] under the *modified* CR and AR strategies are a function of N_c and, therefore, determination of the optimal values of [ii] and N_c would require the use of an iterative solution method.
2. Empirical findings of this research indicate that the optimal values of [ii] under the IR strategy result in values of [ii] which are near-optimal under the *modified* CR and AR strategies, and these near-optimal values of [ii] generally result in the same optimal values of N_c (because of the integer restriction on N_c).

In the remainder of this Chapter, it is assumed that the Min-**V** | Min-**B** optimal values of [ii] for the correlation induction strategies are the same as those obtained for the IR strategy. Therefore, this optimal value of [ii] is obtained by setting the partial derivative of $\mathbf{V}_{IR}$ with respect to [ii] equal to zero, yielding

$$\frac{\partial \mathbf{V}}{\partial [\mathrm{ii}]} = \frac{\partial}{\partial \theta}\left[\frac{k\,w_{ii}}{[\mathrm{ii}]} + \frac{r\,k\,(k-1)\,w_{iijj}}{2\theta\,[\mathrm{ii}]} + \frac{\theta\,(r+k-1)}{\theta\,(r+k-1) - r\,k\,[\mathrm{ii}]}\right]$$

$$+ \frac{\partial}{\partial \theta}\left[\frac{-2\,k\,r\,w_{ii}}{\theta\,(r+k-1) - r\,k\,[\mathrm{ii}]}\right]$$

$$+ \frac{\partial}{\partial \theta}\left[(k\,r\,w_{iiii})\,\frac{\theta\,(r+k-2) - r\,(k-1)\,[\mathrm{ii}]}{\theta^2\,(r-1)\,(r+k-1)\,[\mathrm{ii}] - r\,k\,\theta\,(r-1)\,[\mathrm{ii}]^2}\right]$$

$$+ \frac{\partial}{\partial \theta}\left[(k\,(k-1)\,r\,w_{iijj})\,\frac{r\,[\mathrm{ii}] - \theta}{\theta^2\,(r-1)\,(r+k-1)\,[\mathrm{ii}] - r\,k\,\theta\,(r-1)\,[\mathrm{ii}]^2}\right] = 0$$

$$0 = 2\theta k\, w_{ii} + r k (k-1)\, w_{iijj}$$

$$- \frac{2\theta^2 r k \ell_1 (r+k-1)[ii]^2 - 4\theta r^2 k^2 \ell\, w_{ii}[ii]^2}{d_1^2 - 2 d_1 e_1 [ii] + e_1^2 [ii]^2} \qquad [4.2.16]$$

$$- \frac{m_1 (2 b_1 e_1 [ii] - b_1 d_1 - c_1 e_1 [ii]^2) + n_1 (r e_1 [ii]^2 + \theta d_1 - 2\theta e_1 [ii])}{d_1^2 - 2 d_1 e_1 [ii] + e_1^2 [ii]^2}$$

where $b_1 = \theta r (r+k-2)$

$c_1 = r^2 (k-1)$

$d_1 = \theta^2 (r-1)(r+k-1)$

$e_1 = r k \theta (r-1)$

$j_1 = 2\theta k\, w_{ii} + r k (k-1)\, w_{iijj}$

$\ell_1 = \theta^2 (r-1)^2$

$m_1 = 2\theta k\, w_{iiii}$

$n_1 = 2\theta r k (k-1)\, w_{iijj}$.

Collecting the $[ii]^2$, $[ii]^1$, and $[ii]^0$ terms, the quadratic equation for the partial derivative of **V** set equal to zero becomes

$$0 = \ddot{o}_1 [ii]^2 + p_1 [ii]^1 + q_1 [ii]^0$$

where $\ddot{o}_1 = -2 r k \ell_1 \theta^2 (r+k-1) + 4 r^2 k^2 \ell_1 \theta\, w_{ii} + c_1 e_1 m_1 - n_1 r e_1 + e_1^2 j_1$

$p_1 = -2 j_1 d_1 e_1 - 2 m_1 b_1 e_1 + 2 n_1 \theta e_1$

$q_1 = b_1 d_1 m_1 - n_1 \theta d_1 + d_1^2 j_1$.

Applying the quadratic formula, the Min-**V** | Min-**B** value of [ii] becomes

$$[ii] = \frac{-p_1 \pm \sqrt{p_1^2 - 4 \ddot{o}_1 q_1}}{2 \ddot{o}_1} \qquad [4.2.17]$$

where the "−" root is the optimal value of [ii] because the "+" root is infeasible for the designs examined in this research. (See Figure P-1 on page 421 for a graphical display of **V** versus [ii] and N_c.) The Min-**V** | Min-**B** values of [ii] can be obtained by appropriately choosing the number of center runs. For the CCDs, BBDs, and FACs, the Min-**V** | Min-**B** value of N_c (as shown in Appendix P on pages 415-416) becomes

$$N_c = \frac{g^2(F + 2\alpha^2)}{[ii]} - (F + n_\alpha) \qquad \text{for the CCDs}$$

$$N_c = \frac{f\,F\,g^2}{[ii]} - F \qquad \text{for the BBDs and FACs} \qquad [4.2.18]$$

where g is the Min-**B** value of the scaling factor, [ii] is the Min-**V** | Min-**B** value of the pure second order design moment, and f is the fraction of factorial design points in which a particular design variable is equal to zero; that is

$f = 1/2$ for the $k = 4$ BBD,

$f = 2/5$ for the $k = 5$ BBD,

$f = 3/7$ for the $k = 7$ BBD, and

$f = 2/3$ for the FACs.

The Min-**V** | Min-**B** values of N_c for the second order designs examined in this research are shown in Tables 8 and 9 for spherical and cuboidal regions of interest, respectively. (The exact, non-integer values of N_c are given in Table P-2 on page 419.) The optimal number of center runs for the FACs have been rounded up to $N_c = 0$ because the optimal designs require a negative number of center runs, and the optimal values of N_c for the SCDs are only approximations based on the formula for the CCDs given in equation [4.2.18].

Table 8. **Min-V | Min-B Number of Center Runs in a Spherical Region of Interest. Optimal values of Nc for the CCDs, BBDs, FACs, and SCDs.**

DESIGN CLASS	Min-V | Min-B Values of Nc					
	k=2	k=3	k=4	k=5	k=6	k=7
CCD	2	2	2	3	2	3
BBD	-	-	2	3	-	2
FAC	-	0	0	0	0	0
SCD	-	1	1	1	1	2

Table 9. **Min-V | Min-B Number of Center Runs in a Cuboidal Region of Interest. Optimal values of Nc for the CCDs, BBDs, FACs, and SCDs.**

DESIGN CLASS	Min-V | Min-B Values of Nc					
	k=2	k=3	k=4	k=5	k=6	k=7
CCD	2	2	3	4	3	4
BBD	-	-	3	3	-	3
FAC	-	0	0	0	0	0
SCD	-	2	2	2	2	2

Utilizing the Min-V | Min-B values of N_c shown in Table 8, Figure 24 on page 192 presents a graphical display of the total number of design points (N) versus the number of factors in the model (k) for a spherical region of interest. The figure indicates that the number of design points in the CCDs and BBDs are similar, the SCDs require slightly fewer design points, and the FACs require the largest number of design points.

In the next section, an evaluation of the IR, CR, and AR correlation induction strategies is presented using the Min-V | Min-B second order designs.

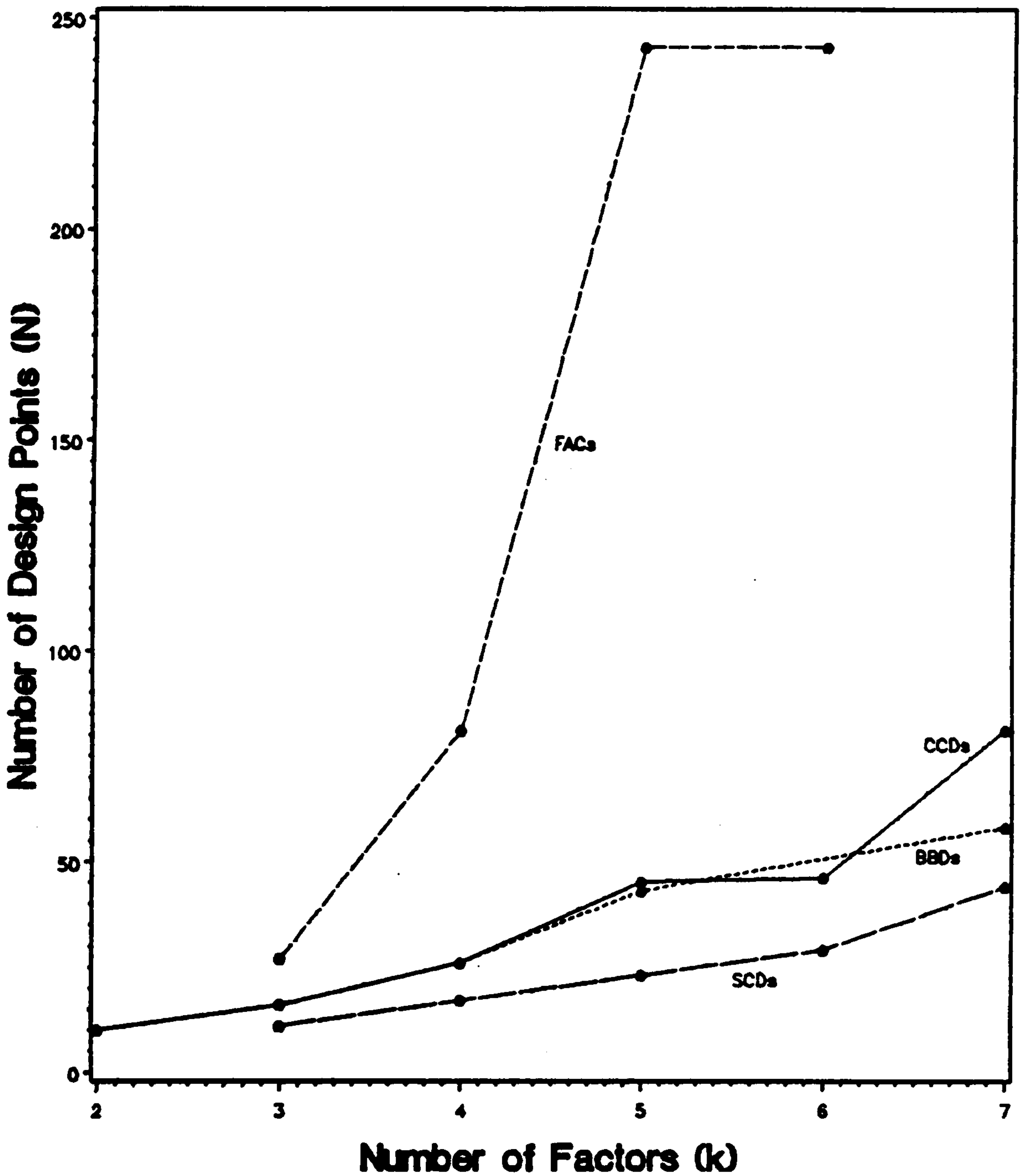


Figure 24. The Optimal Number of Design Points in a Min-V | Min-B Second Order Design.
The region of interest is Spherical.
An additional design point is generally required for a Cuboidal region.
Two additional design points are generally required for a Min-V* | Min-B* design.

4.2.3 Min-V | Min-B Second Order Designs

The optimal second order designs for the MSE of response criteria are Min-**V** | Min-**B** designs, resulting in minimum values of **V**, given minimum values of **B**. The Min-**B** designs are achieved through the use of the scaling factors specified in Tables 6 and 7 on page 179, and the Min-**V** | Min-**B** designs are achieved through the addition of the number of center runs specified in Tables 8 and 9 on page 191. The required levels of the axial design points (prior to scaling) for the CCDs and SCDs are shown in Table 5 on page 178.

In the first order case, the two-level factorial designs were used to evaluate the correlation induction strategies on the basis of minimum values of **J**. In the second order case, the correlation induction strategies can be evaluated (and the design classes compared) on the basis of minimum values of the **V** component alone. Three reasons for the use of **V**, instead of $\mathbf{J} = \mathbf{V} + \mathbf{B}$, for evaluation of the strategies in the second order case are as follows:

1. Computation of **B** requires specification of three unknown cubic parameters (Θ_1, Θ_2, and Θ_3), but the Min-**B** values of θ and g do not depend on the values of these parameters. Therefore, the optimal second order design plans remain unchanged for varying amounts of third order bias and varying magnitudes of **B**.

2. In the case of ordinary least squares estimation, the equation for **B** is the same under the **IR**, **CR**, and **AR** strategies. Therefore, only the **V** component of **J** is affected by the use of correlation induction.

3. Empirical evidence of this research indicates that the values of **B** are fairly constant from one design class to the next. The results for the $k=4$ designs shown in Table 10 on page 194 indicate that the values of **B** are similar for the CCDs and BBDs, but slightly larger for the FACs.

Table 10. Values of $\mathbf{B}$ for the $k=4$ Second Order Designs in a Spherical Region. Θ_1, Θ_2, and Θ_3 are standardized measures of cubic curvature. The second order designs are Min-V | Min-B designs.

DESIGN CLASS	Values of $\mathbf{B}$:						
	$\Theta_1=500$ $\Theta_2=0$ $\Theta_3=0$	$\Theta_1=0$ $\Theta_2=500$ $\Theta_3=0$	$\Theta_1=0$ $\Theta_2=0$ $\Theta_3=500$	$\Theta_1=500$ $\Theta_2=500$ $\Theta_3=0$	$\Theta_1=500$ $\Theta_2=0$ $\Theta_3=500$	$\Theta_1=0$ $\Theta_2=500$ $\Theta_3=500$	$\Theta_1=500$ $\Theta_2=500$ $\Theta_3=500$
CCD	0.7	17.5	1.0	18.1	1.7	18.5	19.2
BBD	1.5	17.2	1.0	18.8	2.6	18.2	19.8
FAC	9.4	14.6	1.0	24.1	10.4	15.6	25.0

Note: The SCDs are not included in the table because $\mathbf{B}$ is a function of the individual β coefficients in $\underline{\beta}_2$ (not just the Θ_1, Θ_2, and Θ_3 parameters).

The remainder of the research in this section presents an evaluation of the three correlation induction strategies and the four second order design classes using the Min-V | Min-B second order designs. Section 4.2.3.1 evaluates the CR strategy by comparing the values of **V** under the IR and CR strategies and section 4.2.3.2 evaluates the AR strategy by comparing the values of **V** under the IR and AR strategies. The results presented here correspond to a spherical region of interest; the results for a cuboidal region (shown in Appendices Q, R, and S) are similar.

4.2.3.1 Min-V | Min-B Designs Under the CR Strategy

The performance of the CR strategy relative to the IR strategy is presented in this section. The values of **V** for each strategy are computed using equations [4.2.9], [4.2.10] and [4.2.11] on page 186. The equation for $\mathbf{V}_{mod\,CR}$ reduces to $\mathbf{V}_{pure\,CR}$ when the number of center runs is less than two, and both $\mathbf{V}_{CR}$ equations reduce to $\mathbf{V}_{IR}$ when $\rho_+ = 0$. The preferred strategy (IR or CR) and the preferred design class (CCD, BBD, FAC, or SCD) are those which result in the least amount of variance error, **V**.

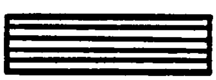
The correlation induction strategies can be evaluated by viewing plots of $\mathbf{V}_{CR}$ versus ρ_+ for each of the second order design classes. When $\rho_+ = 0$, the results correspond to the **IR** strategy, and when $\rho_+ > 0$, the results correspond to the **CR** strategy. Figures 25-28 on pages 197-200 illustrate the values of $\mathbf{V}_{IR}$ and $\mathbf{V}_{CR}$ for the CCDs, BBDs, FACs, and SCDs, respectively. (The data used to generate these plots are shown in Appendices Q and S on pages 434 and 460). For the CCDs, BBDs, and FACs, the values of $\mathbf{V}_{CR}$ increase as the magnitude of ρ_+ increases, indicating that the **IR** strategy is preferable to the **CR** strategy. For the SCDs, however, the **CR** strategy is preferable to the **IR** strategy because $\mathbf{V}_{CR}$ decreases as ρ_+ increases. Thus, for the second order designs examined in this research (except the SCDs), the **IR** strategy is preferable to the **CR** strategy when the design criteria is minimization of the MSE of response.

In addition to evaluating the correlation induction strategies, the performance of the four design classes are compared by viewing plots of $\mathbf{V}_{CR}$ versus ρ_+. Plots of the $k = 5$ designs are used to compare the design classes; similar results are realized for other values of k. Figure 29 on page 201 illustrates the values of $\mathbf{V}_{CR}$ and Figure 30 on page 202 shows the difference between the values of $\mathbf{V}_{CR}$ and $\mathbf{V}_{IR}$. Figure 29 indicates that the CCDs and BBDs are the preferred design classes under the **CR** strategy when $\rho_+ < .6$, but the SCDs perform slightly better when $\rho_+ > .6$. The FACs are the least preferred design class under the **CR** strategy and the SCDs are the least preferred design class under the **IR** strategy. Figure 30 indicates that as ρ_+ increases, the performance of the CCDs and BBDs deteriorates at the same rate as the performance of the SCDs improves, and the performance of the FACs deteriorates even more rapidly. Thus, for the MSE of response criteria, the CCDs and BBDs are the preferred design classes under the **CR** strategy when $\rho_+ < .6$, and the SCDs are the preferred design class when $\rho_+ > .6$.

The following *line* convention is used to distinguish between design classes in the CR strategy illustrations (Figures 24-30 and 41-46):

	Type of Line	Design Class
1.		CCDs
2.		BBDs
3.		FACs
4.		SCDs

and the following *shading* convention is used to distinguish between design classes in the AR strategy illustrations (Figures 31, 38-39, 47-48, and 53):

	Type of Shading	Design Class
1.		CCDs
2.		BBDs
3.		FACs
4.		SCDs.

The *shading* convention used in Figures 34-37 and Figures 49-52 to distinguish between designs with $k = 2, \ldots, 7$ factors is as follows: vertical (90°) lines for the $k = 7$ designs; 75° lines for the $k = 6$ designs; 60° lines the $k = 5$ designs; ... 15° lines the $k = 2$ designs.

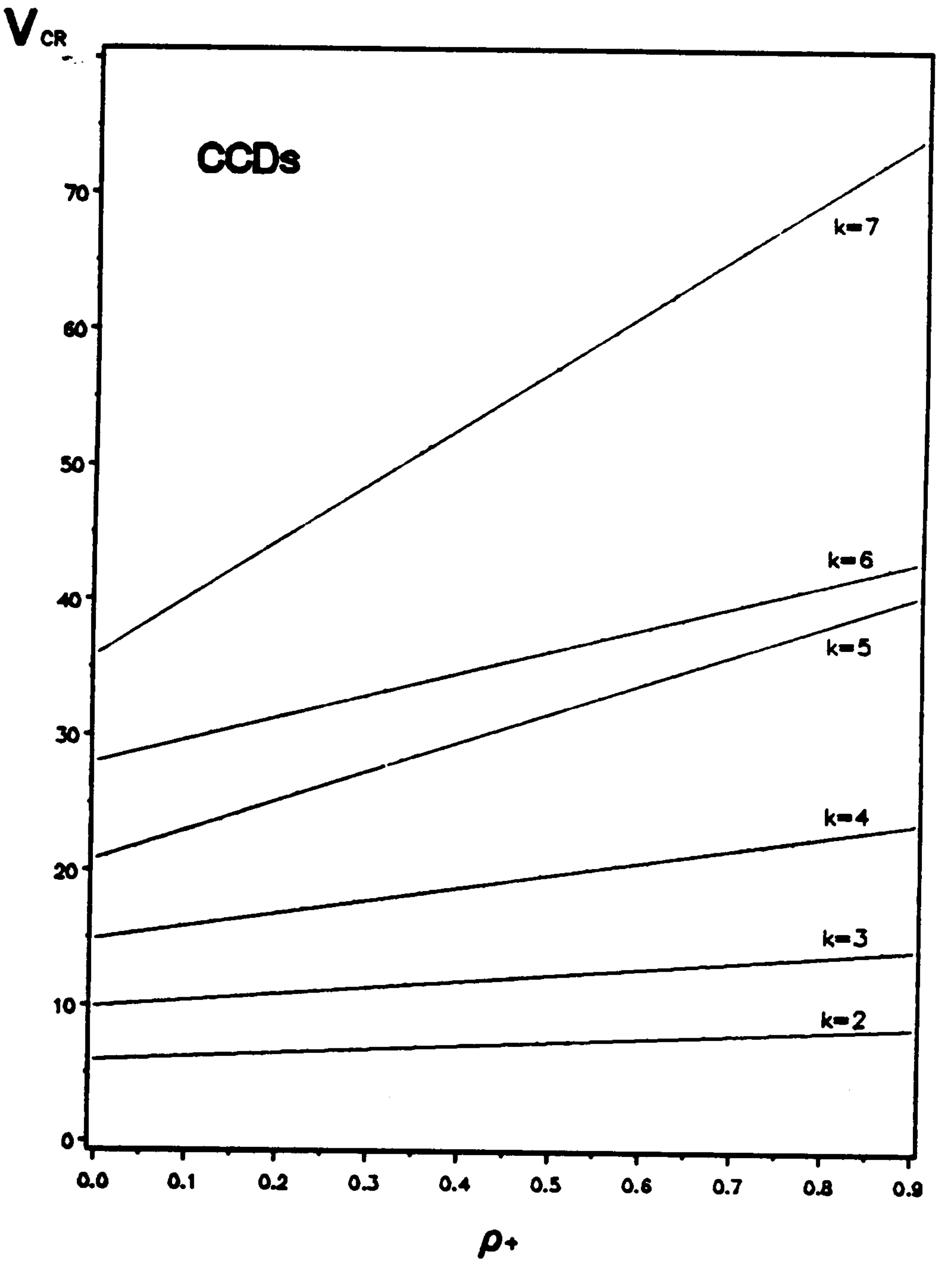


Figure 25. Optimal Values of V for the CCDs under the CR Strategy.

The CCDs are Min-V | Min-B designs.
The region of interest is Spherical.

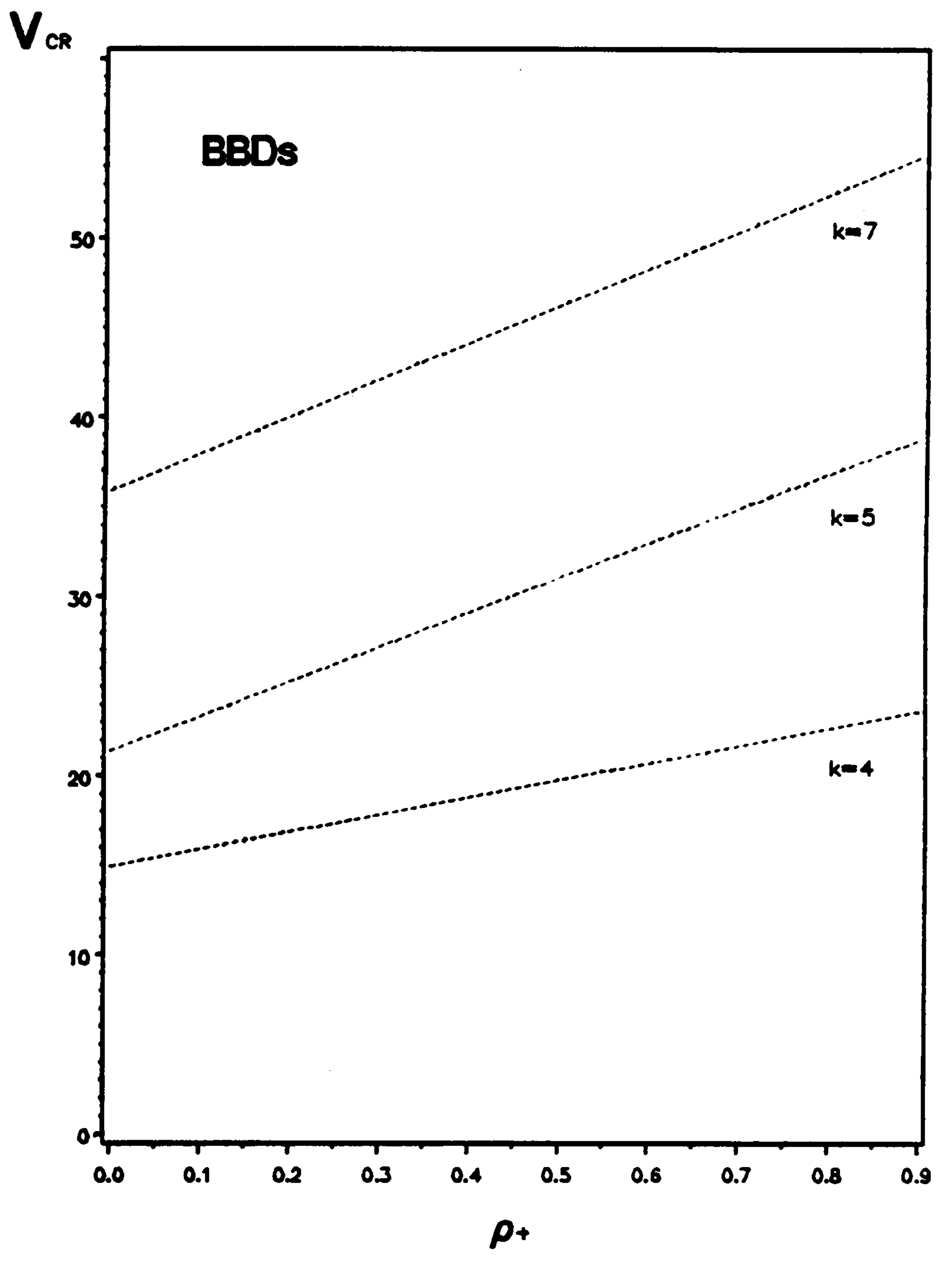


Figure 26. Optimal Values of V for the BBDs under the CR Strategy.

The BBDs are Min-V | Min-B designs.
The region of interest is Spherical.

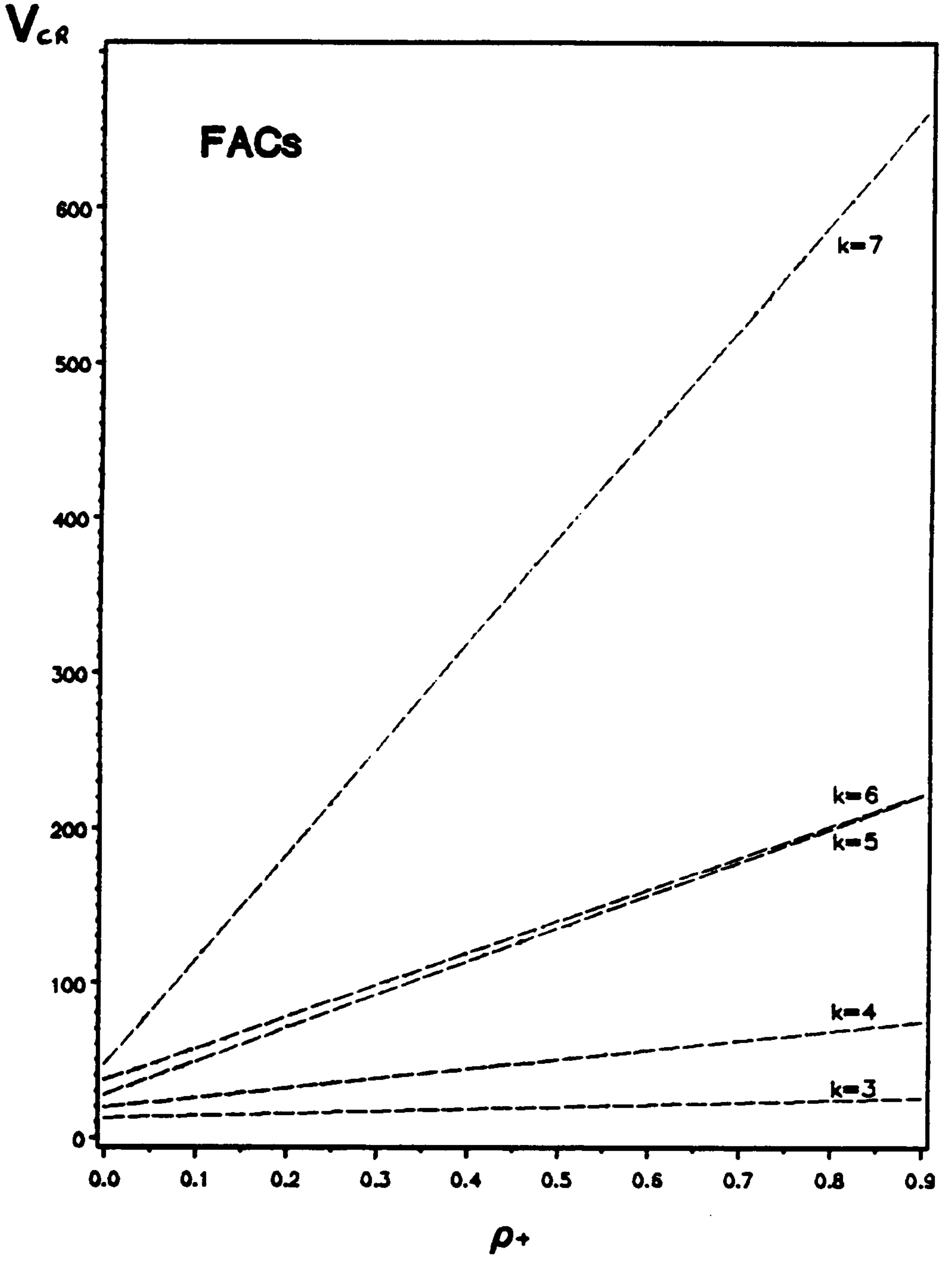


Figure 27. Optimal Values of V for the 3-level Factorial Designs under the CR Strategy.
The FACs (3^k designs) are Min-V | Min-B designs.
The region of interest is Spherical.

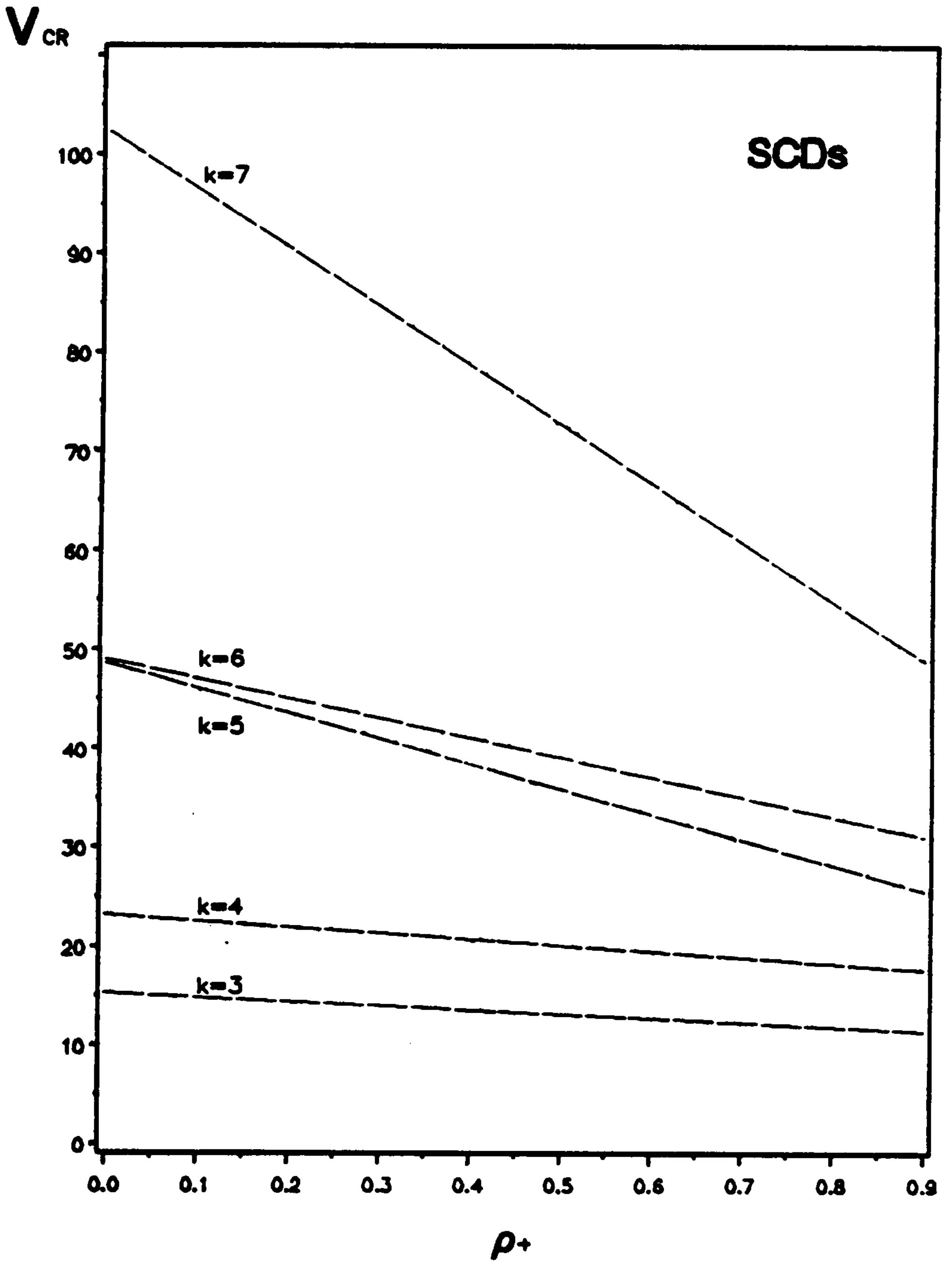


Figure 28. Optimal Values of V for the Small Composite Designs under the CR Strategy.

The SCDs (small composite designs) are Min-V | Min-B designs.
The region of interest is Spherical.

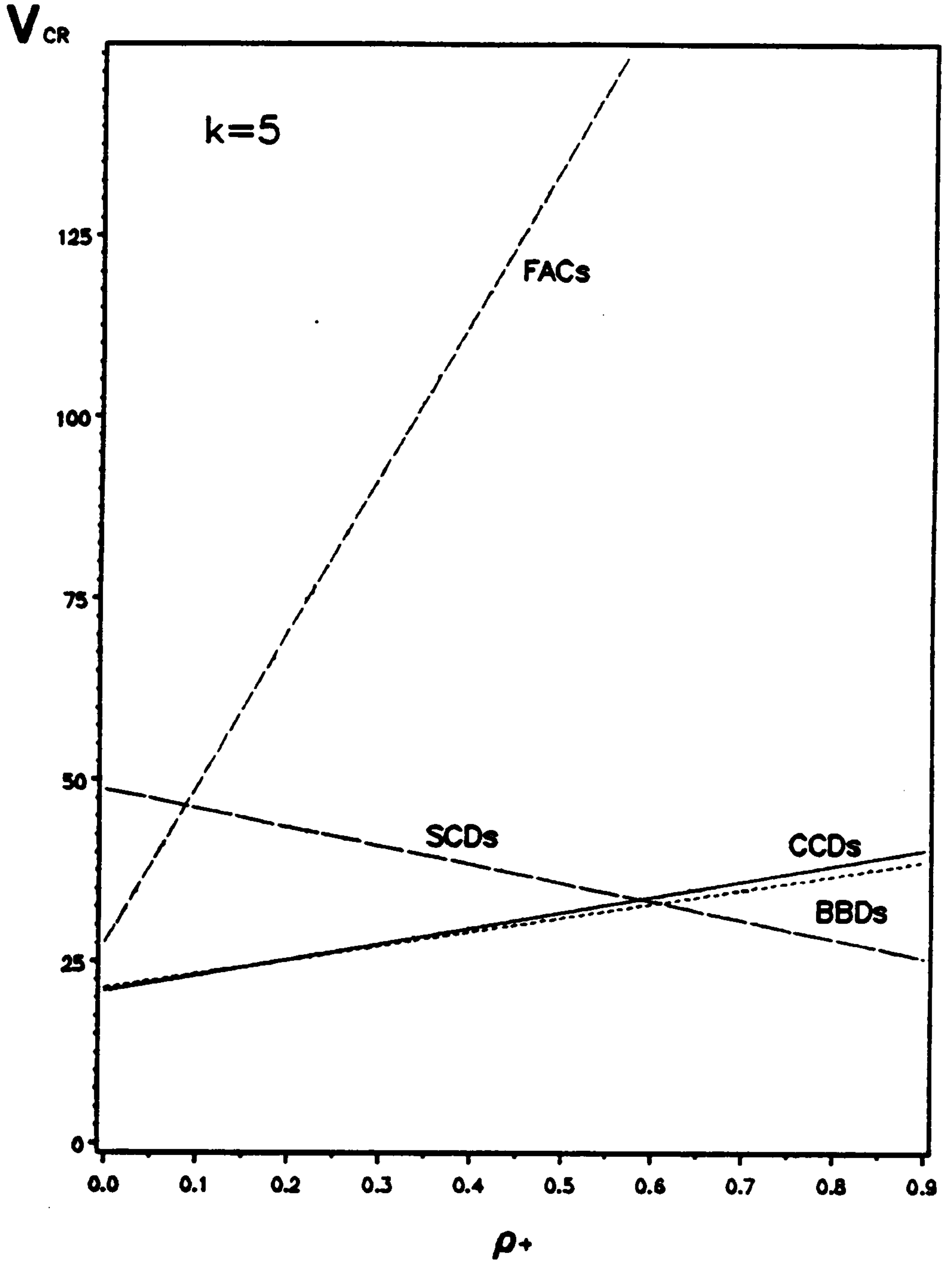


Figure 29. Optimal Values of V for a k=5 Second Order Design under the CR Strategy.
The optimal k=5 second order designs are Min-V | Min-B designs.
The region of interest is Spherical.

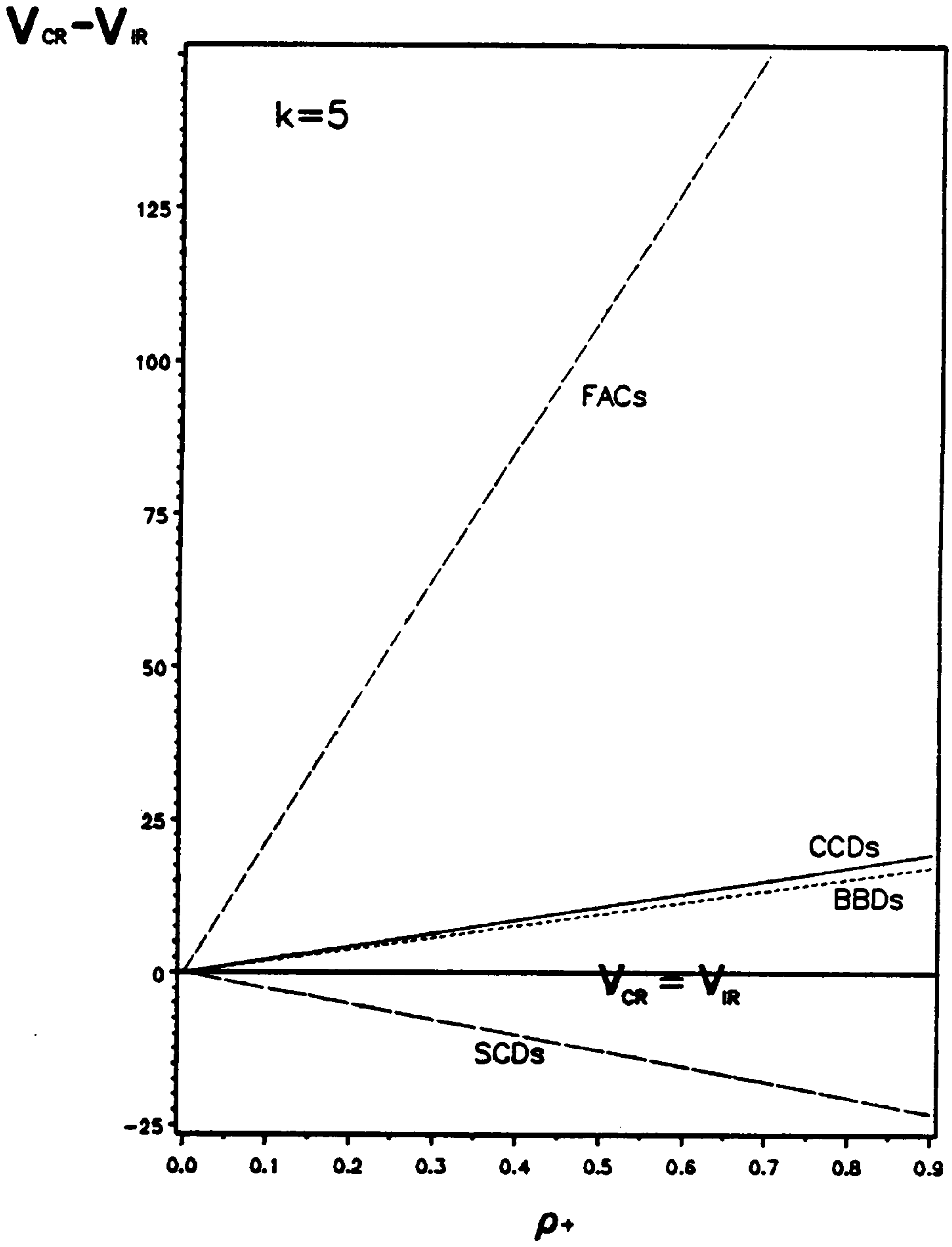


Figure 30. Differences Between the Optimal Values of V under the CR and IR Strategies.

The optimal $k=5$ second order designs are Min-V | Min-B designs.
The region of interest is Spherical.

4.2.3.2 Min-$\mathbf{V}$ | Min-$\mathbf{B}$ Designs Under the AR Strategy

The performance of the AR strategy relative to the IR and CR strategies is considered in this section. The values of $\mathbf{V}$ for the AR strategy are computed using equations [4.2.12], [4.2.13], [4.2.14], and [4.2.15] on pages 186-187. The equation for $\mathbf{V}_{mod.AR}$ reduces to $\mathbf{V}_{pure\,AR}$ when the number of center runs is less than three, the equation for $\mathbf{V}_{bbd4,AR}$ reduces to $\mathbf{V}_{3fac,AR}$ when the number of center runs is zero, and all of the $\mathbf{V}_{AR}$ equations reduce to $\mathbf{V}_{IR}$ when $\rho_+ = \rho_- = 0$. The preferred strategy (IR, CR, or AR) and the preferred design class (CCD, BBD, FAC, or SCD) are those which result in the least amount of variance error, $\mathbf{V}$.

The relative performance of the correlation induction strategies can be evaluated by viewing plots of $\mathbf{V}$ versus ρ_+ for various magnitudes of ρ_-. The values of $\mathbf{V}_{AR}$ are indicated by shaded triangular areas on the plots, with the upper portions of the shaded areas corresponding to low values of ρ_- and the lower portions corresponding to high values of ρ_-. The largest values of $\mathbf{V}$ are realized when $\rho_- = 0$, or equivalently, when $R = 0$, where R is defined as

$$R = \frac{\rho_-}{\rho_+} \quad ; \qquad 0 \le R \le 1 \qquad [4.2.19]$$

and the smallest values of $\mathbf{V}$ are achieved when $\rho_- = \rho_+$, or equivalently, when $R = 1$.

Figures 31 and 32 illustrate the two different situations that arise when evaluating the IR and AR strategies, and Figure 33 illustrates the similarity of the results for spherical and cuboidal regions of interest. Figure 31 on page 206 shows the values of $\mathbf{V}$ for the $k=2$ CCD under the IR, CR, and AR correlation induction strategies. The values of $\mathbf{V}$ are the highest under the CR strategy, and $\mathbf{V}_{CR}$ increases as the magnitude of ρ_+ increases. The values of $\mathbf{V}$ are the lowest under the AR strategy, and $\mathbf{V}_{AR}$ decreases as the magnitudes of ρ_+ and ρ_- increase (as R approaches 1). Figure 32 on page 207 illustrates the values of $\mathbf{V}$ for the $k=5$ CCD under each strategy. Again, the CR strategy performs the poorest, but unlike the $k=2$ CCD, the AR strategy is preferable to the IR strategy only when $R > .392$. Figure 33 on page 208 shows a comparison of the values of **V** for the $k=5$ CCD in a spherical and cuboidal region of interest. The values of **V** are slightly lower in a spherical region, but the relative performance of the strategies is the same in both regions.

The performance of the AR and IR correlation induction strategies are compared by viewing plots of $\mathbf{V}_{AR}$ versus ρ_+ for each design class. When $\rho_+ = \rho_- = 0$, the results correspond to the IR strategy, and when $\rho_+ \geq \rho_- > 0$, the results correspond to the AR strategy. Figures 34-37 on pages 209-212 illustrate the values of $\mathbf{V}_{AR}$ for the CCDs, BBDs, FACs, and SCDs, respectively. (The data used to generate these plots are shown in Appendices Q and S on pages 436 and 460). For the $k=2, 3, 4, 6$ CCDs, the BBDs, the $k=3$ FAC, and the SCDs, the values of $\mathbf{V}_{AR}$ decrease as the magnitudes of ρ_+ and ρ_- increase, indicating that the AR strategy is preferable to the IR strategy. However, for the $k=5, 7$ CCDs and the $k=4, 5, 6, 7$ FACs, the relative performance of the AR and IR strategies depends on the magnitudes of ρ_+ and ρ_-. This result is due to a deterioration in the performance of the AR strategy for designs with a large number of experimental runs.

Tables 11 and 12 on page 205 show the minimum values of $R = \rho_- / \rho_+$ needed for the AR strategy to be preferable to the IR strategy in spherical and cuboidal regions of interest, respectively. The tables indicate that the AR strategy is preferable to the IR strategy for all of the second order designs except the $k=5, 7$ CCDs and the $k=4, 5, 6, 7$ FACs. For these designs, the AR strategy is preferable to the IR strategy only when the magnitudes of ρ_+ and ρ_- are similar. As the number of experimental runs increases, larger ratios of ρ_- to ρ_+ are necessary for the AR strategy to be preferred to the IR strategy.

Table 11. Minimum Values of $R = \rho_- / \rho_+$ for the AR Strategy in a Spherical Region.

DESIGN CLASS	Minimum Values of $R = \rho_- / \rho_+$					
	k=2	k=3	k=4	k=5	k=6	k=7
CCD	0	0	0	.39	0	.80
BBD	-	-	0	0	-	0
FAC	-	0	.28	.66	.55	.81
SCD	-	0	0	0	0	0

Note: The second order designs are Min-V | Min-B designs.
The range of possible values for R is: $0 \leq R \leq 1$.
The AR strategy is preferable to the IR strategy when $R > R_{min}$.

Table 12. Minimum Values of $R = \rho_- / \rho_+$ for the AR Strategy in a Cuboidal Region.

DESIGN CLASS	Minimum Values of $R = \rho_- / \rho_+$					
	k=2	k=3	k=4	k=5	k=6	k=7
CCD	0	0	0	.41	0	.82
BBD	-	-	0	0	-	0
FAC	-	0	.42	.73	.64	.85
SCD	-	0	0	0	0	0

Note: The AR strategy is always preferable to the IR Strategy for the $k = 2, 3, 4, 6$ CCDs, BBDs, $k = 3$ FAC, and SCDs.

The relative performance of the four design classes under the AR strategy is illustrated in Figures 38 and 39 on pages 213-214. These figures indicate that the CCDs and BBDs are the preferred design classes under the AR strategy, but the SCDs also perform well when $\rho_+ > .6$. The CCDs tend to perform slightly better than the BBDs for large magnitudes of ρ_-, and the reverse is true for small magnitudes of ρ_-. The FACs are the least preferred design class under the AR strategy, but their performance improves rapidly as ρ_- increases. Thus, similar to the results for the IR and CR strategies, the CCDs and BBDs are the preferred design classes under the AR strategy when the design criteria is minimization of the MSE of response.

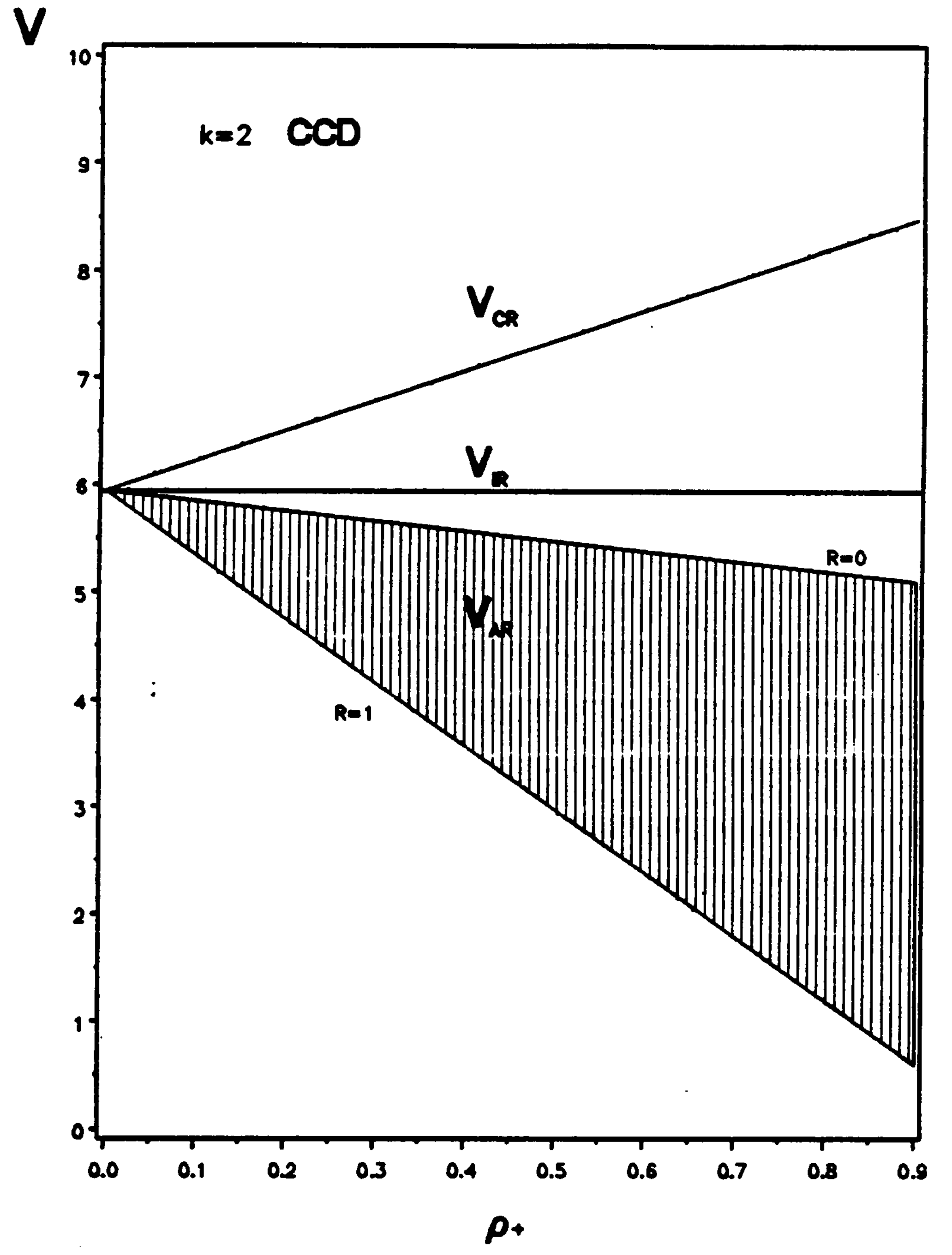


Figure 31. Optimal Values of V for the k=2 CCD under the IR, CR, and AR Strategies.

The $k=2$ CCD is a Min-V | Min-B design.

The region of interest is Spherical.

The shaded V_{AR} region corresponds to values of $R=\rho_-/\rho_+$ between 0 and 1.

The AR strategy always performs better than the IR strategy in this situation.

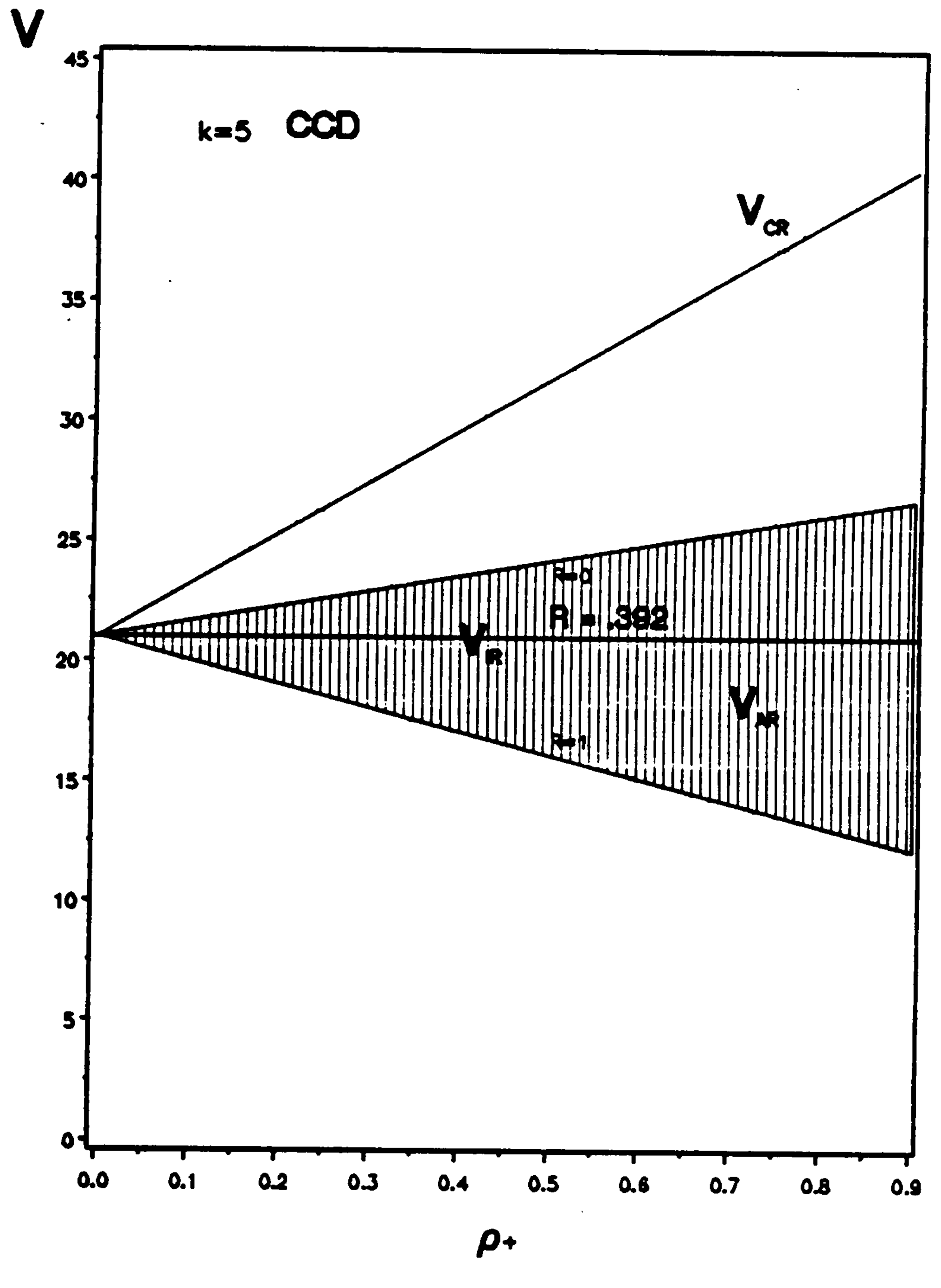


Figure 32. Optimal Values of V for the k = 5 CCD under the IR, CR, and AR Strategies.
The $k=5$ CCD is a Min-V | Min-B design.
The region of interest is Spherical.
The shaded V_{AR} region corresponds to values of $R = \rho_- / \rho_+$ between 0 and 1.
For this design, the AR strategy performs better than the IR strategy when $R > .392$.

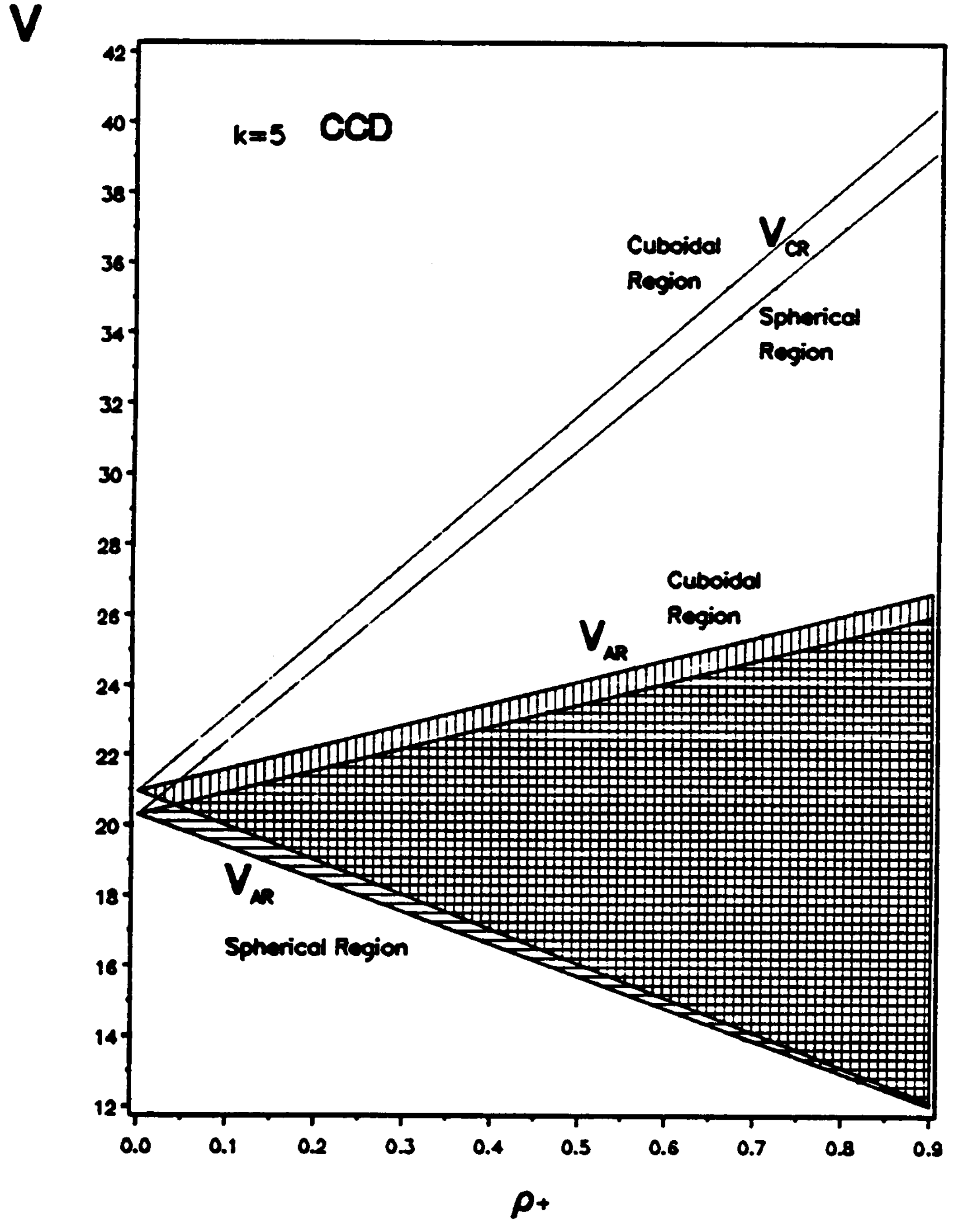


Figure 33. Optimal Values of V for the $k=5$ CCD in a Spherical or Cuboidal Region.

The $k=5$ CCD is a Min-V | Min-B design.

The shaded V_{AR} regions correspond to values of $R=\rho_-/\rho_+$ between 0 and 1.

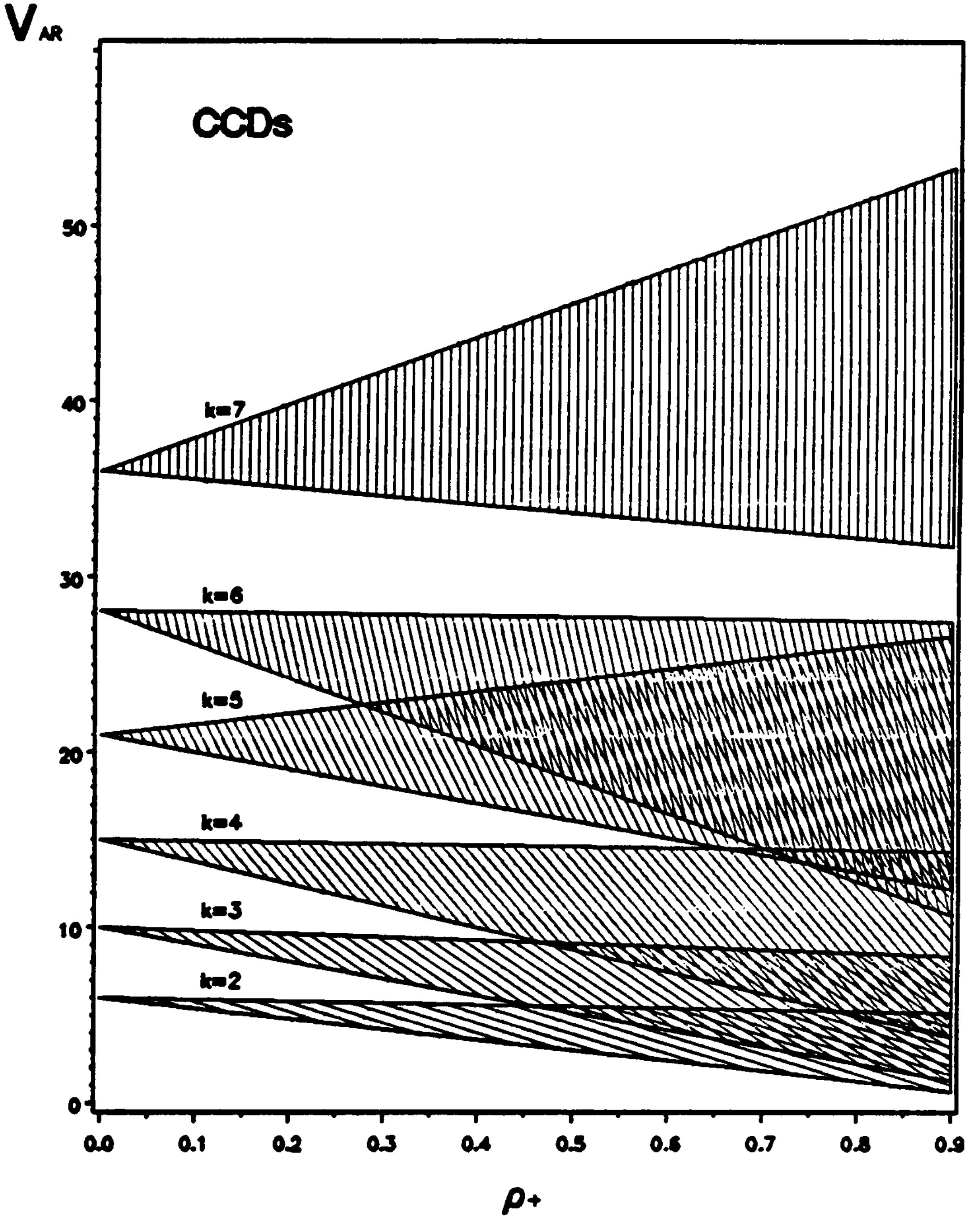


Figure 34. Optimal Values of V for the CCDs under the AR Strategy.

The CCDs are Min-V | Min-B designs.

The region of interest is Spherical.

R is defined as: $R = \rho_- / \rho_+ = 1$ and $0 \leq R \leq 1$.

The upper triangular lines correspond to $R = 0$ and the lower correspond to $R = 1$.

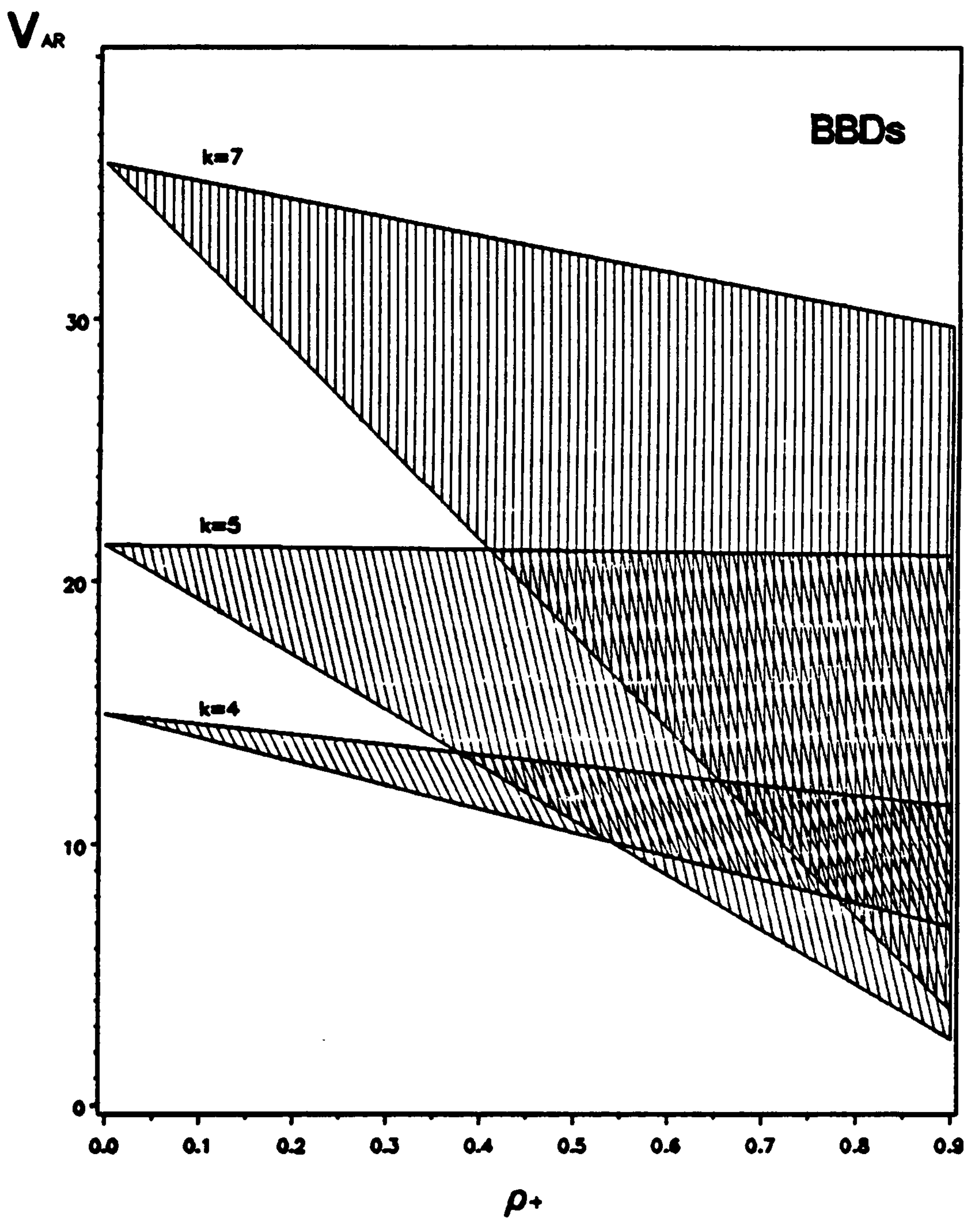


Figure 35. Optimal Values of V for the BBDs under the AR Strategy.

The BBDs are Min-V | Min-B designs.

The region of interest is Spherical.

R is defined as: $R = \rho_- / \rho_+ = 1$ and $0 \leq R \leq 1$.

The upper triangular lines correspond to $R = 0$ and the lower correspond to $R = 1$.

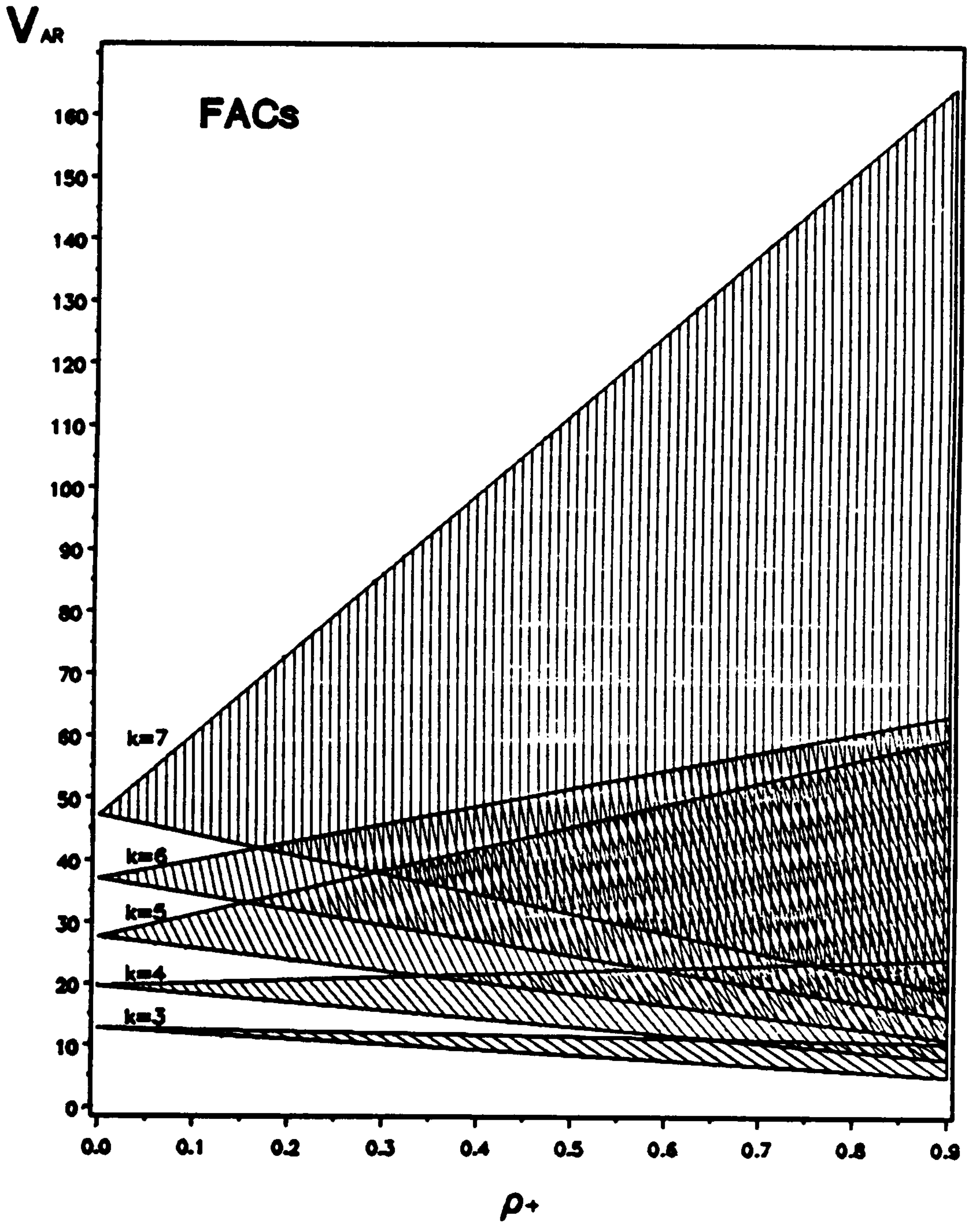


Figure 36. Optimal Values of V for the 3-level Factorial Designs under the AR Strategy.

The FACs (3^k designs) are Min-V | Min-B designs.

The region of interest is Spherical.

R is defined as: $R = \rho_- / \rho_+ = 1$ and $0 \leq R \leq 1$.

The upper triangular lines correspond to $R = 0$ and the lower correspond to $R = 1$.

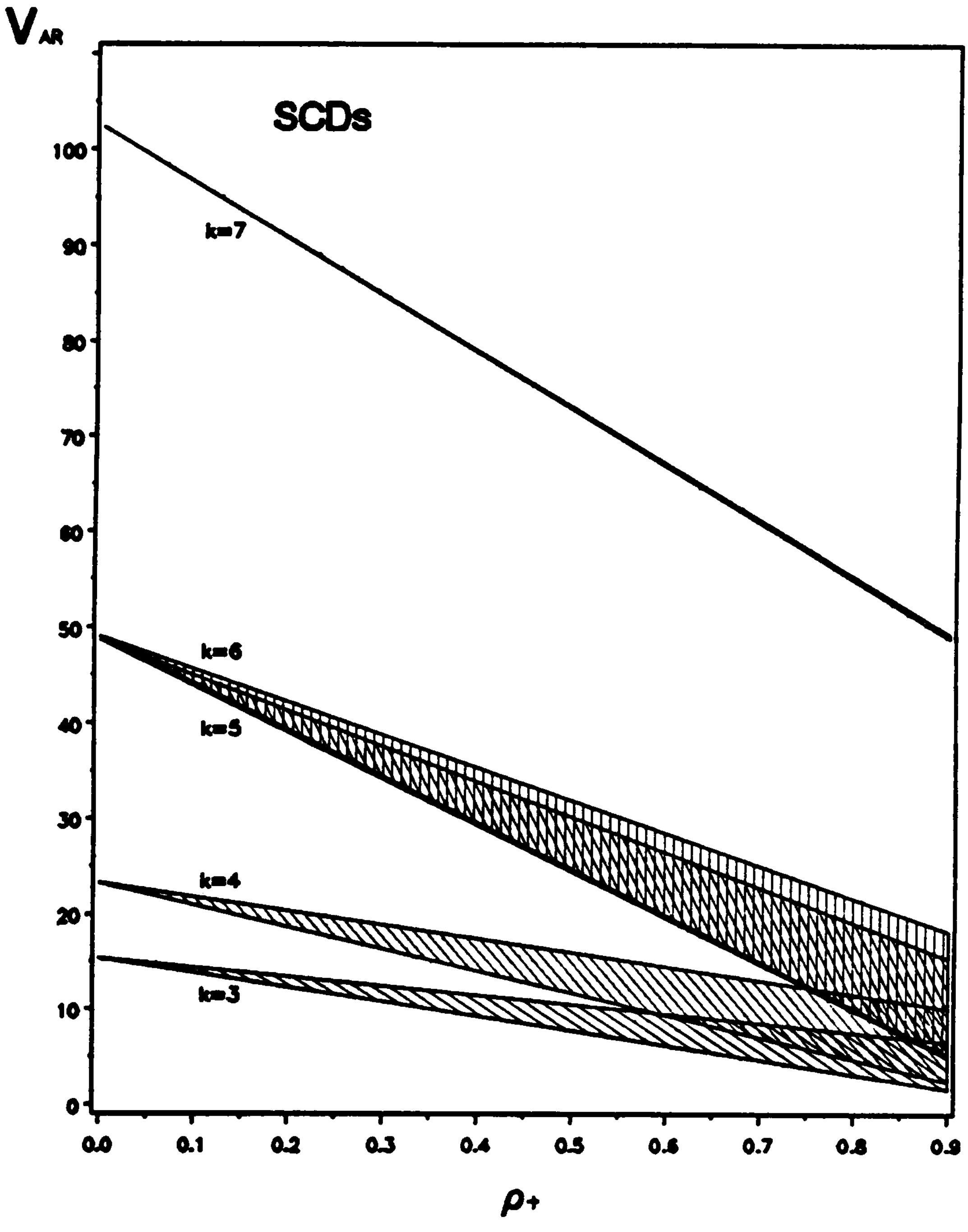


Figure 37. Optimal Values of V for the Small Composite Designs under the AR Strategy.
The SCDs (small composite designs) are Min-V | Min-B designs.
The region of interest is Spherical.
R is defined as: $R = \rho_- / \rho_+ = 1$ and $0 \leq R \leq 1$.
The upper triangular lines correspond to $R = 0$ and the lower correspond to $R = 1$.

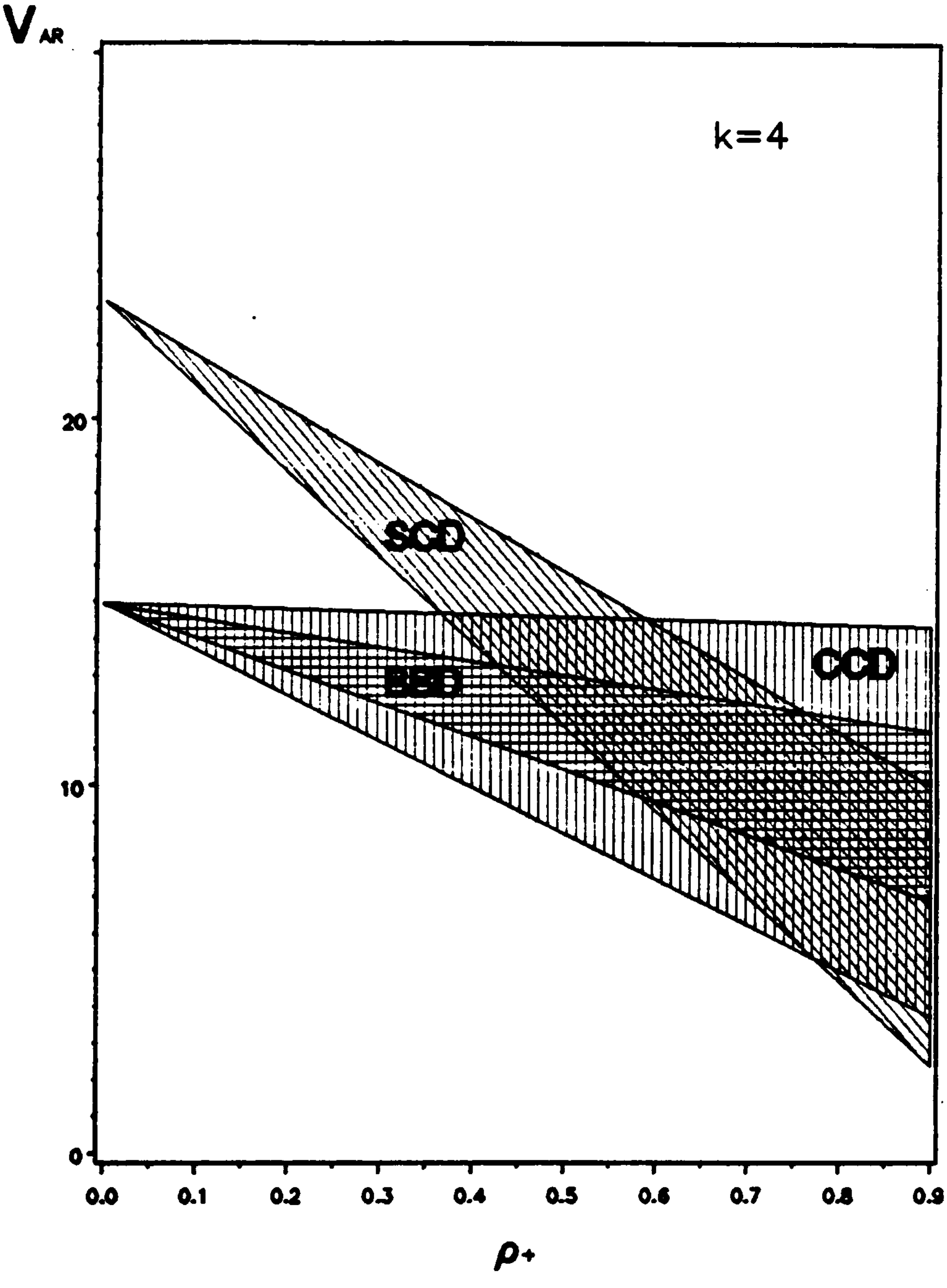


Figure 38. Optimal Values of V for a k = 4 Second Order Design under the AR Strategy.
The optimal designs are Min-V | Min-B designs.
The region of interest is Spherical.
The 3^k design is not shown here (see Figure 36).

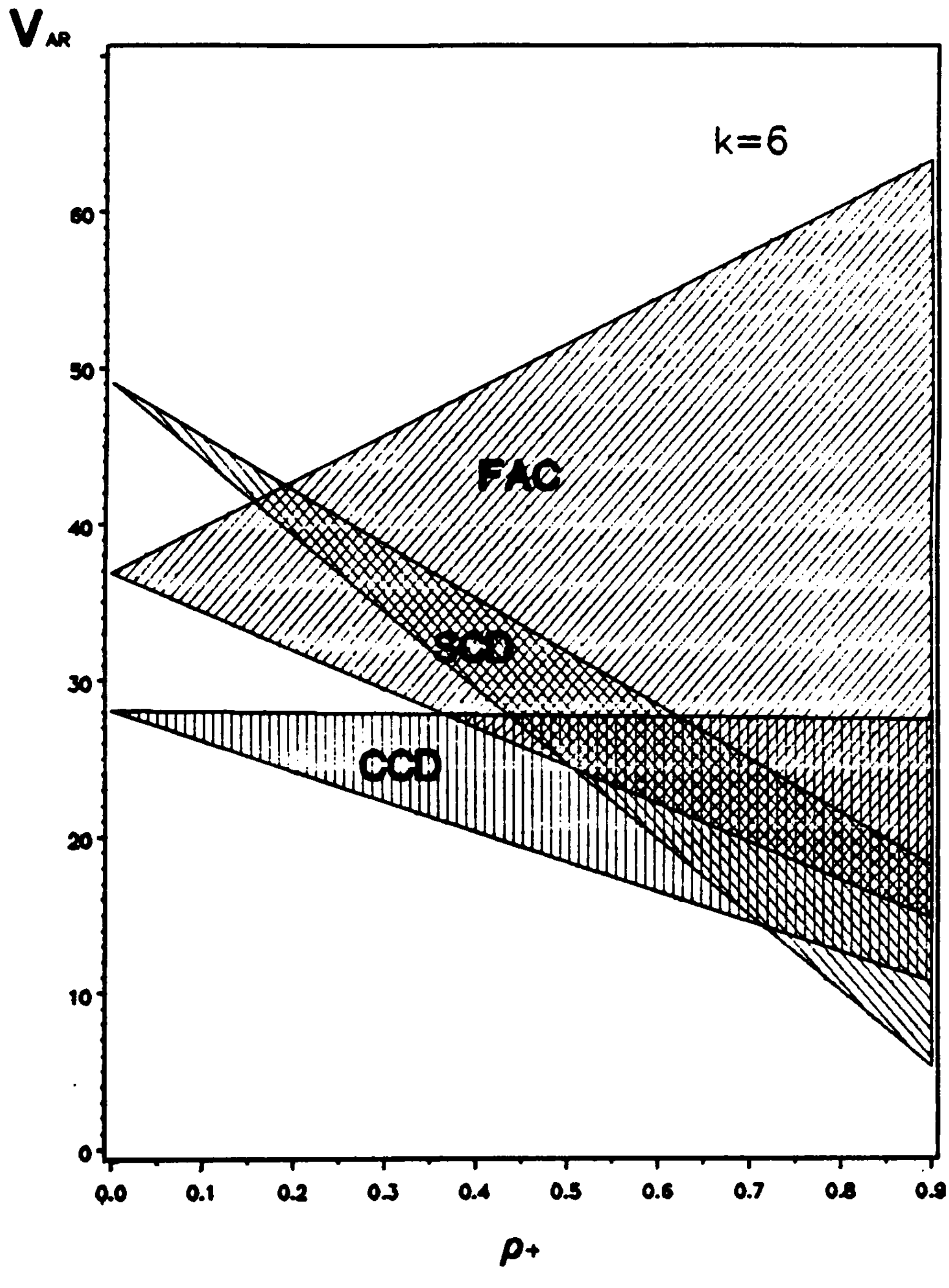


Figure 39. Optimal Values of V for a k=6 Second Order Design under the AR Strategy.

The optimal designs are Min-V | Min-B designs.

The region of interest is Spherical.

The $k=6$ BBD is not examined in this research and, therefore, not shown here.

4.3 MSE of Slope Criteria in the Second Order Case

This section focuses on the evaluation of the correlation induction strategies in terms of the MSE of slope criteria using the CCD, BBD, FAC, and SCD second order design classes. The MSE of slope criteria (discussed on pages 73-78) calls for minimizing $\mathbf{J}^*$, the average, normalized, mean squared error of the slopes of the response function. The $\mathbf{J}^*$ criteria is often useful prior to the location of the region containing the optimum response when it is of interest to minimize the errors associated with estimating the *change* in the value of the response variable.

The optimal designs in the first order case were Min-$\mathbf{J}^*$ designs, utilizing the values of [ii] which minimized $\mathbf{V}^* + \mathbf{B}^*$. In the second order case, $\mathbf{J}^*$ cannot be minimized with respect to [ii] alone, and therefore the optimal second order designs considered here are Min-$\mathbf{V}^*$ | Min-$\mathbf{B}^*$ designs. Similar to the optimal designs for the MSE of response criteria, the $\mathbf{J}^*$-optimal designs utilize the values of the scaling factor which minimize $\mathbf{B}^*$ and the number of center runs which minimize $\mathbf{V}^*$ (for the Min-$\mathbf{B}^*$ designs). Because the optimal number of center runs must be rounded to the nearest integer value, the $\mathbf{J}^*$-optimal second order designs are *near* Min-$\mathbf{V}^*$ | Min-$\mathbf{B}^*$ designs.

The research of this Chapter assumes that an experimenter is fitting second order polynomial regression model ($d_1 = 2$) and desires protection against bias in the fitted model coefficients due to unfitted cubic and third order interaction terms ($d_2 = 3$). The equation representing the slopes, or partial derivatives, of the fitted and true response functions, as shown in equations [2.2.50] and [2.2.49] on pages 75-76, are written as

$$\hat{\underline{\gamma}}_{(\underline{x})} = \frac{\partial \hat{y}}{\partial \underline{x}} = \Lambda'_{1\,(\underline{x})}\, \underline{b}_1$$

$$\underline{\gamma}_{(\underline{x})} = \frac{\partial y}{\partial \underline{x}} = \Lambda'_{1\,(\underline{x})}\, \underline{\beta}_1 + \Lambda'_{2\,(\underline{x})}\, \underline{\beta}_2$$

where $\Lambda'_{1(\omega)}$ is a matrix of the partial derivatives of the fitted first and second order terms, and $\Lambda'_{2(\omega)}$ is a matrix of the partial derivatives of the unfitted cubic and third order interaction terms. Both of these matrices are illustrated in Appendix K (pages 353-358). The MSE of slope, $\mathbf{J}^*$, is the average, normalized mean squared error of $\hat{\underline{y}}_{(\omega)}$, computed as the sum of the bias and variance components, $\mathbf{B}^*$ and $\mathbf{V}^*$.

The next two sections present the $\mathbf{B}^*$ and $\mathbf{V}^*$ components of the mean squared error of $\hat{\underline{y}}_{(\omega)}$. Similar to the MSE of response criteria research, the developments are applicable to the CCDs, BBDs, and FACs, but do not necessarily hold for the SCDs (because of their non-zero odd order design moments).

4.3.1 Bias Component of $\mathbf{J}^*$ in the Second Order Case

The bias component of the MSE of slopes, $\mathbf{B}^*$, is computed as the average squared bias of the partial derivative of the vector of predicted responses, normalized with respect to the number of design points, N, and the experimental error variance, σ^2. Myers and Lahoda (48) present a mathematical development of $\mathbf{B}^*$, but these authors invoke a sufficient condition for minimization of $\mathbf{B}^*$ and, as a result, restrict their research to rotatable CCDs. This research extends the work of Myers and Lahoda by developing the $\mathbf{B}^*$ component for non-rotatable designs in the CCD, BBD, and FAC design classes.

Myers and Lahoda (48) define the bias component of $\mathbf{J}^*$, as shown in equation [2.2.56] on page 78, as

$$\mathbf{B}^* = \frac{N\Omega_r}{\sigma^2} \int_R \text{Bias}^2 [\hat{y}(\underline{x})]\, d\underline{x}$$

$$= \frac{N}{\sigma^2}\, \underline{\beta}'_2 \left\{ A' \mu^*_{11} A - 2\mu^*_{21} A + \mu^*_{22} \right\} \underline{\beta}_2$$

where Ω_r is the inverse of the volume of the centered and scaled region of interest, A is the alias matrix, and the μ^* terms are the region matrices of the partial derivatives of the fitted and unfitted model terms.

For the second order designs examined here, which have the two design moment properties shown on page 170, the equation for $\mathbf{B}^*$ can be simplified to a scalar equation involving three unknown third order parameters (Θ_1, Θ_2, and Θ_3). The $\mathbf{B}^*$ component of $\mathbf{J}^*$ is obtained by summing the $A' \mu^*_{11} A$, $2\mu^*_{21} A$, and μ^*_{22} component matrices of $\mathbf{B}^*$ shown in Appendix O (pages 396-407), yielding

$$\mathbf{B}^* = \frac{N}{\sigma^2}\, \underline{\beta}_2' \left[\begin{array}{c|c} G_{B^*} & \mathbf{0} \\ \hline \mathbf{0} & 3\, w_{iiij}\, \mathbf{I}_{\binom{k}{3}} \end{array} \right] \underline{\beta}_2$$

where the $\mathbf{0}$'s indicate null matrices and G_{B^*} is a ($k^2 \times k^2$) block diagonal matrix, with k identical blocks of the ($k \times k$) H_{B^*} matrix

$$H_{B^*} = \left[\begin{array}{c|c} h^*_1 & h^*_2\, \underline{1}'_{k-1} \\ \hline h^*_3\, \underline{1}_{k-1} & (h^*_4 - h^*_5)\, \mathbf{I}_{k-1} + h^*_5\, \underline{1}_{k-1}\underline{1}_{k-1}' \end{array} \right]$$

where the h_i^* terms are scalar quantities defined as

$$\begin{aligned}
h_1^* &= \theta^2 - 6\theta w_{ii} + 9 w_{iiii} \\
h_2^* &= \theta^2 / r - 6\theta w_{ii} / r + 3 w_{iiii} \\
h_3^* &= \theta^2 / r - 2\theta w_{ii} + 3 w_{iiii} \\
h_4^* &= \theta^2 / r^2 - 2\theta w_{ii} / r + w_{iiii} + 4 w_{iijj} \\
h_5^* &= \theta^2 / r - 2\theta w_{ii} / r^2 + w_{iijj}
\end{aligned}$$

and θ and r are ratios of design moments, as defined in equations [4.2.1] and [4.2.2] on page 174.

The $\mathbf{B}^*$ component of $\mathbf{J}^*$ can be written in equation form as

$$\mathbf{B}^* = \frac{N}{\sigma^2}\left[\{h_1^* + (k-1) h_3^*\} \sum_{i=1}^{k} \beta_{iii}^2 + \{h_2^* + h_4^* + (k-2) h_5^*\} \sum_{i \neq j}\sum \beta_{ijj}^2 + 3 w_{iijj} \sum_{i \neq j \neq k}\sum\sum \beta_{ijk}^2\right]$$

$$= \phi_1^* \Theta_1 + \phi_2^* \Theta_2 + \phi_3^* \Theta_3 \qquad [4.3.1]$$

where $\phi_1^* = h_1^* + (k-1) h_3^*$ $\qquad \Theta_1 = \frac{N}{\sigma^2} \sum_{i=1}^{k} \beta_{iii}^2$

$\phi_2^* = h_2^* + h_4^* + (k-2) h_5^*$ $\qquad \Theta_2 = \frac{N}{\sigma^2} \sum_{i \neq j}\sum \beta_{ijj}^2$

$\phi_3^* = 3 w_{iijj}$ $\qquad \Theta_3 = \frac{N}{\sigma^2} \sum_{i \neq j \neq k}\sum\sum \beta_{ijk}^2$.

Equation [4.3.1] indicates that $\mathbf{B}^*$ is a function of the following parameters and constants:

$$\begin{aligned}
\mathbf{B}^* &= f(\phi_1^*, \phi_2^*, \phi_3^*, \Theta_1, \Theta_2, \Theta_3) \\
&= f(h_1^*, h_2^*, h_3^*, h_4^*, h_5^*, k, w_{iijj}, \Theta_1, \Theta_2, \Theta_3) \\
&= f(\theta, r, k, w_{ii}, w_{iiii}, w_{iijj}, \Theta_1, \Theta_2, \Theta_3) .
\end{aligned}$$

The values of r, k, w_{ii}, w_{iiii}, and w_{iijj} are fixed and the Θ parameters are unknown, leaving θ as the only term which can be altered to minimize $\mathbf{B}^*$. The value of θ which results in a Min-$\mathbf{B}^*$ design is determined by setting the partial derivative of $\mathbf{B}^*$ equal to zero, yielding

$$\frac{\partial \mathbf{B}^*}{\partial \theta} = \frac{\partial}{\partial \theta}\left[\frac{\theta^2 (r+k-1)\Theta_1}{r} - 2\theta w_{ii}(k+2)\Theta_1 + 3 w_{iiii}(k+2)\Theta_1\right]$$

$$+\frac{\partial}{\partial \theta}\left[\frac{\theta^2 (r+k-1)\Theta_2}{r^2} - \frac{2\theta w_{ii}(k+2)\Theta_2}{r} + 4 w_{iiii}\Theta_2 + w_{iijj}(k+2)\Theta_2\right]$$

$$+\frac{\partial}{\partial \theta}[3 w_{iijj}\Theta_3] = 0$$

$$0 = \frac{2\theta(r+k-1)}{r}\left[\Theta_1 + \frac{\Theta_2}{r}\right] - 2 w_{ii}(k+2)\left[\Theta_1 + \frac{\Theta_2}{r}\right] . \qquad [4.3.2]$$

Solving equation [4.3.2] for θ, the Min-$\mathbf{B}^*$ optimal value of θ becomes

$$\theta = \frac{r w_{ii}(k+2)}{(r+k-1)} . \qquad [4.3.3]$$

The second partial derivative of $\mathbf{B}^*$ with respect to θ is a positive quantity, and therefore the optimal value of θ minimizes $\mathbf{B}^*$. The Min-$\mathbf{B}^*$ value of θ can be obtained with an appropriate choice of the scaling factor, g. For the CCDs, BBDs, and FACs, the Min-$\mathbf{B}^*$ value of g (as shown in equation [4.2.7] on page 177) is

$$g = \sqrt{\frac{\theta(F+2\alpha^2)}{(F+2\alpha^4)}} \qquad \text{for the CCDs}$$

$$g = \sqrt{\theta} \qquad \text{for the BBDs and FACs}$$

where θ is the value of [iiii]/[ii] which minimizes $\mathbf{B}^*$, F is the number of factorial design points, and α is the level of the axial design points (shown in Table 5 on page 178).

The Min-B^* values of g for the second order designs are shown in Tables 13 and 14 for spherical and cuboidal regions of interest, respectively, and Figure 40 on page 221 graphically displays the optimal values of the scaling factor. The optimal design variable levels are smaller than the ± 1 extremes typically used in practice. In addition, the optimal values of g for the MSE of slope criteria are slightly larger than those for the MSE of response criteria, and the optimal values for a cuboidal region are slightly larger than those for a spherical region.

Table 13. Min-B^* Values of the Scaling Factor in a Spherical Region of Interest. Optimal values of g for the CCDs, BBDs, FACs, and SCDs.

DESIGN CLASS	Min-B^* Values of g					
	k=2	k=3	k=4	k=5	k=6	k=7
CCD	.71	.57	.49	.44	.41	.38
BBD	-	-	.71	.71	-	.58
FAC	-	.65	.58	.52	.48	.45
SCD	-	.59	.50	.45	.41	.38

Table 14. Min-B^* Values of the Scaling Factor in a Cuboidal Region of Interest. Optimal values of g for the CCDs, BBDs, FACs, and SCDs.

DESIGN CLASS	Min-B^* Values of g					
	k=2	k=3	k=4	k=5	k=6	k=7
CCD	.82	.74	.70	.68	.66	.65
BBD	-	-	1.0	1.1	-	1.0
FAC	-	.84	.82	.80	.78	.77
SCD	-	.76	.71	.68	.66	.65

The bias component of J^* derived in this section applies to the CCD, BBD, and FAC design classes under the IR, CR, and AR correlation induction strategies. Because the SCDs have some non-zero odd order design moments, the Min-B^* values of g for these designs were approximated using the value of θ in equation [4.3.3] on page 219.

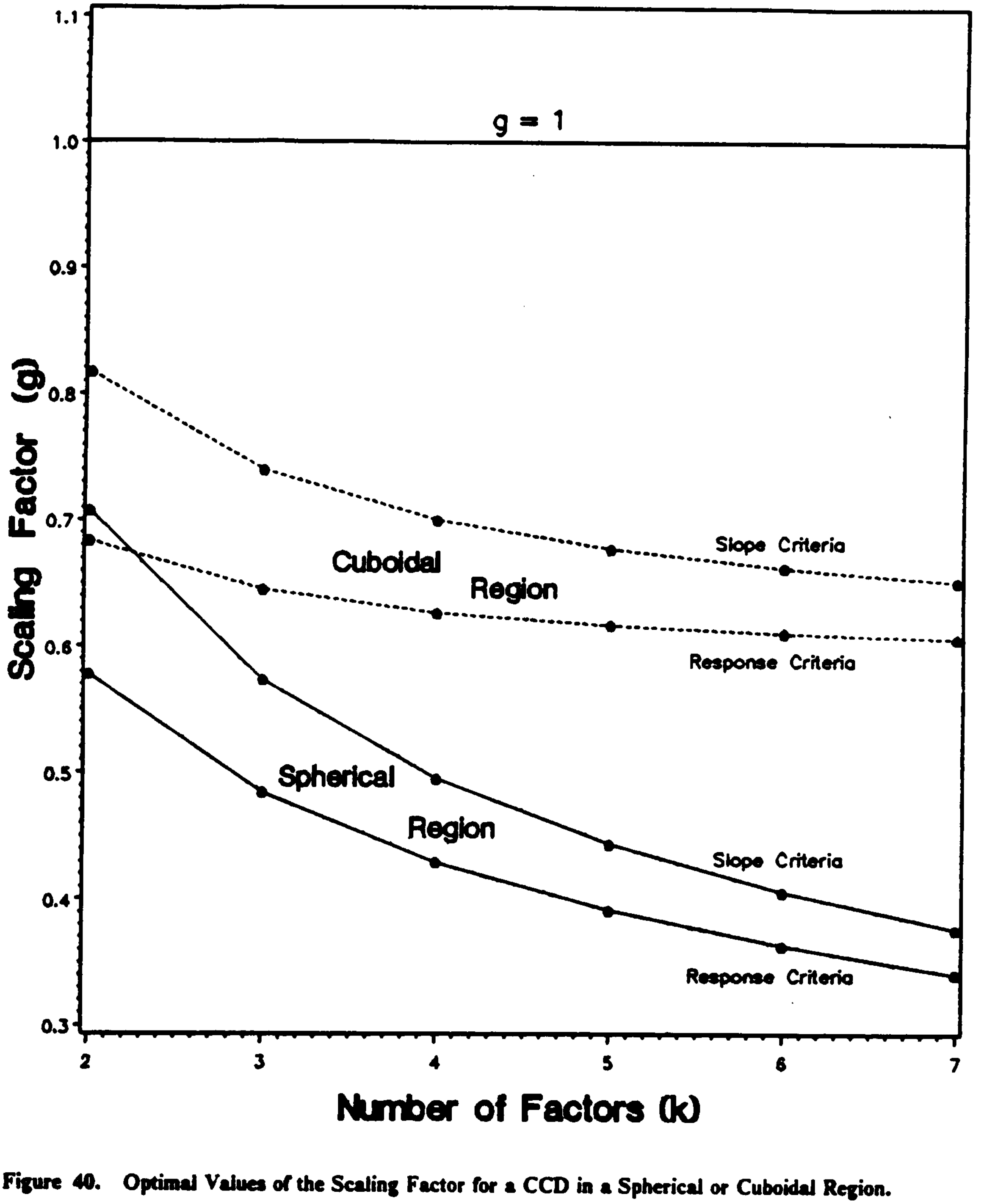


Figure 40. Optimal Values of the Scaling Factor for a CCD in a Spherical or Cuboidal Region.

The horizontal line at $g = 1$ indicates the value of the scaling factor which is typically used in practice.

The Response Criteria values of g are those required for a Min-B design.

The Slope Criteria values of g are those required for a Min-B* design.

4.3.2 Variance Component of J* in the Second Order Case

The variance component of **J*** is computed as the average variance of the partial derivative of the fitted response function, normalized with respect to N and σ^2. Myers and Lahoda (48) present a mathematical development of **V*** for situations in which the responses are uncorrelated and the designs are rotatable. Hussey, Myers and Houck (29) expand on the previous work by allowing for correlated responses and non-rotatable designs, and these authors develop the equations for **V*** for CCDs and BBDs under the pure **CR** and **AR** strategies. This research expands on the work of Hussey et.al. in several ways (as noted on page 182).

Myers and Lahoda (48) define the **V*** component, as shown in equation [2.2.55] on page 77, as

$$\mathbf{V}^* = \frac{N\Omega_r}{\sigma^2} \int_R \mathrm{Var}\,[\,\hat{\underline{y}}_{(\underline{x})}\,]\,d\underline{x}$$

$$= \frac{N}{\sigma_2} \;\mathrm{Trace}\,\{\mathrm{Var}\,[\,\underline{b}_1\,]\;\mu^*_{11}\}$$

where Ω_r^{-1} is the volume of the region of interest (defined on page 71), $\mathrm{Var}\,[\,\underline{b}_1\,]$ is the variance-covariance matrix of the least squares estimators of the fitted model coefficients (defined on page 72), and μ^*_{11} is the region matrix of the partial derivatives of the fitted model terms (defined on page 77).

The variance error, **V***, depends on the shape of the region of interest through μ^*_{11}. For the second order designs examined in this research, the $(p_1 \times p_1)$ μ^*_{11} matrix becomes

$$\mu_{11}^{*} = \left[\begin{array}{c|c|c|c} 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I}_k & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & 4\,w_{ii}\,\mathbf{I}_k & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & 2\,w_{ii}\,\mathbf{I}_{\binom{k}{2}} \end{array}\right]$$

where the $\mathbf{0}$'s indicate null matrices, the $\mathbf{I}$'s indicate identity matrices, and w_{ii} is a region moment of the design (defined on pages 289 and 291).

The $\mathbf{V}^{*}$ component is obtained by taking the *trace* of the product of the μ_{11}^{*} and Var [$\underline{b}_1$]. For the IR, CR, and AR correlation induction strategies, the component parts of the Var [$\underline{b}_1$] matrices are shown in Appendix M (pages 381-388) and the equations for $\mathbf{V}^{*}$ are derived in Appendix N (pages 389-395). Only the results of the derivations in the Appendix are presented here. Under the IR strategy, V is equal to the identity matrix, and the equation for $\mathbf{V}^{*}_{IR}$ becomes

$$\begin{aligned} \mathbf{V}^{*}_{IR} &= N \ \text{Trace}\,\{ (X'_1\,X_1)^{-1}\ \mu_{11}^{*} \} \\ &= \frac{k}{[ii]} + \frac{4\,b\,k\,w_{ii}}{D_2} + \frac{k\,(k-1)\,r\,w_{ii}}{\theta\,[ii]} \end{aligned} \qquad [4.3.4]$$

where b and D_2 are terms in the $(X'_1\,X_1)^{-1}$ matrix (defined on page 184) and r and θ are ratios of design moments (defined in equations [4.2.1] and and [4.2.2] on page 174).

Under the CR strategy, two forms of the variance error are relevant. For the designs utilizing less than two center runs, the equation for $\mathbf{V}^{*}$ under the *pure* CR strategy becomes

$$V^*_{pure\ CR} = (1-\rho_+)\, V^*_{IR} \qquad [4.3.5]$$

and for the designs utilizing two or more center runs, the equation for **V*** under the *modified* CR strategy becomes

$$V^*_{mod.\ CR} = (1-\rho_+)\, V^*_{IR} \;+\; 4\,k\,w_{ii}\,N\,\rho_+\,[\,t^2 + r^2\,] \qquad [4.3.6]$$

where t and r (defined on page 185) are elements of the Var [$\underline{b}_1$] matrix.

Under the AR strategy, four forms of the variance error are relevant. For the designs which partition into two orthogonal blocks and utilize less than three center runs, the equation for **V*** under the *pure* AR strategy becomes

$$V^*_{pure\ AR} = (1-\rho_+)\, V^*_{IR} \qquad [4.3.7]$$

which is identical to equation [4.3.5] for the *pure* CR strategy.

The second form of the **V*** component under the AR strategy applies to the designs which partition into two orthogonal blocks and utilize more than two center runs. The equation for **V*** under the *modified* AR strategy becomes

$$V^*_{mod.\ AR} = (1-\rho_+)\, V^*_{IR} \qquad [4.3.8]$$

$$+\; 4\,k\,w_{ii}\Big[\; \tfrac{1}{2}\,N\,t^2\,(\rho_+ - \rho_-) \;+\; N\,r^2\,\rho_+ \;+\; \tfrac{1}{2}\,N\,m^2\,(\rho_+ + \rho_-)\;\Big]$$

where t, r, and m (defined on page 185) are elements of the Var [$\underline{b}_1$] matrix.

The third form of V^* under the AR strategy applies to the FACs. These designs partition into three orthogonal blocks and are not augmented with center runs. The equation for V^*_{AR} for the 3^k designs becomes

$$V^*_{AR,\,3fac} = (1 - h\,\rho_+)\,V^*_{IR} \qquad [4.3.9]$$

where $h = 2\,/\,(\text{number of blocks})$ is a term which takes the additional design block into account.

The fourth form of V^* under the AR strategy applies to the $k = 4$ BBD. This design partitions into three orthogonal blocks and all of the center runs are placed in the 3rd block. The equation for V^*_{AR} for the $k = 4$ BBD becomes

$$V^*_{AR,\,bbd4} = (1 - \rho_+)\,V^*_{IR} + 8\,w_{ii}\,N\,t^2\,(\rho_+ - \rho_-) \qquad [4.3.10]$$

$$+ \rho_+\left[\frac{4}{3\,[ii]} + \frac{4\,r\,w_{ii}}{\theta\,[ii]} + 16\,N\,w_{ii}\,e\right]$$

where t, r, and m (defined on page 185) are elements of the Var $[\,\underline{b}_1\,]$ matrix.

The variance component of J^* under the IR strategy, the pure and modified CR strategies, the pure, modified, and 3-block AR strategies, as shown in equations [4.3.4]-[4.3.10], indicate that V^* is a function of the following parameters and constants:

$$V^* = f\,(\theta\,,[ii]\,,r\,,k\,,w_{ii}\,,\rho_+\,,\rho_-\,,N\,,N_{1a}\,,N_{1b})\ .$$

The values of r, k, N, N_{1a}, and N_{1b} are fixed for a given design plan, w_{ii} is fixed for a given region shape, ρ_+ and ρ_- are fixed for a given simulation study, and θ is fixed for a Min-B^* design. Therefore, [ii] is the only term in the equations for V^* which can be altered to minimize V^*. The value of [ii] which results in a Min-V^* | Min-B^* design is determined by setting the partial derivative of V^* with respect to [ii] equal to zero. The partial derivatives of $V^*_{pure\,CR}$, $V^*_{pure\,AR}$, and $V^*_{AR,\,3fac}$ result in the same optimal values of [ii] as V^*_{IR} because the $(1 - \rho_+)$ and $(1 - h\,\rho_+)$

coefficients of V_{IR} in equations [4.3.5], [4.3.7], and [4.3.9] disappear when the partial derivatives are set equal to zero. However, the additional terms in $V^*_{mod.\,CR}$, $V^*_{mod.\,AR}$, and $V^*_{AR,\,bbdt}$ are functions of [ii] through the Δ, e, m, Γ, and t terms. Similar to the optimal designs for the MSE of response criteria, the values of [ii] which are optimal for the IR strategy are approximately optimal for the CR and AR strategies. (Justification for the use of the V^*_{IR} optimal values of [ii] is given on page 188.)

The V*-optimal value of [ii] can be obtained by setting the partial derivative of V^*_{IR} with respect to [ii] equal to zero, yielding

$$\frac{\partial V^*}{\partial [ii]} = \frac{\partial}{\partial \theta}\left[\frac{k}{[ii]} + \frac{k(k-1)\,w_{ii}\,r}{\theta\,[ii]} + (4\,k\,w_{ii})\,\frac{\theta\,(r+k-2) - r\,(k-1)\,[ii]}{\theta^2\,(r-1)\,(r+k-1)\,[ii] - r\,k\,\theta\,(r-1)\,[ii]^2}\right]$$

$$0 = k\,\theta + a_1 + 4\,k\,\theta\,w_{ii}\left[\frac{b_1\,d_1 - 2\,b_1\,e_1\,[ii] + c_1\,e_1\,[ii]^2}{d_1^{\,2} - 2\,d_1\,e_1\,[ii] + e_1^{\,2}\,[ii]^2}\right] \qquad [4.3.11]$$

where b_1, c_1, d_1, and e_1 are defined on page 189, and

$$a_1 = k\,(k-1)\,w_{ii}\,r$$

$$f_1 = \frac{k\,\theta + a_1}{4\,k\,\theta\,w_{ii}} \; .$$

Collecting the $[ii]^2$, $[ii]^1$, and $[ii]^0$ terms, the quadratic equation for the partial derivative of V^*_{IR} set equal to zero becomes

$$0 = g_1\,[ii]^2 + h_1\,[ii]^1 + i_1\,[ii]^0$$

where $g_1 = c_1\,e_1 + f_1\,e_1^{\,2}$

$h_1 = -2\,b_1\,e_1 - 2\,d_1\,e_1\,f_1$

$i_1 = b_1\,d_1 + f_1\,d_1^{\,2}$.

Applying the quadratic formula, the Min-V* | Min-B* value of [ii] becomes

$$[\text{ii}] = \frac{-h_1 \pm \sqrt{h_1^2 - 4 g_1 i_1}}{2 g_1} \qquad [4.3.12]$$

where the "$-$" root is the **V***-optimal value of [ii] because the "$+$" root is infeasible for the designs examined in this research. (See Figure P-2 on page 422 for a graphical display of **V*** versus [ii] and N_c.) The optimal values of [ii] can be obtained through an appropriate choice for the number of center runs. For the CCDs, BBDs, and FACs, the Min-**V*** | Min-**B*** value of N_c (as shown in equation [4.2.18] on page 190) becomes

$$N_c = \frac{g^2 (F + 2\alpha^2)}{[\text{ii}]} - (F + n_\alpha) \qquad \text{for the CCDs}$$

$$N_c = \frac{f F g^2}{[\text{ii}]} - F \qquad \text{for the BBDs and FACs}$$

where g is the Min-**B*** value of the scaling factor, [ii] is the Min-**V*** | Min-**B*** value of the pure second order design moment, and f is the fraction of factorial design points in which a particular design variable is equal to zero (defined on page 190).

The Min-**V*** | Min-**B*** values of N_c for the second order designs are shown in Table 15 on page 228, and the exact, non-integer values of N_c are given in Table P-2 on page 419. Unlike the results for the MSE of response criteria, the optimal number of center runs for the MSE of slope criteria are the same for spherical and cuboidal regions. The values of N_c for the FACs have been rounded down to $N_c = 0$ because the addition of one center run (the optimal number) to these large designs has almost no effect on the value of [ii], and the equation for $\mathbf{V}^*_{AR}$ becomes considerably more complicated when center runs are added to the FACs. The optimal values of N_c for the SCDs are approximated using equation [4.3.12] for the values of [ii].

Table 15. Min-V*| Min-B* Number of Center Runs in a Spherical or Cuboidal Region. Optimal values of Nc for the CCDs, BBDs, FACs, and SCDs.

DESIGN CLASS	Min-V*| Min-B* Values of Nc					
	k=2	k=3	k=4	k=5	k=6	k=7
CCD	3	3	4	5	5	7
BBD	-	-	4	5	-	5
FAC	-	0	0	0	0	0
SCD	-	2	3	3	3	3

The next section presents an evaluation of the IR, CR, and AR correlation induction strategies using the Min-V* | Min-B* second order designs.

4.3.3 Min-V* | Min-B* Second Order Designs

The optimal second order designs for the MSE of slope criteria are Min-V* | Min-B* designs, resulting in minimum values of V*, given minimum values of B*. The Min-B* designs are achieved through the use of the scaling factors specified in Tables 13 and 14 on page 220, and the Min-V* | Min-B* designs are achieved through the addition of the number of center runs specified in Table 15. The required levels of the axial design points (prior to scaling) for the CCDs and SCDs are shown in Table 5 on page 178.

Similar to the developments for the MSE of response criteria, the research of this section uses the variance error as the basis for evaluating the correlation induction strategies. The justification for evaluating the strategies using V*, instead of $J^* = V^* + B^*$, is identical to the justification for using V instead of J (see page 193). Similar to the results for B in Table 10 on page 194, Table 16 on page 229 indicates that the values of B* are fairly constant from one design class to the next.

Table 16. Values of B^* for the $k=4$ Second Order Designs in a Spherical Region. Θ_1, Θ_2, and Θ_3 are standardized measures of cubic curvature. The second order designs are Min-V^*\| Min-B^* designs.

DESIGN CLASS	Values of B^*:						
	$\Theta_1=100$ $\Theta_2=0$ $\Theta_3=0$	$\Theta_1=0$ $\Theta_2=100$ $\Theta_3=0$	$\Theta_1=0$ $\Theta_2=0$ $\Theta_3=100$	$\Theta_1=100$ $\Theta_2=100$ $\Theta_3=0$	$\Theta_1=100$ $\Theta_2=0$ $\Theta_3=100$	$\Theta_1=0$ $\Theta_2=100$ $\Theta_3=100$	$\Theta_1=100$ $\Theta_2=100$ $\Theta_3=100$
CCD	60	21	6	82	67	28	88
BBD	62	21	6	83	69	27	90
FAC	79	15	6	94	85	22	101

Note: The SCDs are not included in the table because B^* is a function of the individual β coefficients in β_2 (not just the Θ_1, Θ_2, and Θ_3 parameters).

The research presented here is an evaluation of the three correlation induction strategies and the four second order design classes using the Min-V^* | Min-B^* second order designs. The CR strategy is evaluated in section 4.3.3.1 and the AR strategy is evaluated in section 4.3.3.2.

4.3.3.1 Min-V^* | Min-B^* Designs Under the CR Strategy

The Min-V^* | Min-B^* second order designs are used to evaluate the performance of the CR strategy relative to the IR strategy in this section. The values of V^* for each strategy are computed using equations [4.3.4], [4.3.5], and [4.3.6] on pages 223-224. The equation for $V^*_{mod.CR}$ reduces to $V^*_{pure\,CR}$ when the number of center runs is less than two and both V^*_{CR} equations reduce to V^*_{IR} when $\rho_+=0$. The preferred strategy (IR or CR) and the preferred design class (CCD, BBD, FAC, or SCD) are those which result in the least amount of variance error, V^*.

The correlation induction strategies can be evaluated by viewing plots of V^*_{CR} versus ρ_+ for each of the second order design classes. The IR strategy is indicated whenever $\rho_+=0$, and the CR strategy is indicated whenever $\rho_+ > 0$. Figures 41-44 on pages 231-234 illustrate the values of V^*_{CR} for the CCDs, BBDs, FACs, and SCDs, respectively. (The data used to generate these plots is shown in Appendices R and S on pages 448 and 462). The values of V^*_{CR} decrease as the magnitude of ρ_+ increases, indicating that the CR strategy is preferable to the IR strategy. Thus, for the second order designs considered here, the CR strategy is preferable to the IR strategy when the design criteria is minimization of the MSE of slope.

In addition to evaluating the performance of the correlation induction strategies, the performance of the four design classes are compared by viewing plots of V^*_{CR} versus ρ_+ for the second order designs. Figure 45 on page 235 illustrates the values of V^*_{CR} and Figure 46 on page 236 shows the difference between the values of V^*_{IR} and V^*_{CR}. Figure 45 indicates that the CCDs and BBDs are the preferred design classes when $\rho_+ < .4$, and the FACs are preferred when $\rho_+ > .4$. The SCDs are the least preferred design class under both the IR and CR strategies. Figure 46 indicates that the performance of the SCDs improves rapidly as ρ_+ increases but the performance of the CCDs and BBDs improves slowly. Thus, for the MSE of slope criteria, the CCDs and BBDs are the preferred design classes for the CR strategy when $\rho_+ < .4$, and the FACs are the preferred design class when $\rho_+ > .4$.

The *line* convention used in the figures which follow is shown on page 196.

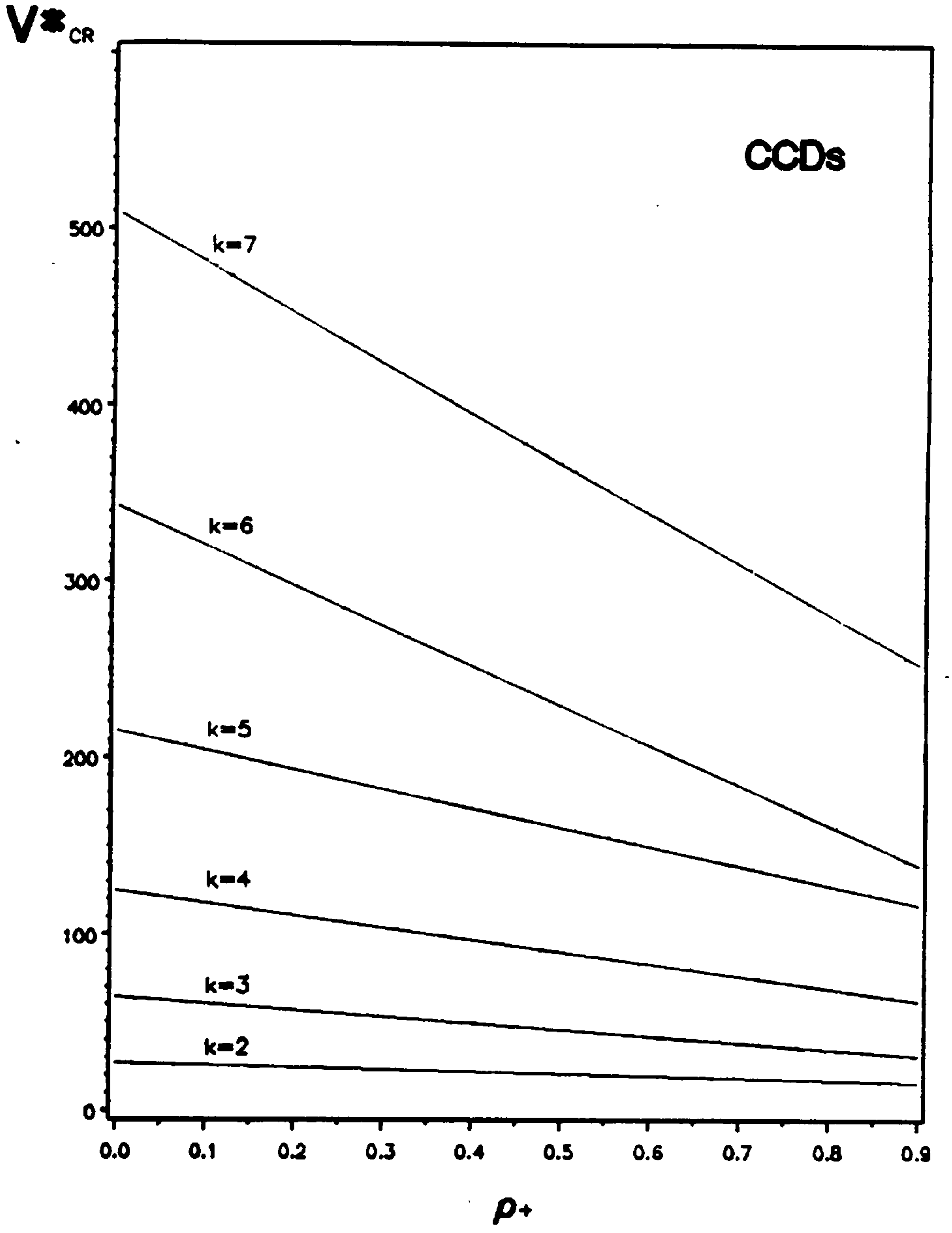


Figure 41. Optimal Values of V* for the CCDs under the CR Strategy.
The CCDs are Min-V* | Min-B* designs.
The region of interest is Spherical.

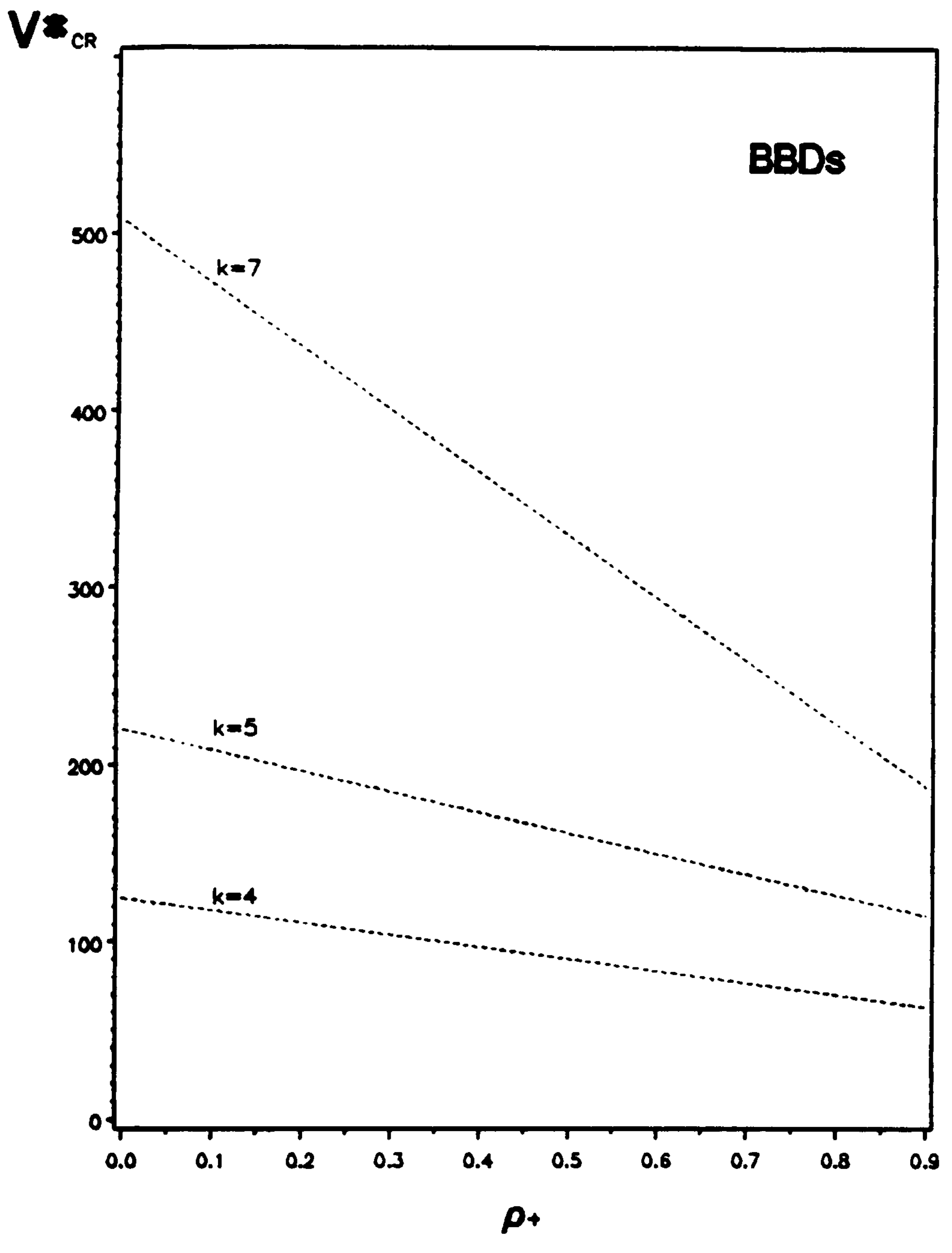


Figure 42. Optimal Values of V* for the BBDs under the CR Strategy.
The BBDs are Min-V* | Min-B* designs.
The region of interest is Spherical.

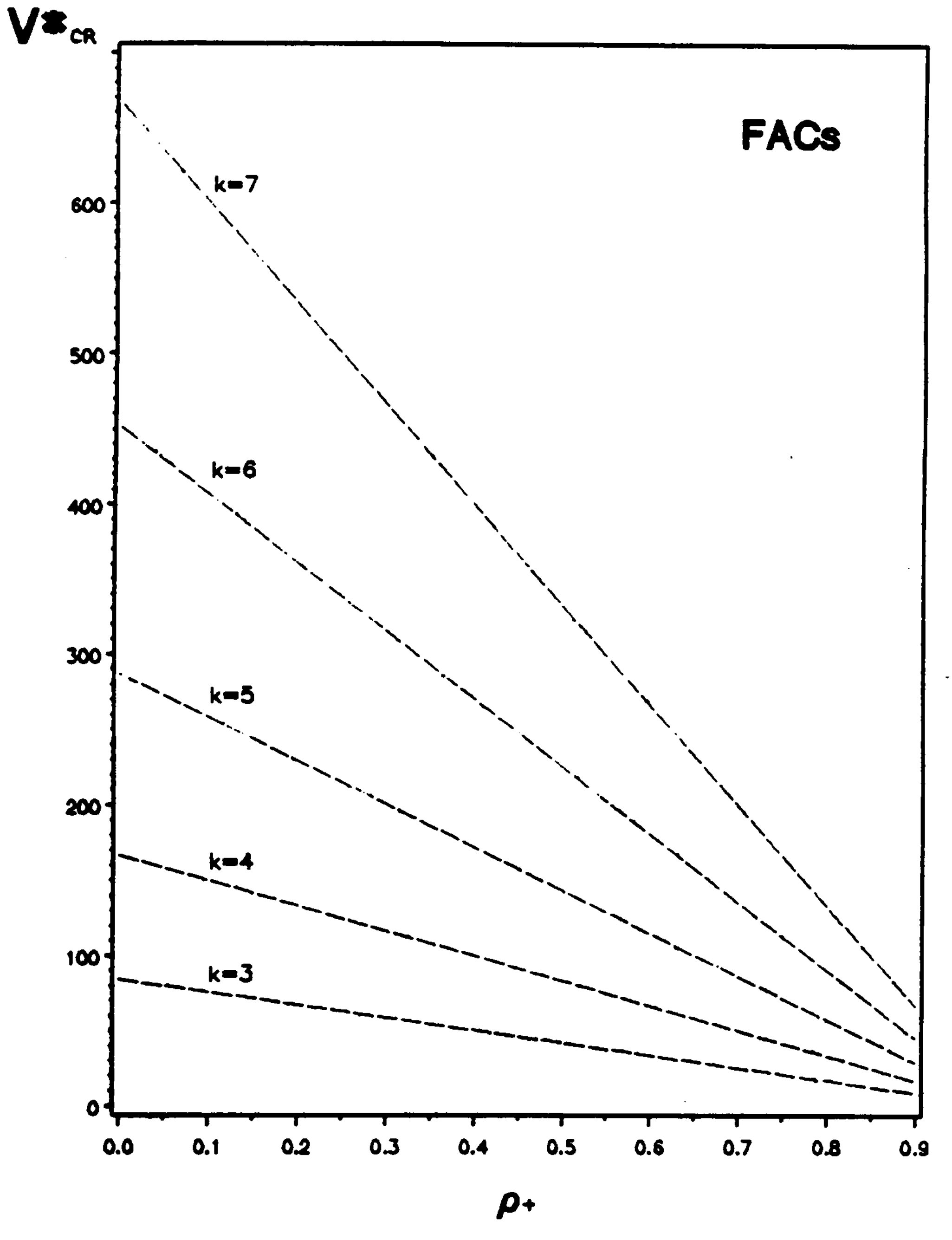


Figure 43. **Optimal Values of V^* for the 3-level Factorial Designs under the CR Strategy.**
The FACs (3^k designs) are Min-V^* | Min-B^* designs.
The region of interest is Spherical.

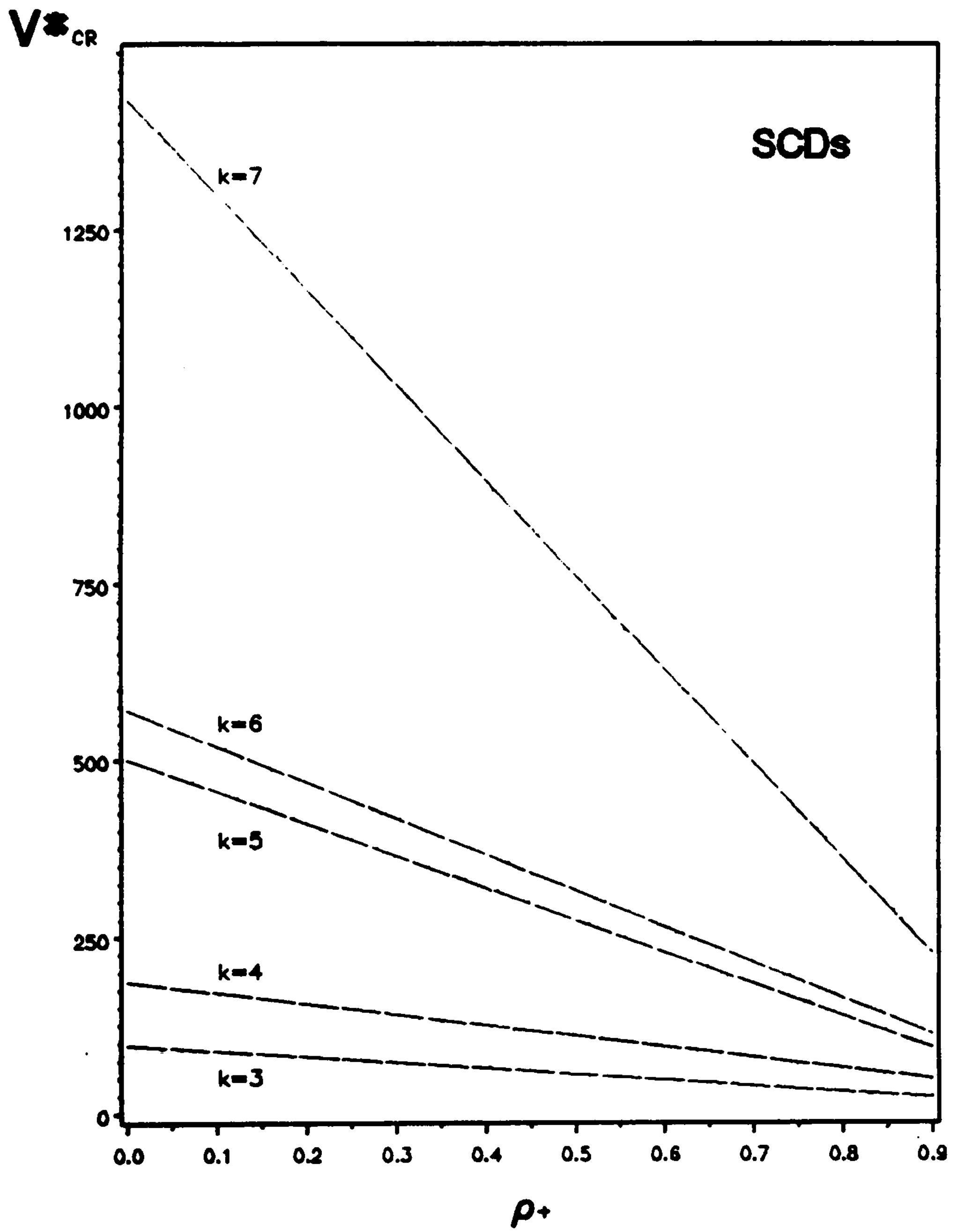


Figure 44. Optimal Values of V* for the Small Composite Designs under the CR Strategy.
The SCDs (small composite designs) are Min-V* | Min-B* designs.
The region of interest is Spherical.

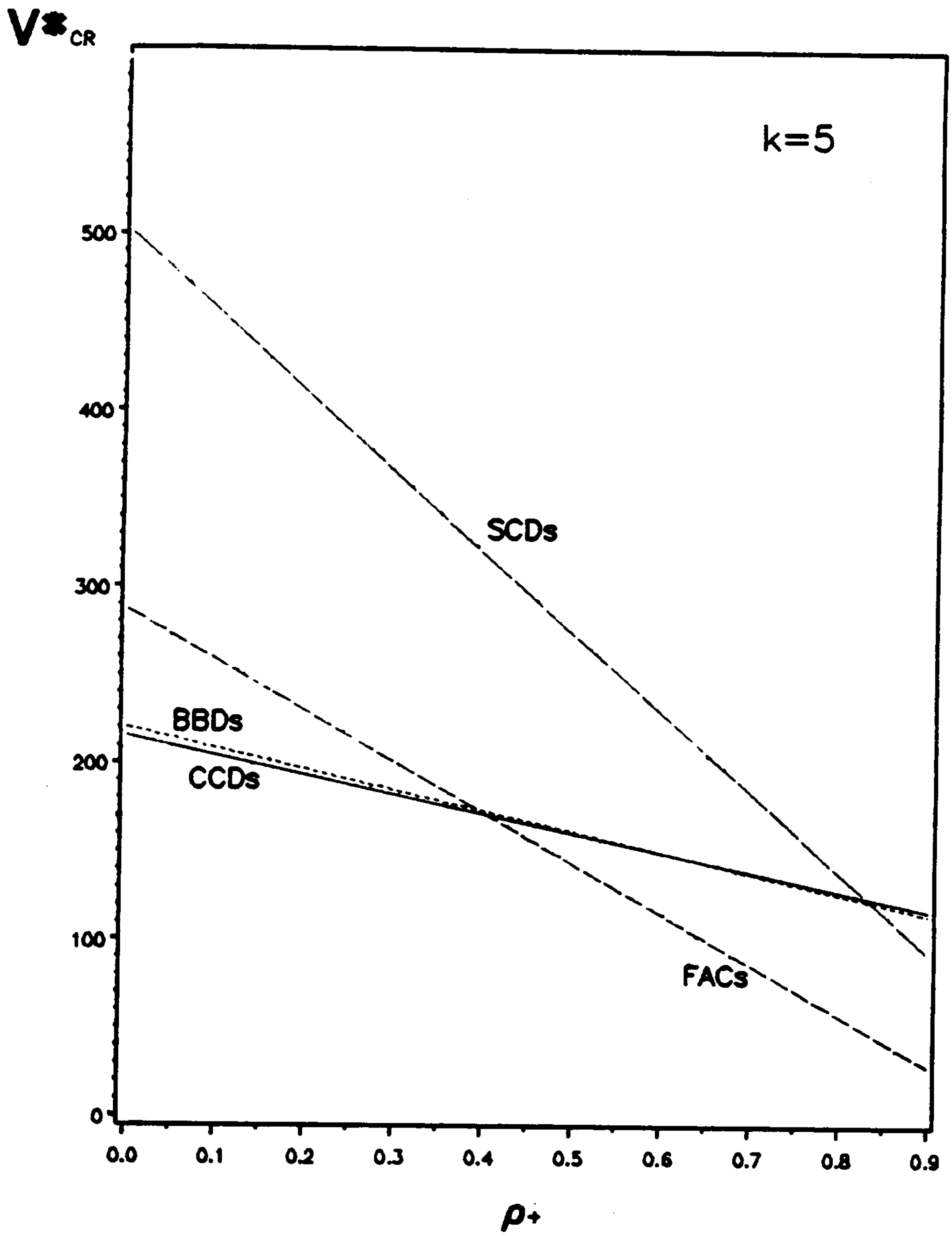


Figure 45. **Optimal Values of V* for a k=5 Second Order Design under the CR Strategy.**
The optimal k=5 second order designs are Min-V* | Min-B* designs.
The region of interest is Spherical.

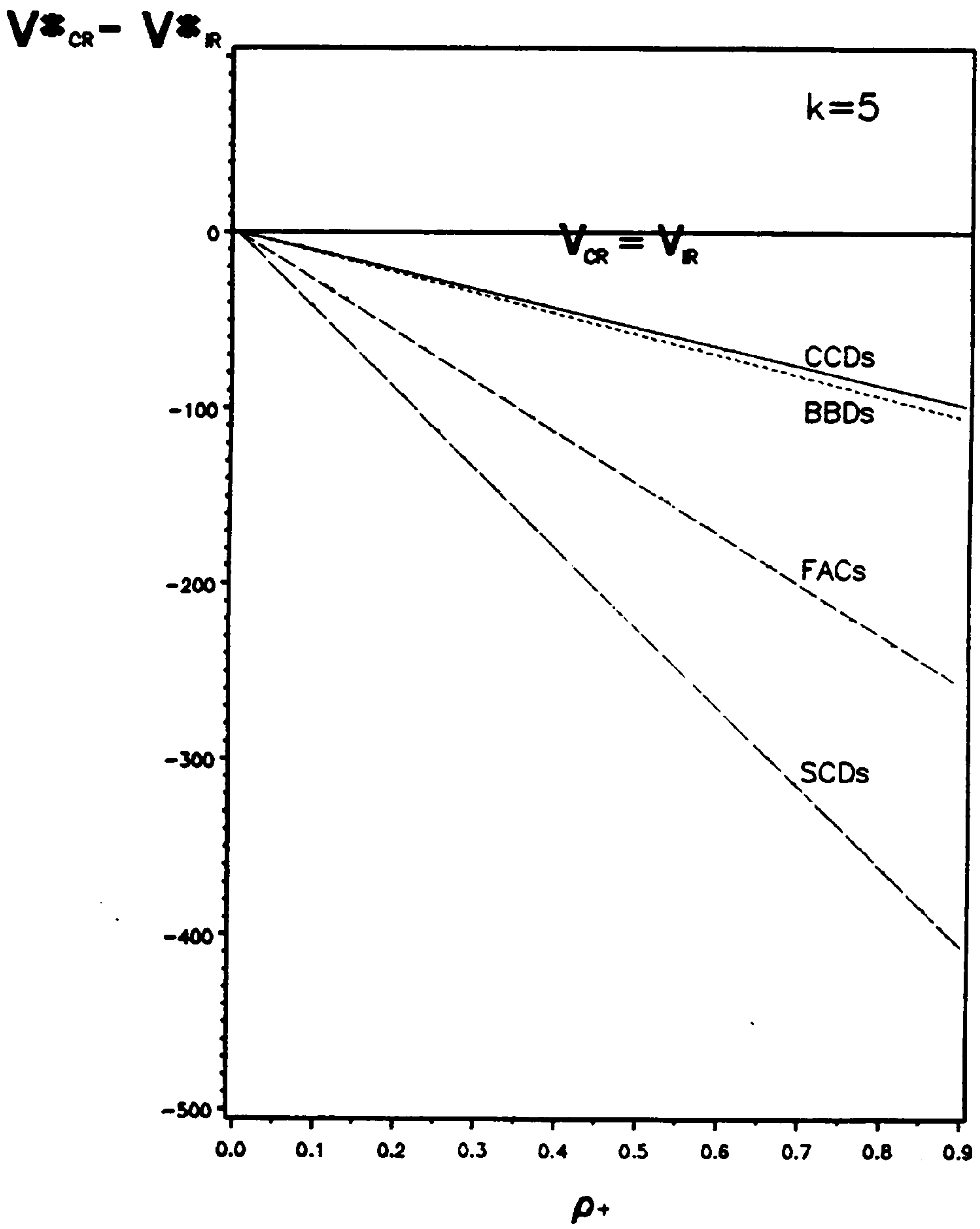


Figure 46. Differences Between the Optimal Values of V* under the CR and IR Strategies.
The optimal $k=5$ second order designs are Min-V* | Min-B* designs.
The region of interest is Spherical.

4.3.3.2 *Min*-$\mathbf{V}^*$ | *Min*-$\mathbf{B}^*$ *Designs Under the AR Strategy*

The performance of the **AR** strategy relative to the **IR** and **CR** strategies is evaluated in this section. The values of $\mathbf{V}^*_{\mathrm{AR}}$ are computed using equations [4.3.7], [4.3.8], [4.3.9], and [4.3.10] on pages 224-225. The equation for $\mathbf{V}^*_{mod.\mathrm{AR}}$ reduces to $\mathbf{V}^*_{pure\,\mathrm{AR}}$ when the number of center runs is less than three, the equation for $\mathbf{V}^*_{bbd4,\mathrm{AR}}$ reduces to $\mathbf{V}^*_{3fac,\mathrm{AR}}$ when the number of center runs is zero, and all of the $\mathbf{V}^*_{\mathrm{AR}}$ equations reduce to $\mathbf{V}^*_{\mathrm{IR}}$ when $\rho_+ = \rho_- = 0$. The preferred strategy (**IR**, **CR**, or **AR**) and the preferred design class (CCD, BBD, FAC, or SCD) are those which result in the least amount of variance error, $\mathbf{V}^*$.

The correlation induction strategies are evaluated by viewing plots of $\mathbf{V}^*_{\mathrm{AR}}$ versus ρ_+ for various magnitudes of ρ_-. The values of $\mathbf{V}^*_{\mathrm{AR}}$ are indicated by shaded triangular areas on the plots, with the upper boundaries of the shaded areas corresponding to $R = 0$, and the lower boundaries corresponding to $R = 1$ (where $R = \rho_- / \rho_+$ is defined in equation [4.2.19] on page 203). Figure 47 on page 239 illustrates the values of $\mathbf{V}^*$ for the $k = 2$ CCD under the **IR**, **CR**, and **AR** correlation induction strategies. The figure indicates that **IR** strategy results in the largest values of $\mathbf{V}^*$ and the **AR** strategy results in the smallest values of $\mathbf{V}^*$. Also, the values of $\mathbf{V}^*_{\mathrm{CR}}$ become smaller as the magnitude of ρ_+ increases, and the values of $\mathbf{V}^*_{\mathrm{AR}}$ become smaller as the magnitudes of both ρ_+ and ρ_- increase (as R approaches 1). Figure 48 on page 240 illustrates the values of $\mathbf{V}^*$ for the $k = 5$ CCD in spherical and cuboidal regions of interest. The values of $\mathbf{V}^*$ are lower in a cuboidal region, but the relative performance of the strategies is the same for both regions.

The **AR** and **IR** correlation induction strategies are compared by viewing plots of $\mathbf{V}^*_{\mathrm{AR}}$ versus ρ_+ for various magnitudes of ρ_-. When $\rho_+ = \rho_- = 0$, the results correspond to the **IR** strategy, and when $\rho_+ \geq \rho_- > 0$, the results correspond to the **AR** strategy. Figures 49-52 on pages 241-244 illustrate the values of $\mathbf{V}^*_{\mathrm{AR}}$ for the CCDs, BBDs, FACs, and SCDs, respectively. (The data used to generate these plots is shown in Appendices R and S on pages 450 and 462). The values of

$\mathbf{V}^*_{AR}$ decrease as the magnitudes of ρ_+ and ρ_- increase, indicating that the AR strategy is preferable to the IR strategy for the second order designs considered here.

In addition to evaluating the correlation induction strategies, the performances of the four design classes under the AR strategy are compared by viewing plots of $\mathbf{V}^*_{AR}$ versus ρ_+ for various magnitudes of ρ_-. Figure 53 on page 245, which illustrates the values of $\mathbf{V}^*_{AR}$ for the $k=5$ designs, indicates that the CCDs and BBDs are the preferred design classes and the SCDs are the least preferred. When compared to the results for the MSE of response criteria (Figures 38 and 39 on pages 213-214), the values of $\mathbf{V}^*_{AR}$ are less influenced by the magnitude of ρ_-, as indicated by the smaller triangular areas in Figure 53. The SCDs are the least preferable design class for the AR strategy, but the performance of these designs improves as the magnitude of ρ_+ increases. Thus, similar to the results for the IR and CR strategies, the CCDs and BBDs are the preferred design classes for the AR strategy when the design criteria is minimization of the MSE of slope.

The *shading* convention used in the figures which follow is shown on page 196.

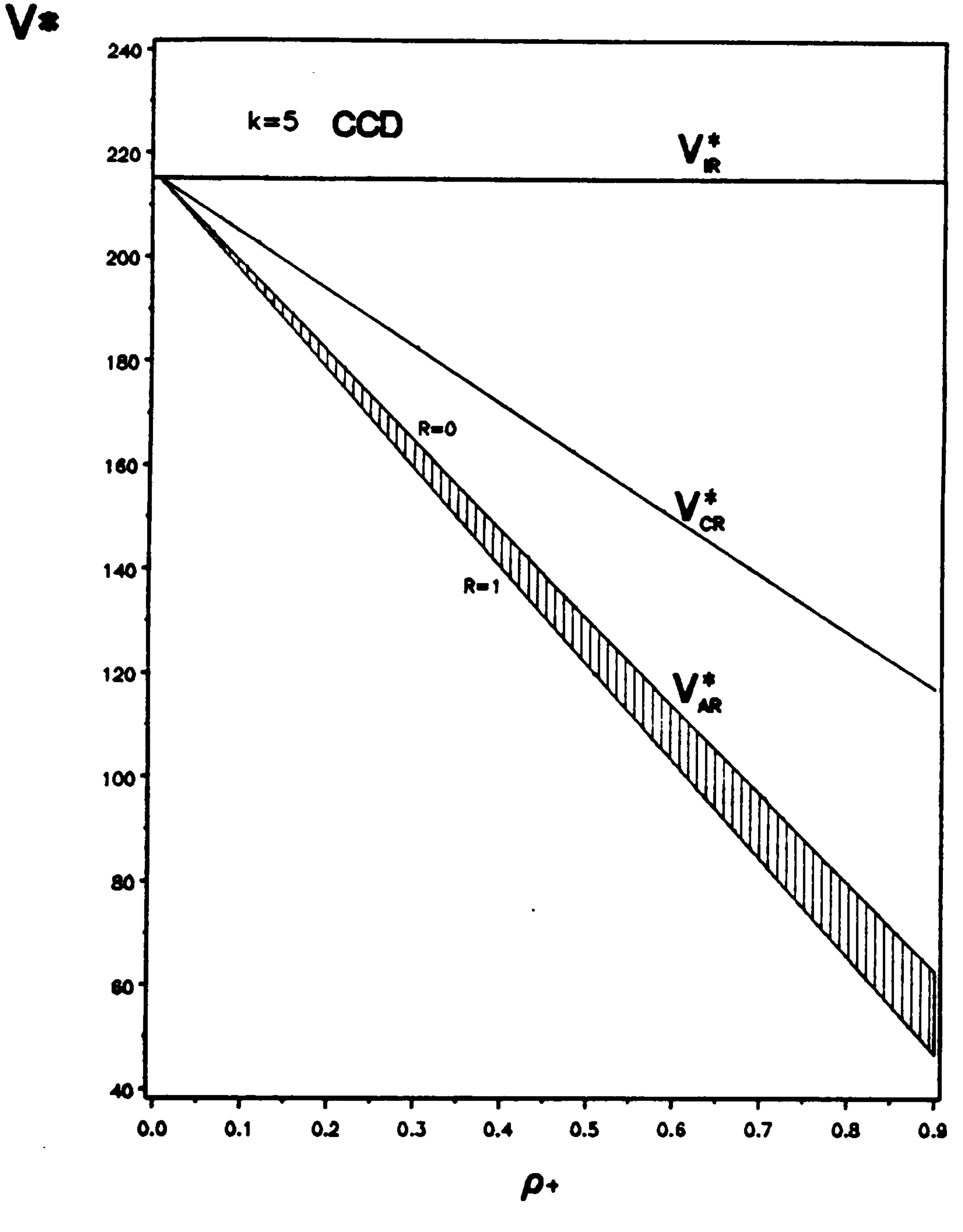


Figure 47. Optimal Values of V* for the k=5 CCD under the IR, CR, and AR Strategies.
The $k=5$ CCD is a Min-V* | Min-B* design.
The region of interest is Spherical.
The shaded V^*_{AR} region corresponds to values of $R=\rho_-/\rho_+$ between 0 and 1.
The AR and CR strategies always perform better than the IR strategy in this situation.

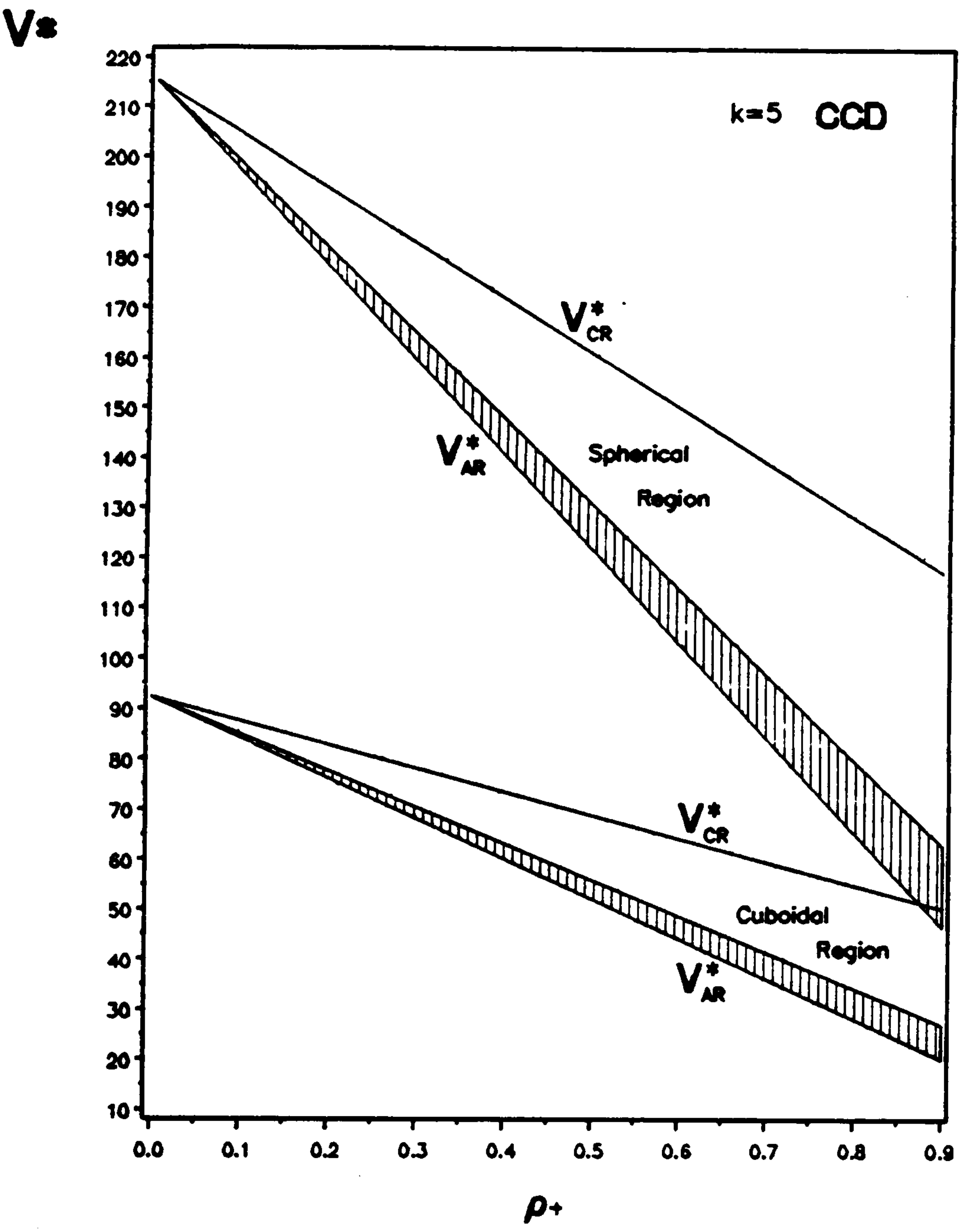


Figure 48. Optimal Values of V* for the k = 5 CCD in a Spherical or Cuboidal Region.
The $k=5$ CCD is a Min-V* | Min-B* design.
The shaded V^*_{AR} regions correspond to values of $R = \rho_- / \rho_+$ between 0 and 1.

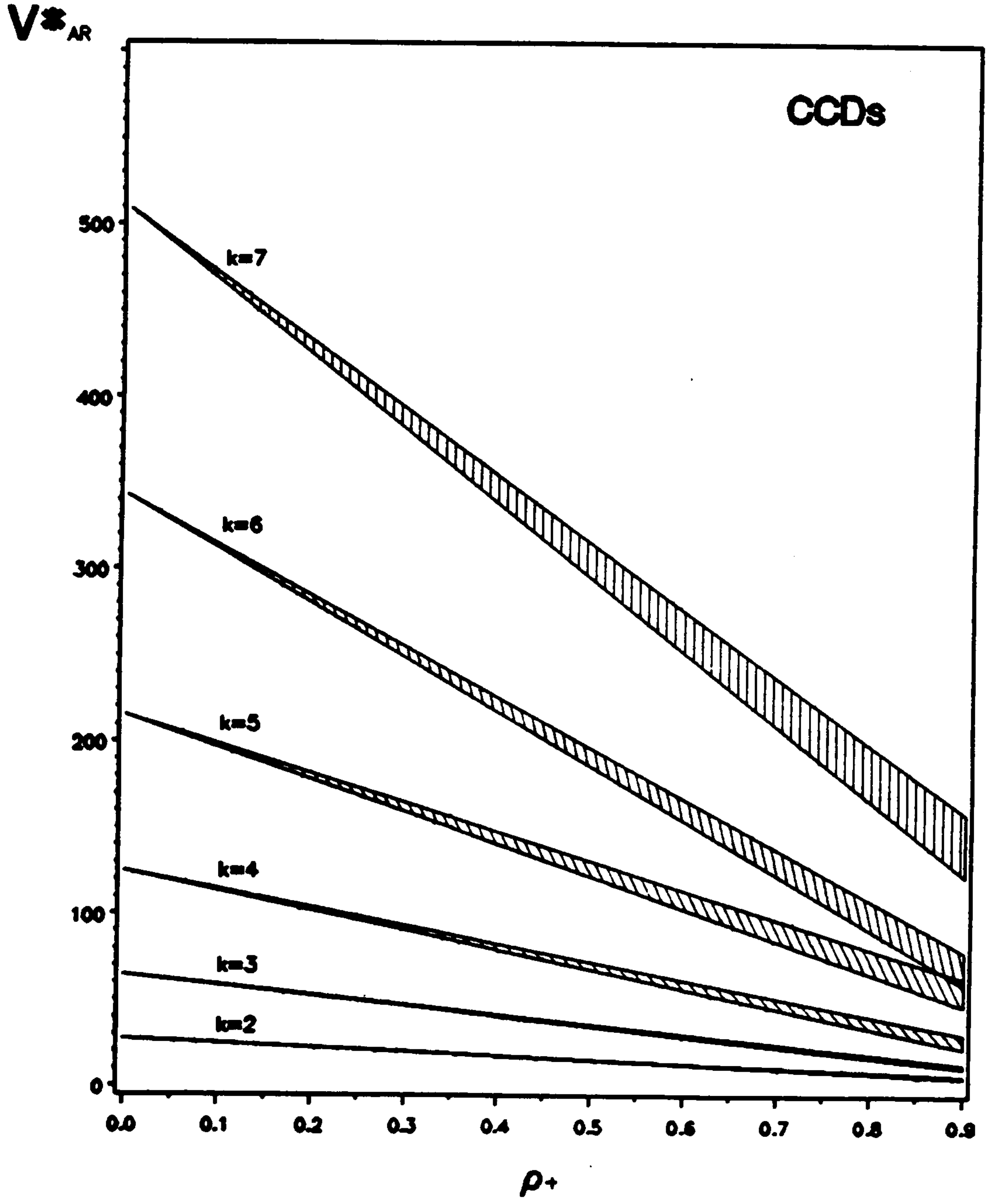


Figure 49. Optimal Values of V* for the CCDs under the AR Strategy.

The CCDs are Min-V* | Min-B* designs.

The region of interest is Spherical.

R is defined as: $R = \rho_- / \rho_+ = 1$ and $0 \leq R \leq 1$.

The upper triangular lines correspond to $R = 0$ and the lower correspond to $R = 1$.

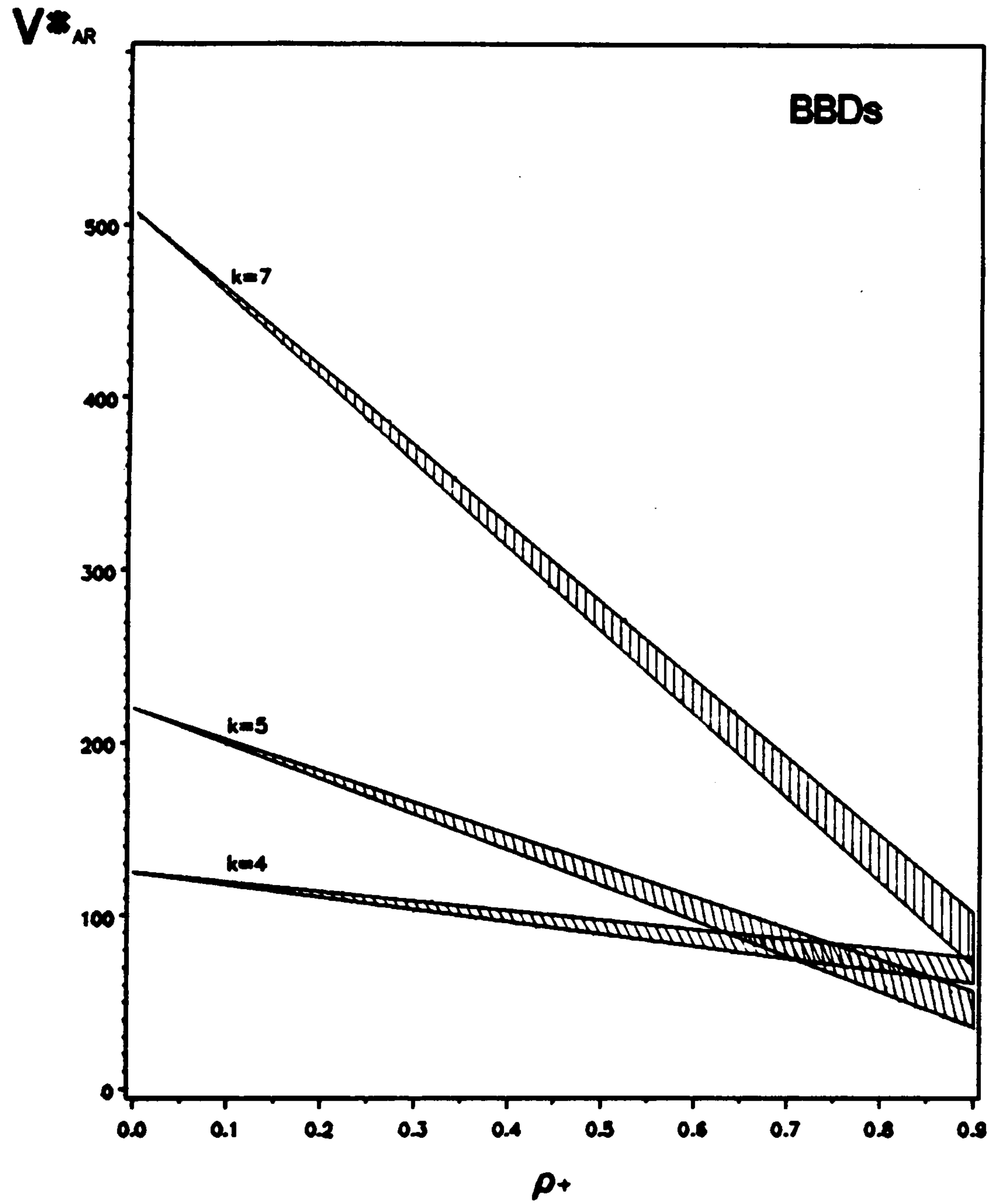


Figure 50. Optimal Values of V* for the BBDs under the AR Strategy.

The BBDs are Min-V* | Min-B* designs.

The region of interest is Spherical.

R is defined as: $R = \rho_- / \rho_+ = 1$ and $0 \leq R \leq 1$.

The upper triangular lines correspond to $R = 0$ and the lower correspond to $R = 1$.

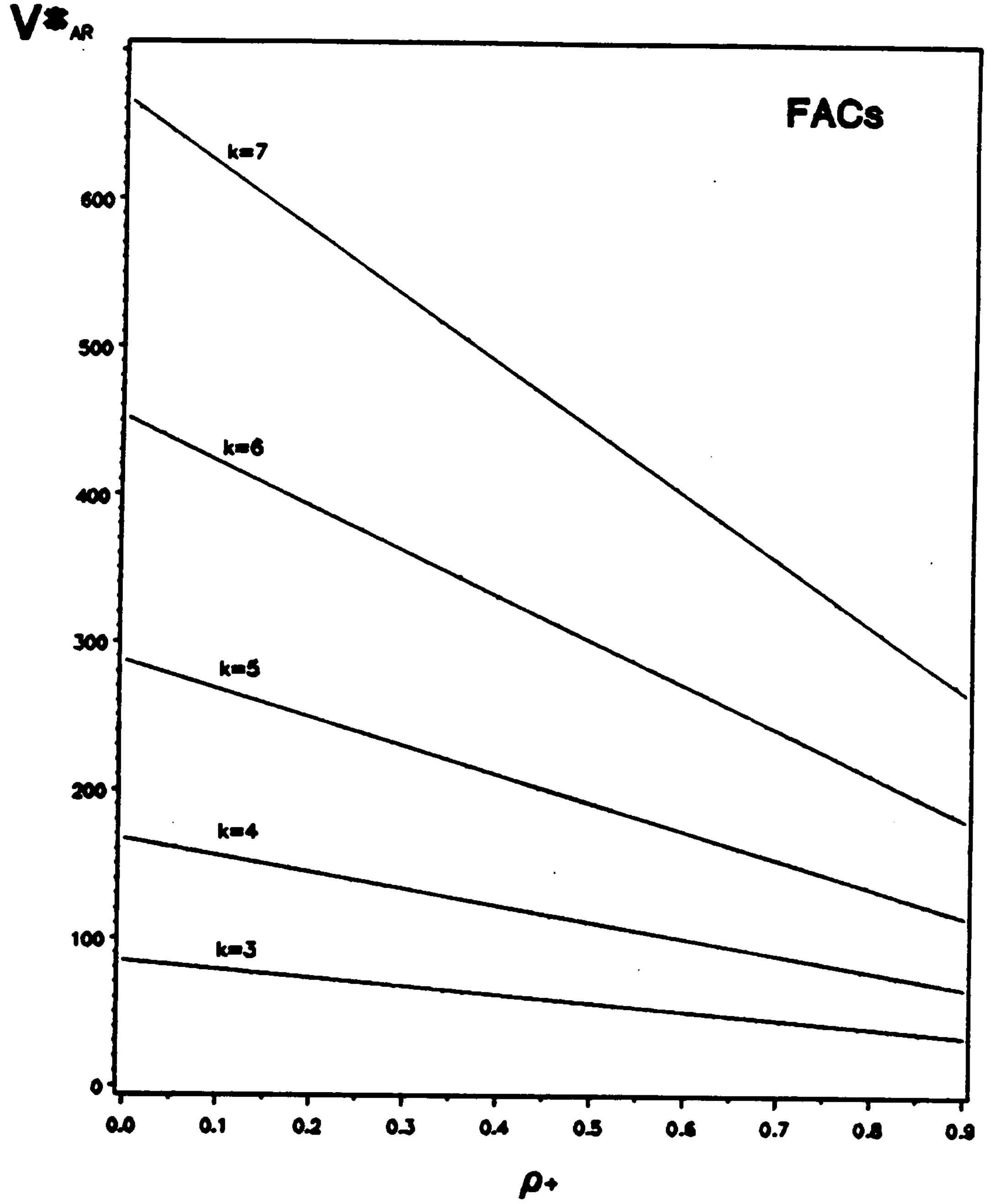


Figure 51. Optimal Values of V* for the 3-level Factorial Designs under the AR Strategy.

The FACs (3^k designs) are Min-V* | Min-B* designs.

The region of interest is Spherical.

R is defined as: $R = \rho_- / \rho_+ = 1$ and $0 \leq R \leq 1$.

The upper triangular lines correspond to $R = 0$ and the lower correspond to $R = 1$.

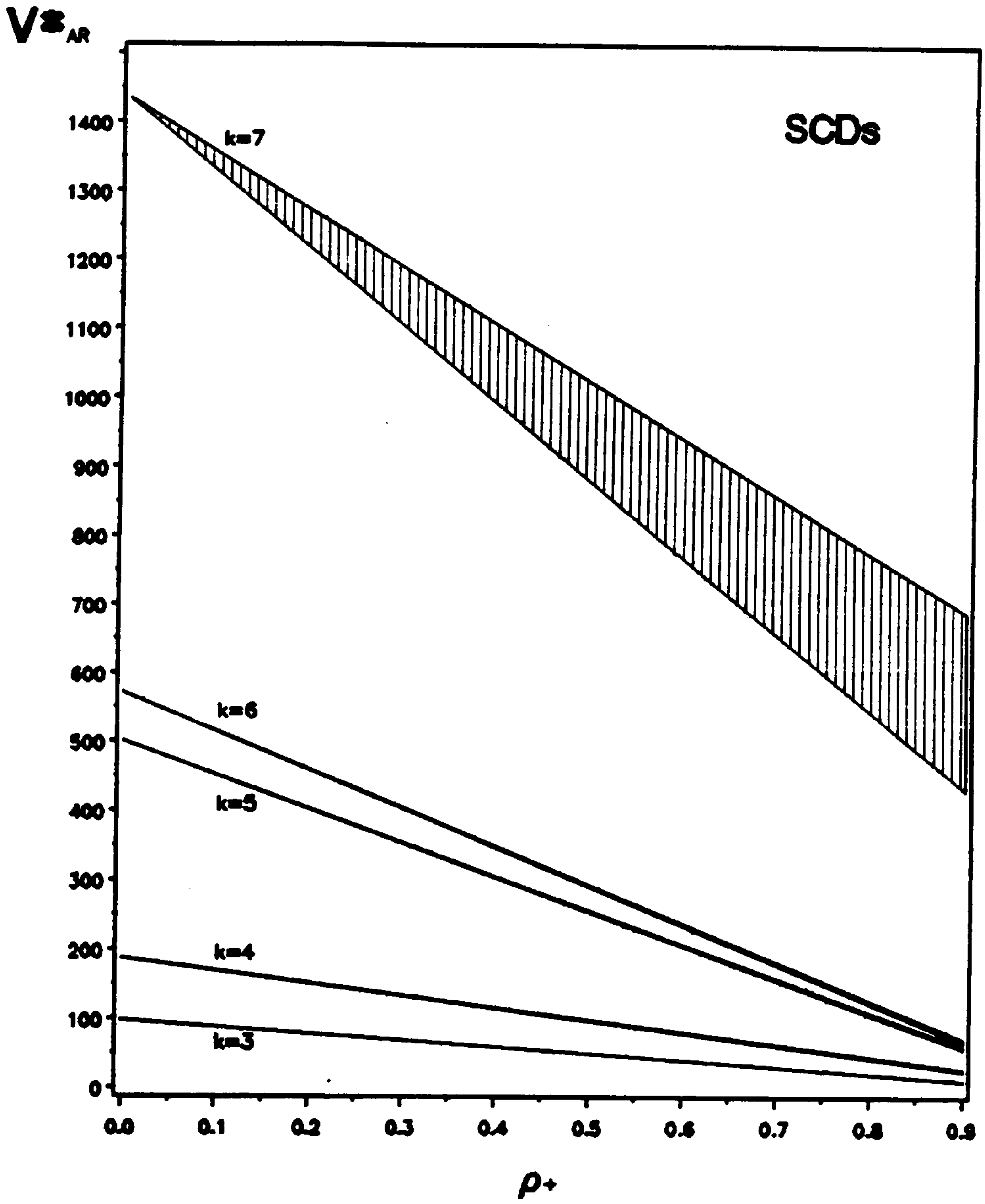


Figure 52. Optimal Values of V^* for the Small Composite Designs under the AR Strategy.

The SCDs (small composite designs) are Min-V^* | Min-B^* designs.

The region of interest is Spherical.

R is defined as: $R = \rho_- / \rho_+ = 1$ and $0 \leq R \leq 1$.

The upper triangular lines correspond to $R = 0$ and the lower correspond to $R = 1$.

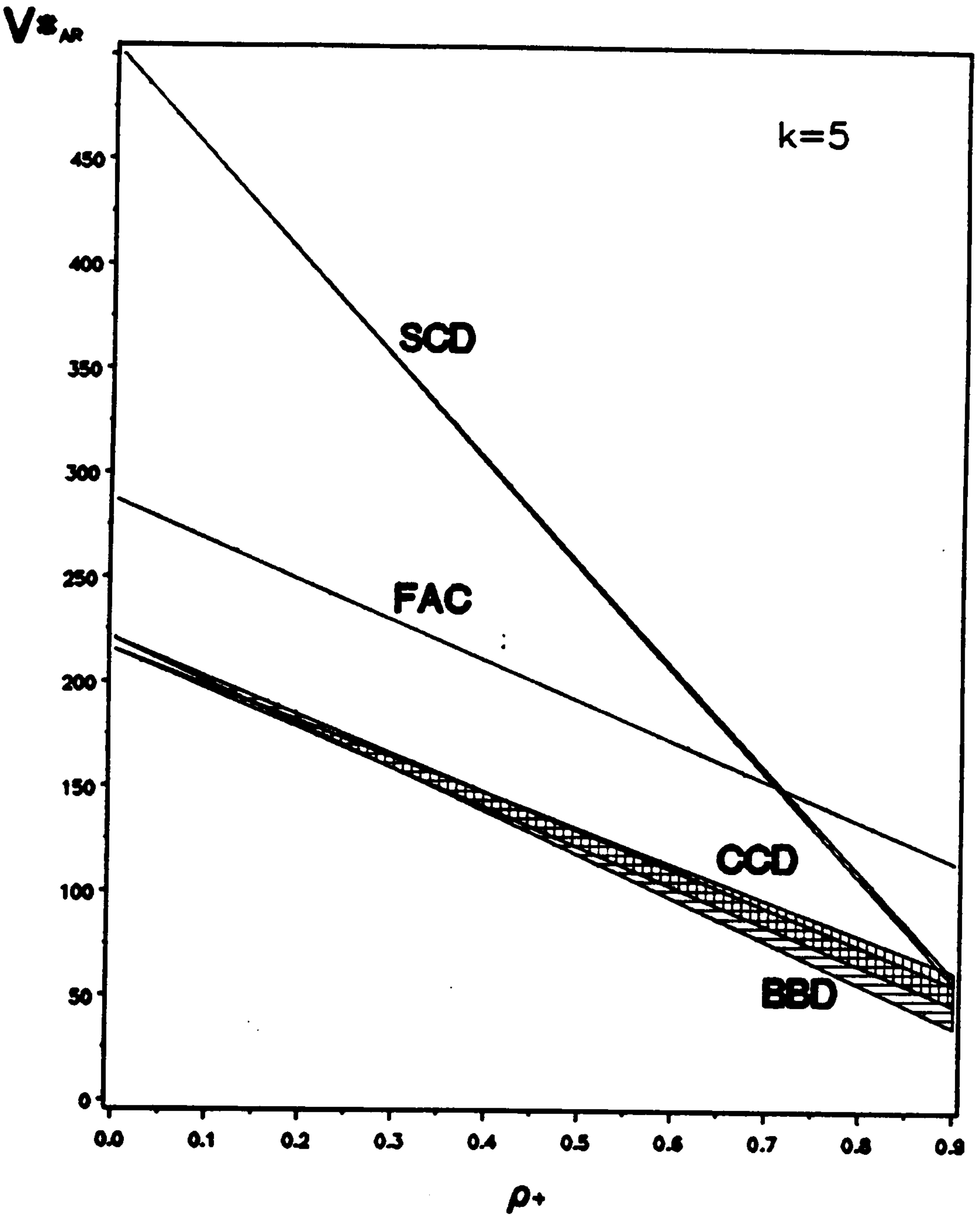


Figure 53. Optimal Values of V* for a k=5 Second Order Design under the AR Strategy.
The optimal designs are Min-V* | Min-B* designs.
The region of interest is Spherical.

4.4 Summary of Results in the Second Order Case

Three correlation induction strategies were evaluated in this Chapter for the case of fitting a second order polynomial model while protecting against third order bias. The three strategies included: the use of independent random number streams for each design point (IR); the use of common streams for each non-replicated design point (CR); and the use of common and antithetic streams for non-replicated design points in opposite blocks (AR). Two design criteria, MSE of response and MSE of slope, served as performance measures using four second order design classes; central composite designs (CCDs), Box-Behnken designs (BBDs), three-level factorial designs (FACs), and small composite designs (SCDs).

The optimal second order designs utilized the levels of the design variables which minimized the bias error ($\mathbf{B}$ or $\mathbf{B}^*$) and the number of center runs which minimized the variance error ($\mathbf{V}$ or $\mathbf{V}^*$) of the minimum bias designs. The optimal designs for the MSE of response criteria were termed Min-$\mathbf{V}$ | Min-$\mathbf{B}$ designs, and the optimal designs for the MSE of slope criteria were termed Min-$\mathbf{V}^*$ | Min-$\mathbf{B}^*$ designs. Tables 17 and 18 on pages 247-248 indicate the optimal number of center runs and the optimal placement of the design points for each of the second order designs considered. The optimal designs in Table 17 correspond to a spherical region of interest, and the optimal designs in Table 18 correspond to a cuboidal region. (The entire design matrices are shown in Appendix L on pages 363-380.)

The remainder of this section is a summary of the relative performance of the three correlation induction strategies in the second order case. Results for the MSE of response criteria are presented first, followed by the results for the MSE of slope criteria.

Table 17. Min-V | Min-B and Min-V* | Min-B* Second Order Designs in a Spherical Region.
Nc is the optimal number of center runs.
g is the level of the scaled *factorial* design points in the optimal designs.
αg is the level of the scaled *axial* points in the optimal CCDs and SCDs.

DESIGN	Min-V | Min-B Values of:			Min-V* | Min-B* Values of:		
	N_c	g	α g	N_c	g	α g
$k=2$ CCD	2	.5774	.8165	3	.7071	1.0000
$k=3$ CCD	2	.4840	.8538	3	.5727	1.0102
$k=4$ CCD	2	.4287	.8822	4	.4950	1.0187
$k=5$ CCD	3	.3916	.9033	5	.4435	1.0243
$k=6$ CCD	2	.3626	.9103	5	.4054	1.0177
$k=7$ CCD	3	.3401	.9243	7	.3760	1.0219
$k=4$ BBD	2	.6124	-	4	.7071	-
$k=5$ BBD	3	.6236	-	5	.7071	-
$k=7$ BBD	2	.5222	-	5	.5774	-
$k=3$ FAC	0	.5533	-	0	.6547	-
$k=4$ FAC	0	.5000	-	0	.5774	-
$k=5$ FAC	0	.4606	-	0	.5222	-
$k=6$ FAC	0	.4297	-	0	.4804	-
$k=7$ FAC	0	.4045	-	0	.4472	-
$k=3$ SCD	1	.5170	.8009	2	.5889	0.9854
$k=4$ SCD	1	.4194	.8896	3	.5000	1.0000
$k=5$ SCD	1	.3853	.9036	3	.4456	1.0041
$k=6$ SCD	1	.3585	.9140	3	.4065	1.0055
$k=7$ SCD	2	.3399	.9146	3	.3758	1.0112

- The $k=6, 7$ CCDs use $^1/_2$ fractional replicates for the factorial portions of the designs, the $k=6, 7$ FACs are $^1/_3$ fractional replicates of the 3^k designs, the $k=3, 4, 6$ SCDs are Hartley designs, and the $k=5, 7$ SCDs are Draper designs.
- Under the CR strategy, $N_c - 1$ (or zero) center runs use independent streams.
- Under the AR strategy, $N_c - 2$ (or zero) center runs use independent streams, except for the $k=4$ BBD, in which N_c center runs use independent streams.
- The single center run of the Min-V | Min-B $k=3$ SCD is placed in the factorial block, and the single center run of the $k=4, 5, 6$ SCDs are placed in the axial block.

Table 18. Min-V | Min-B and Min-V*| Min-B* Second Order Designs in a Cuboidal Region.
Nc is the optimal number of center runs.
g is the level of the scaled *factorial* design points in the optimal designs.
αg is the level of the scaled *axial* points in the optimal CCDs and SCDs.

DESIGN		Min-V | Min-B Values of :			Min-V* | Min-B* Values of :		
		N_c	g	α g	N_c	g	α g
k=2	CCD	2	.6831	0.9661	3	.8165	1.1547
k=3	CCD	2	.6446	1.1369	3	.7394	1.3042
k=4	CCD	3	.6261	1.2885	4	.7000	1.4406
k=5	CCD	4	.6167	1.4242	5	.6775	1.5646
k=6	CCD	3	.6103	1.5322	5	.6620	1.6619
k=7	CCD	4	.6063	1.6477	7	.6513	1.7699
k=4	BBD	3	.8944	-	4	1.0000	-
k=5	BBD	3	.9832	-	5	1.0801	-
k=7	BBD	3	.9309	-	5	1.0000	-
k=3	FAC	0	.7368	-	0	.8452	-
k=4	FAC	0	.7303	-	0	.8165	-
k=5	FAC	0	.7261	-	0	.7977	-
k=6	FAC	0	.7232	-	0	.7845	-
k=7	FAC	0	.7211	-	0	.7746	-
k=3	SCD	2	.6628	1.1091	2	.7603	1.2722
k=4	SCD	2	.6325	1.2649	3	.7071	1.4142
k=5	SCD	2	.6196	1.3962	3	.6807	1.5338
k=6	SCD	2	.6120	1.5138	3	.6639	1.6420
k=7	SCD	2	.6059	1.6305	3	.6508	1.7514

- The *k=6, 7* CCDs use ½ fractional replicates for the factorial portions of the designs, the *k=6, 7* FACs are ⅓ fractional replicates of the 3^k designs, the *k=3, 4, 6* SCDs are Hartley designs, and the *k=5, 7* SCDs are Draper designs.
- Under the CR strategy, $N_c - 1$ (or zero) center runs use independent streams.
- Under the AR strategy, $N_c - 2$ (or zero) center runs use independent streams, except for the *k=4* BBD, in which N_c center runs use independent streams.

The results for the MSE of response criteria, **J**, presented in section 4.2, indicated the following order of preference for the correlation induction strategies:

1. **AR** strategy — performs the best of the three strategies and its performance improves as the magnitudes of ρ_+ and ρ_- increase. However, for the $k=5, 7$ CCDs and $k \neq 3$ FACs, the **AR** strategy only performs better than the **IR** strategy when ρ_- / ρ_+ is larger than approximately ½. (See Tables 11 and 12 on page 205.)

2. **IR** strategy — performs better than the **CR** strategy, but not as well as the **AR** strategy, for the CCD, BBD, and FAC design classes. For the SCDs, however, the **CR** strategy performs better than the **IR** strategy.

3. **CR** strategy — performs the poorest of the three strategies for the CCDs, BBDs, and FACs, and its performance deteriorates as the magnitude of ρ_+ increases. For the SCDs, however, the **CR** strategy performs better than the **IR** strategy, and its performance improves as the magnitude of ρ_+ increases.

The results for the MSE of response criteria indicated the following order of preference for the Min-**V** | Min-**B** second order designs:

1. CCDs and BBDs — perform the best of the four design classes under the **IR** and **AR** correlation induction strategies. However, the SCDs perform slightly better under the **CR** strategy when $\rho_+ > .6$.
   - — the CCDs tend to perform better than the BBDs when the magnitude of ρ_+ is small and ρ_- / ρ_+ is near one.
   - — the BBDs tend to perform better than the CCDs when the magnitude of ρ_+ is large and ρ_- / ρ_+ is near zero.
   - — the $k=5, 7$ BBDs perform slightly better than the corresponding CCDs.

2. SCDs — perform better than the CCDs and BBDs under the **CR** strategy when $\rho_+ > .6$. However, the SCDs do not perform as well as the CCDs and BBDs under the **AR** and **IR** strategies.

3. FACs — perform the poorest of the four design classes under the **CR** and **AR** strategies, but perform better than the SCDs under the **IR** strategy.

Thus, in order to minimize the mean squared error of the predicted responses, **J**, the results of this research suggest the use of the **AR** correlation induction strategy and a Min-**V** | Min-**B** central composite or Box-Behnken design. The BBDs are recommended for fitting models with $k=5, 7$

factors, and the CCDs are recommended for models with $k=2, 3, 4, 6$ factors. The least desirable correlation induction strategy for minimization of **J** is the CR strategy, and the least desirable designs are the FACs.

The results for the MSE of slope criteria, **J***, presented in section 4.3, indicated the following order of preference for the correlation induction strategies:

1. AR strategy — performs the best of the three correlation induction strategies and its performance improves as the magnitudes of ρ_+ and ρ_- increase.
2. CR strategy — performs better than the IR strategy but not as well as the AR strategy, and its performance improves as the magnitude of ρ_+ increases.
3. IR strategy — performs the poorest of the three correlation induction strategies.

The results for the MSE of slope criteria indicated the following order of preference for the Min-**V*** | Min-**B*** second order designs:

1. CCDs and BBDs — perform the best of the four design classes under the IR and AR correlation induction strategies. However, the FACs perform slightly better under the CR strategy when $\rho_+ > .4$.
   - the CCDs tend to perform better than the BBDs when the magnitude of ρ_+ is small, and the BBDs tend to perform better than the CCDs when the magnitude of ρ_+ is large.
   - the $k=5, 7$ BBDs perform slightly better than the corresponding CCDs.
2. FACs — perform better than the CCDs and BBDs under the CR strategy when $\rho_+ > .4$. However, the FACs do not perform as well as the CCDs and BBDs under the AR and IR strategies.
3. SCDs — perform the poorest of the four design classes under the IR strategy. However, when $\rho_+ > .8$, the SCDs perform better than the FACs under the AR strategy and better than the CCDs and BBDs under the CR strategy.

Similar to the findings for the MSE of response criteria, the results for the MSE of slope criteria suggest the use of the AR correlation induction strategy and a Min-**V*** | Min-**B*** central composite or Box-Behnken design. The BBDs are recommended for fitting models with $k=5, 7$ factors, and the CCDs are recommended for models with $k = 2, 3, 4, 6$ factors. The least desirable correlation induction strategy for minimization of **J*** is the **IR** strategy, and the least desirable design class is the SCDs.

When taken together, the results for the MSE of response and MSE of slope criteria indicate that the preferred correlation induction strategy is the assignment rule blocking strategy and the preferred second order design classes are the central composite and Box-Behnken designs. The strategy of common random number streams and/or the use of a 3^k design are poor choices when prediction of the response important. The strategy of independent streams and/or the use of a small composite design are poor choices when estimation of the response function gradient is important.

The research findings suggest that the correlation induction strategies can be useful techniques for reducing the mean squared errors of the predicted responses and the slopes of the response function. The bias components, **B** and **B***, are unaffected by correlation induction under ordinary least squares estimation and, therefore, the reductions in MSE are achieved through the variance components, **V** and **V***. The correlation induction strategies lower the variance components by reducing the variances of the fitted model coefficients. Tables 19 and 20 on pages 252-253 show the variances of b_0, b_i, b_{ii}, and b_{ij} (the intercept, linear, quadratic, and two-way interaction coefficients). Table 19 presents the variances for the $k = 4$ Min-**V** | Min-**B** designs in a spherical region and Table 20 presents the coefficient variances for the $k = 5$ Min-**V*** | Min-**B*** designs in a cuboidal region. The percentage decreases ↓, or increases ↑, relative to the **IR** strategy are indicated in parentheses. For the **CR** and **AR** strategies, induced positive and negative correlations of $(\rho_+, \rho_-) = (.7, .2)$ and $(\rho_+, \rho_-) = (.3, .1)$ are used.

Tables 19 and 20 indicate that the FACs result in the lowest variances of the fitted model coefficients and the SCDs result in the highest variances. The large differences between the magnitudes of the coefficient variances for these designs are due to the large (small) number of design points in the FACs (SCDs), thereby yielding more (less) precise estimates of the β coefficients. However,

Table 19. Variances of the Fitted Model Coefficients for the Optimal k = 4 Second Order Designs.

The optimal designs are Min-V | Min-B designs.

The region of interest is Spherical.

b_0, b_i, b_{ii}, and b_{ij} are the intercept, linear, quadratic, and two-way interaction coefficients.

Correlation Induction Strategy	Variance of Model Coefficients	k = 4 Designs			
		CCD	BBD	FAC	SCD
IR	$Var(b_0)/\sigma^2$	.496	.500	.111	.946
	$Var(b_i)/\sigma^2$	.222	.222	.074	.334
	$Var(b_{ii})/\sigma^2$	1.57	1.63	.889	2.39
	$Var(b_{ij})/\sigma^2$	1.85	1.78	.444	4.04
CR $\rho_+ = .7$	$Var(b_0)/\sigma^2$	.499 (0.6% ↑)	.500 (no chg.)	.733 (560% ↑)	.984 (4% ↑)
	$Var(b_i)/\sigma^2$	.067 (70% ↓)	.067 (70% ↓)	.022 (70% ↓)	.100 (70% ↓)
	$Var(b_{ii})/\sigma^2$	1.08 (31% ↓)	1.11 (32% ↓)	.267 (70% ↓)	.716 (70% ↓)
	$Var(b_{ij})/\sigma^2$	.555 (70% ↓)	.533 (70% ↓)	.133 (70% ↓)	1.21 (70% ↓)
AR $\rho_+ = .7$ $\rho_- = .2$	$Var(b_0)/\sigma^2$	.441 (11% ↓)	.500 (no chg.)	.170 (53% ↑)	.535 (44% ↓)
	$Var(b_i)/\sigma^2$	.067 (70% ↓)	.119 (47% ↓)	.040 (47% ↓)	.100 (70% ↓)
	$Var(b_{ii})/\sigma^2$	.472 (70% ↓)	1.48 (9% ↓)	.474 (47% ↓)	.716 (70% ↓)
	$Var(b_{ij})/\sigma^2$	.555 (70% ↓)	.533 (70% ↓)	.237 (47% ↓)	1.21 (70% ↓)
CR $\rho_+ = .3$	$Var(b_0)/\sigma^2$	.497 (0.3% ↑)	.500 (no chg.)	.378 (240% ↑)	.962 (2% ↑)
	$Var(b_i)/\sigma^2$	.156 (30% ↓)	.156 (30% ↓)	.052 (30% ↓)	.234 (30% ↓)
	$Var(b_{ii})/\sigma^2$	1.36 (13% ↓)	1.41 (14% ↓)	.622 (30% ↓)	1.67 (30% ↓)
	$Var(b_{ij})/\sigma^2$	1.30 (30% ↓)	1.24 (30% ↓)	.311 (30% ↓)	2.83 (30% ↓)
AR $\rho_+ = .3$ $\rho_- = .1$	$Var(b_0)/\sigma^2$	.466 (6% ↓)	.500 (no chg.)	.133 (20% ↑)	.763 (19% ↓)
	$Var(b_i)/\sigma^2$	.156 (30% ↓)	.178 (20% ↓)	.059 (20% ↓)	.234 (30% ↓)
	$Var(b_{ii})/\sigma^2$	1.10 (30% ↓)	1.56 (4% ↓)	.711 (20% ↓)	1.67 (30% ↓)
	$Var(b_{ij})/\sigma^2$	1.30 (30% ↓)	1.24 (30% ↓)	.356 (20% ↓)	2.83 (30% ↓)

Table 20. Variances of the Fitted Model Coefficients for the Optimal k = 5 Second Order Designs.

The optimal designs are Min-V* | Min-B* designs.

The region of interest is Cuboidal.

b_0, b_i, b_{ii}, and b_{ij} are the intercept, linear, quadratic, and two-way interaction coefficients.

Correlation Induction Strategy	Variance of Model Coefficients	k = 5 Designs			
		CCD	BBD	FAC	SCD
IR	$Var(b_0)/\sigma^2$	.199	.200	.0045	.333
	$Var(b_i)/\sigma^2$	.051	.054	.0097	.191
	$Var(b_{ii})/\sigma^2$	.108	.090	.0046	.142
	$Var(b_{ij})/\sigma^2$	.148	.184	.0229	1.00
CR	$Var(b_0)/\sigma^2$	.200 (0.5% ↑)	.200 (no chg.)	.0084 (85% ↑)	.333 (no chg.)
$\rho_+ = .7$	$Var(b_i)/\sigma^2$	.015 (70% ↓)	.016 (70% ↓)	.0029 (70% ↓)	.052 (70% ↓)
	$Var(b_{ii})/\sigma^2$	.134 (24% ↑)	.130 (44% ↑)	.0014 (70% ↓)	.128 (10% ↓)
	$Var(b_{ij})/\sigma^2$	.044 (70% ↓)	.055 (70% ↓)	.0068 (70% ↓)	.300 (70% ↓)
AR	$Var(b_0)/\sigma^2$	.202 (1% ↑)	.184 (8% ↓)	.0035 (22% ↓)	.290 (13% ↓)
$\rho_+ = .7$	$Var(b_i)/\sigma^2$	.015 (70% ↓)	.016 (70% ↓)	.0052 (47% ↓)	.057 (70% ↓)
$\rho_- = .2$	$Var(b_{ii})/\sigma^2$	.071 (34% ↓)	.059 (35% ↓)	.0024 (47% ↓)	.062 (56% ↓)
	$Var(b_{ij})/\sigma^2$	.044 (70% ↓)	.055 (70% ↓)	.0122 (47% ↓)	.300 (70% ↓)
CR	$Var(b_0)/\sigma^2$	.199 (0.2% ↑)	.200 (no chg.)	.0062 (36% ↑)	.333 (no chg.)
$\rho_+ = .3$	$Var(b_i)/\sigma^2$	.036 (30% ↓)	.037 (30% ↓)	.0068 (30% ↓)	.134 (30% ↓)
	$Var(b_{ii})/\sigma^2$	.119 (10% ↑)	.107 (19% ↑)	.0032 (30% ↓)	.136 (4% ↓)
	$Var(b_{ij})/\sigma^2$	.104 (30% ↓)	.129 (30% ↓)	.0016 (30% ↓)	.700 (30% ↓)
AR	$Var(b_0)/\sigma^2$	.199 (.2% ↑)	.192 (4% ↓)	.0041 (10% ↓)	.312 (7% ↓)
$\rho_+ = .3$	$Var(b_i)/\sigma^2$	.036 (30% ↓)	.037 (30% ↓)	.0078 (20% ↓)	.134 (30% ↓)
$\rho_- = .1$	$Var(b_{ii})/\sigma^2$	.092 (15% ↓)	.076 (15% ↓)	.0037 (20% ↓)	.107 (24% ↓)
	$Var(b_{ij})/\sigma^2$	.104 (30% ↓)	.129 (30% ↓)	.0183 (20% ↓)	.700 (30% ↓)

large designs are penalized under the MSE criteria because **J** and **J*** are normalized with respect to the number of design points, *N*. Therefore, the magnitudes of the coefficient variances under the **CR** and **AR** strategies are less important than the percentage changes in the variances relative to the **IR** strategy.

The variances of the fitted model coefficients under the **CR** and **AR** strategies are compared to the variances under the **IR** strategy in Tables 19 and 20. These tables indicate that the intercept term (b_0) achieves the least variance reduction (if any), the linear and two-way interaction coefficients (b_i and b_{ij}) achieve the greatest variance reductions, and the quadratic coefficients (b_{ii}) generally achieve slight reductions in variance. Under the **CR** strategy, the variance of b_0 tends to increase relative to the **IR** strategy, the variances of b_i and b_{ij} are reduced by approximately $100\,\rho_+\,\%$, and the variance of b_{ii} is generally reduced slightly. Under the **AR** strategy, the variance of b_0 changes very little relative to the **IR** strategy, the variances of b_i and b_{ij} are reduced by approximately $100\,h\,\rho_+\,\%$ (where $h = \frac{2}{\#\ \text{blocks}}$), and the variance of b_{ii} is reduced slightly.

The findings of this Chapter (summarized on pages 249 and 250 for the MSE of response and MSE of slope criteria, respectively) indicate that the **AR** strategy performs better than the **CR** strategy in terms of both MSE design criteria. For the MSE of response criteria, the **IR** strategy performs better than the **CR** strategy, but for the MSE of slope criteria, the **CR** strategy performs better than the **IR** strategy. These relative performance findings for the three correlation induction strategies, under the two MSE criteria, are the result of the changes in the variances of the fitted model coefficients when correlation is induced among the responses.

Under the MSE of slope criteria, the variance of the intercept term (b_0) does not affect the relative performance of the three correlation induction strategies because b_0 is eliminated upon taking the partial derivative of the response function. Therefore, the **CR** strategy, which results in large variances of b_0, performs better than the **IR** strategy, which has larger variances of b_i and b_{ij}. For the MSE of response criteria, however, the large variances of b_0 lead to poor performance of the **CR** strategy relative to the **IR** strategy. The variances of b_0 are similar for the **IR** and **AR** strategies and, therefore, the superior performance of the AR strategy is the result of variance reductions in the linear (b_i) and two-way interaction (b_{ij}) coefficients. The variances of these coefficients are also reduced under the CR strategy, but the reductions are offset by large increases in the variance of b_0. The variances of the quadratic coefficients (b_{ii}) are generally reduced slightly under both the CR and AR strategies, thereby improving the performances of these strategies relative to the IR strategy.

For the designs which partition into three blocks, the reductions in the variances of b_i and b_{ij} under the AR strategy are less than those achieved by the two-block designs. This result is consistent with the findings on pages 249-250 which indicated that the $k=4$ CCD (two blocks) is preferable to the $k=4$ BBD (three blocks). In addition, the FACs are generally the least preferred design class under the AR strategy due to their three-block partitioning, their large number of design points, and their large increases in the variance of b_0 when correlation is induced.

In summary, the second order results presented here suggest that the preferred correlation induction strategy is the assignment rule blocking strategy and the preferred design classes are the central composite and Box-Behnken designs. Second order designs are generally used during the latter stages of the RSM optimization process, in conjunction with methods such canonical and ridge analysis, when the objective is prediction of the optimum response. Therefore, the AR strategy, which leads reduced variances of the prediction equation coefficients (relative to the IR strategy), would be particularly appealing during the final stages of the optimization process.

In the next Chapter, an overview of the findings in the first order and second order cases is presented. The general trends which have emerged for both first and second order designs are discussed in this final Chapter.

Chapter 5

CONCLUSIONS

This dissertation has investigated the performance of three strategies for assigning pseudorandom number streams to the stochastic components of simulation models. These strategies can affect reductions in the variances of certain parameter estimates by inducing correlation among the simulated responses. The findings are applicable to situations in which response surface methodology is being used to optimize a computer simulation model. The first section of this Chapter presents an overview of this dissertation and the following section summarizes the major findings and conclusions of this research investigation.

5.1 Overview

Computer simulation models are often used in the study of real-world systems when controlled experiments on the actual system cannot be performed. The random behavior of a computer simulation model is achieved through the assignment of pseudorandom number streams to the stochastic components of the simulation model. This research evaluated the performance of three assignment strategies in the context of a response surface optimization study.

It was assumed that a valid simulation model had been developed and response surface methodology (RSM) was the tool employed for identification of the optimum operating conditions. RSM is a collection of experimental design and statistical analysis procedures which provide experimenters with a framework for optimization of system response. Sequentially-fitted, low-order polynomial models are used to approximate the relationship between the response variable and a set of input variables within restricted regions of the factor space. In a simulation environment, the assignment of pseudorandom number streams to the design points is an additional experimental design consideration. Random number seed values, which generate streams of uniform (0,1) deviates, must be assigned to the stochastic model components. For any given simulation run, a different seed value is assigned to each stochastic model component. Consequently, the seed values assigned to these components are controlled by the experimenter. The three strategies for assigning sets of random number streams to the simulation runs are summarized below.

1. Independent streams (IR):

   - a different set of pseudorandom number streams (different seed values) is assigned to each simulation run, or equivalently, to each design point.

2. Common streams (CR):

   - the same set of streams (same set of seed values) is assigned to each non-replicated design point, and

   - different sets of streams are assigned to replicated design points.

3. Assignment rule blocking strategy (AR):

   - a common set of streams (same set of seed values) is assigned to each non-replicated design point in the first orthogonal block,

   - an antithetic set of streams is assigned to each non-replicated design point in the second block (same seed values, but the random numbers are "one minus" the uniform deviates of the first block), and

- different sets of streams are assigned to all design points in the third block (if a third block exists) and to replicated design points within blocks.

The **IR** strategy is based on the assumption that the assignment of a different set of pseudorandom number streams to each simulation run results in independent sample responses. The **CR** strategy assumes that the assignment of the same set of random number streams to each simulation run results in positively correlated sample responses. In this research, it was assumed that the magnitude of the induced positive correlation ($0 \leq \rho_{+} < 1$) was unknown but estimable. Underlying the **AR** strategy is the assumption that the assignment of common and antithetic streams to the design points in opposite blocks results in positive correlation among sample responses in the same block and negative correlation among sample responses in opposite blocks. The magnitudes of the induced negative and positive correlations ($0 \leq \rho_{-} \leq \rho_{+} < 1$) were assumed to be unknown but estimable. The **AR** strategy is equivalent to the **IR** strategy whenever $\rho_{+} = \rho_{-} = 0$, and the **CR** strategy is equivalent to the **IR** strategy whenever $\rho_{+} = 0$.

Numerous criteria have been developed for evaluating response surface designs, but two mean squared error criteria, MSE of response and MSE of slope, denoted by **J** and **J***, are of particular importance. The MSE criteria take bias error, denoted by **B** and **B***, as well as variance error, denoted by **V** and **V***, into account. Variance error represents the variability in the predicted responses (or slopes) when the fitted model is the correct model, and bias error represents the difference between the true and predicted responses (or slopes) when the fitted model is incorrect. The inclusion of bias error in the performance criteria is particularly important in RSM studies because the fitted polynomial models are only meant to approximate the true relationship between the response and the input variables within restricted regions of the factor space. Therefore, the possibility of an inadequate model, resulting in one or more biased model coefficients, exists in response surface optimization studies.

The research presented in Chapters 3 and 4 assumed that the experimenter desired protection against bias in the predicted responses (or slopes) resulting from the true response function being of order one degree higher than the fitted model. Therefore, the fitted first order polynomial models of Chapter 3 included linear terms (β_0 and $\beta_i x_i$) and the protection models additionally included quadratic ($\beta_{ii} x_i^2$) and two-way interaction terms ($\beta_{ij} x_i x_j$). The fitted second order polynomial models of Chapter 4 included linear, quadratic, and two-way interaction terms, and the protection models additionally included cubic ($\beta_{iii} x_i^3$) and third order interaction terms ($\beta_{ijj} x_i x_j^2$ and $\beta_{ijk} x_i x_j x_k$).

In the first order case, the class of two-level factorial designs were used to evaluate the correlation induction strategies. When the number of input variables was less than five, full factorial designs were used, and when there were five or more input variables, fractional factorial designs were used. The first order designs examined in this research included:

1. Full two-level factorial designs with $k=1, 2, 3, 4$ factors
2. ½-fractional replications of the $k=5, 6$ designs
3. ¼-fractional replication of the $k=7$ design.

In the second order case, four design classes were used to evaluate the correlation induction strategies. When the number of input variables was greater than five, fractional replicates of the central composite and three-level factorial designs were used. The following second order designs were considered:

1. Central composite designs
    - Full factorial replications of the $k=2, 3, 4, 5$ designs
    - ½-fractional replications of the $k=6, 7$ designs
2. Box-Behnken designs
    - designs with $k=4, 5, 7$ factors

3. Three-level factorial designs

    - Full factorial replications of the $k = 3, 4, 5$ designs
    - $^1/_3$-fractional replications of the $k = 6, 7$ designs

4. Small composite designs

    - Hartley designs with $k = 3, 4, 6$ factors
    - Draper designs with $k = 5, 7$ factors.

Designs with $k = 2$ factors do not exist in the Box-Behnken and small composite design classes, and the $k = 2$ three-level factorial design does not block orthogonally (a requirement of the AR correlation induction strategy). Therefore, the only $k = 2$ second order design examined in this research was the central composite design. The $k = 3$ and $k = 6$ Box-Behnken designs were not pursued because the $k = 3$ design does not block orthogonally and the $k = 6$ design does not have a constant [iiii]/[iijj] ratio (required for analytical determination of the optimal design). The small composite designs considered were those developed by Hartley (24) and Draper (14), whichever design had the fewest design points for a particular value of k.

In the first order case, the correlation induction strategies were evaluated under ordinary least squares (OLS) and weighted least squares (WLS) parameter estimation. For the independent streams strategy, the OLS and WLS techniques were equivalent; for the common streams strategy, the OLS and WLS techniques were equivalent when the CR strategy was *pure* (for designs with fewer than two center runs); and for the assignment rule strategy, the OLS and WLS techniques were equivalent for the designs which partitioned into two orthogonal blocks and had fewer than three center runs (the *pure* AR strategy). However, when the CR and AR strategies were *modified* to allow for additional center runs and/or a third orthogonal block, the OLS and WLS techniques were not equivalent. Therefore, both OLS and WLS estimation were used to evaluate the correlation induction strategies in the first order case.

In the second order case, the OLS and WLS techniques yielded identical results for the pure CR and AR strategies but the techniques were not equivalent for the modified CR and AR strategies.

However, for the following reasons, the correlation induction strategies were not evaluated under WLS estimation in the second order case: (1) the first order research findings indicated that correlation induction was more beneficial under WLS than under OLS, and therefore it would be more difficult for the CR and AR strategies to perform better than the IR strategy under OLS estimation; (2) the non-diagonal $(X_1' X_1)^{-1}$ matrices of the second order designs resulted in equations for the bias and variance errors which were too complicated to be minimized analytically under WLS estimation; (3) the optimal second order design plans typically required the use of two or three center runs, resulting in near-pure CR and AR strategies and near-equivalence of OLS and WLS; and (4) WLS estimation is rarely used in practice because the variance-covariance matrix of the response must be known in order to estimate the model parameters. Under OLS estimation, the model parameters can be estimated using only the sample response data and the levels of the input variables, whereas the magnitudes of the induced correlations must be known for WLS estimation of the parameters.

The optimal first order design plans identified were termed **J**- and **J***-optimal designs. These designs utilized the values of the pure second order design moment, [ii], which resulted in minimum values of **J** and **J***. In the second order case, the optimal design plans could not be specified by the values of [ii] alone because **J** and **J*** were also functions of the pure and mixed fourth order design moments, [iiii] and [iijj]. The optimal second order designs utilized the value of θ ($\theta = [iiii]/[iijj]$) which minimized **B** and **B*** and the value of [ii] which minimized **V** and **V*** (given the Min-**B** and Min-**B*** value of θ). For each second order design, the optimal value of θ was achieved through the use of the appropriate scaling factor, g, and the optimal value of [ii] was achieved through the use of the appropriate number of center runs, N_c. Because the number of center runs added to a design must be integer-valued, the actual values of [ii] were only near-optimal, resulting in *near* Min-**V** | Min-**B** and *near* Min-**V*** | Min-**B*** second order designs.

In the first order case, the correlation induction strategies were evaluated by comparing the optimal values of **J** and **J*** under each strategy. The optimal design sizes, as measured by the pure second order design moment, [ii], were also compared for each strategy. The performance of the three strategies were evaluated in the following situations:

- under OLS and WLS estimation,
- in spherical and cuboidal regions of interest,
- for poorly-conditioned ($\phi \cong 0$) to well-conditioned ($\phi \cong k$) response surfaces,
- using two-level factorial designs with $k=1$ through $k=7$ factors,
- using designs with $N_c=0$ to $N_c=5$ center runs,
- for induced correlation magnitudes of $0 \le \rho_- \le \rho_+ \le .90$, and
- for variance error to bias error ratios of $0 < \mathbf{V}/\mathbf{B},\ \mathbf{V}^*/\mathbf{B}^* \le 10$.

In the second order case the correlation induction strategies were evaluated by comparing the optimal values of **V** and **V*** under each strategy. The performance of the three strategies were evaluated in the following situations:

- under OLS estimation,
- in spherical and cuboidal regions of interest,
- for designs with $k=2$ through $k=7$ factors,
- using four different second order design classes, and
- for induced correlation magnitudes of $0 \le \rho_- \le \rho_+ \le .90$.

The relative performance of the four second order design classes, as well as the optimal size of the designs, as measured by the scaling factor, g, were also compared for each strategy. The values of **J** and **J*** were not computed in the second order case because **B** and **B*** were unaffected by correlation induction under OLS estimation and the optimal values of θ were independent of the magnitudes of **B** and **B***.

In summary, this dissertation evaluated the performances of the IR, CR, and AR correlation induction strategies in terms of two MSE criteria. For the case of fitting a first order response surface model, the strategies were evaluated under OLS and WLS estimation using Min-J and Min-J* two-level factorial designs. In the second order case, the strategies were evaluated under OLS estimation using Min-V | Min-B and Min-V* | Min-B* designs in four design classes; central composite designs, Box-Behnken designs, three-level factorial designs, and small composite designs.

5.2 Research Findings

The focus of this dissertation has been to evaluate three strategies for assigning pseudorandom number streams to the stochastic components of simulation models. The evaluations were performed in the context of response surface optimization studies and utilized both first and second order response surface designs.

Chapter 3 investigated the performance of the three assignment strategies using first order response surface designs. Two mean squared error criteria, MSE of the predicted response variable and MSE of the slopes of the response function, served as performance measures. It was assumed that the experimenter was using a two-level factorial design to fit a first order polynomial model to simulated response data and desired protection against second order bias in the fitted model coefficients. The optimal design plans which were developed resulted in minimum values for the sum of the variance and bias components of the MSE criteria. In the case of the MSE of response criteria, the optimal designs were termed J-optimal designs, and these designs resulted in minimum values for $J = B + V$. In the case of the MSE of slope criteria, the optimal designs were termed J^*-optimal designs and resulted in minimum values for $J^* = B^* + V^*$. The relative performance of the IR, CR, and AR strategies in the first order case are summarized in the list which follows.

1. MSE of Response Criteria in the First Order Case:

   - The CR strategy always results in larger values of J than the IR strategy, and therefore the IR strategy is always preferable to the CR strategy in terms of the MSE of response criteria.

   - The AR strategy results in smaller values of J than the IR strategy when $V/B < 3$. Thus, the AR strategy is preferable to the IR strategy when the variance error is less than three times larger than the bias error.

   - The performance of the AR strategy improves relative to the IR strategy (better over a wider range of V-to-B ratios) as the number of center runs increase, the number of factors in the model decrease, and the magnitudes of the induced correlations increase.

- The **J**-optimal design sizes (as measured by [ii]) are similar to the size of the designs which minimize **B** alone, and much smaller than the size of the designs which minimize **V** alone. The difference between the **J**- and **B**-optimal values of [ii] decrease as the magnitudes of the induced correlations increase.

- The performance of the CR and AR strategies are better under WLS estimation than under OLS estimation. Under WLS, the **B**-optimal design sizes tend to increase as the magnitudes of the induced correlations increase, but under OLS estimation, the **B**-optimal design sizes are unaffected by the induction of correlation.

- The relative performance of the three strategies is the same in spherical and cuboidal regions, however, the optimal design sizes are smaller in a spherical region. Under WLS estimation, the optimal design sizes in a spherical region decrease with k, but the optimal design sizes in a cuboidal region increase with k.

- The optimal design sizes for well-conditioned response surfaces (orthogonal input variables) are smaller than the optimal design sizes for poorly-conditioned response surfaces (redundant input variables).

2. MSE of Slope Criteria in the First Order Case:

   - The CR and AR strategies always result in smaller values of $\mathbf{J}^*$ than the IR strategy, and therefore these strategies are always preferable to the IR strategy in terms of the MSE of slope criteria.

   - The performances of the CR and AR strategies relative to the IR strategy improve as the ratio of $\mathbf{V}^*$-to-$\mathbf{B}^*$ increases and as the magnitude of the induced positive correlation increases.

   - The size of the $\mathbf{J}^*$-optimal designs are the same as the size of the designs which minimize $\mathbf{V}^*$ alone because the $\mathbf{B}^*$ component of $\mathbf{J}^*$ is independent of [ii]. Therefore, the optimal design size does not depend on the amount of unfitted quadratic curvature nor the state of conditioning of the response surface.

   - The OLS and WLS estimation techniques are equivalent for the MSE of slope criteria, and therefore the performances of the CR and AR strategies are unaffected by the addition of center runs.

   - The relative performance of the three strategies is the same in a spherical and a cuboidal region of interest. However, the performances of the strategies tend to improve as k increases in a spherical region, but tend to deteriorate as k increases in a cuboidal region.

Chapter 4 investigated the performance of the correlation induction strategies in terms of the MSE of response and MSE of slope criteria using four classes of second order designs; central composite designs (CCDs), Box-Behnken designs (BBDs), three-level factorial designs (FACs), and small composite designs (SCDs). It was assumed that the fitted model was second order but protection against third order bias was desired. The optimal designs which were developed resulted in minimum values of the variance components, given minimum values of the bias components. In the case of the MSE of response criteria, the optimal designs were termed Min-**V** | Min-**B** designs, and for the MSE of slope criteria, the optimal designs were termed Min-**V*** | Min-**B*** designs. The relative performance of the **IR**, **CR**, and **AR** strategies in the second order case are summarized in the list which follows:

1. MSE of Response Criteria in the Second Order Case:

   - The **AR** strategy performs the best of the three strategies and its performance improves as the magnitudes of the induced positive and negative correlations increase. The performance of the **AR** strategy relative to the **IR** strategy deteriorates slightly for large designs (particularly when the number of design points is more than twice that of a saturated design).

   - The **CR** strategy performs better than the **IR** strategy for the SCD design class only. For the CCDs, BBDs, and FACs, the **IR** strategy is preferable to the **CR** strategy, and the performance of the **CR** strategy deteriorates as the magnitude of the induced correlation increases.

   - The levels of the input variables in the Min-**V** | Min-**B** designs are smaller than the $g = 1$ levels which are typically used in practice. Therefore, the optimal designs require placing the design points closer to the center of the experimental region. As the number of factors in the model increase, the optimal placement of the design points becomes even closer to the center of the region.

   - The optimal levels of the input variables are smaller in a spherical region of interest ($g = .34 - .62$) than in a cuboidal region ($g = .60 - .98$).

   - The optimal number of center runs for the CCDs, BBDs, and SCDs is typically one or two in spherical region of interest. For a cuboidal region, one additional center run is generally required for a Min-**V** | Min-**B** design. The FACs do not require the use of additional center runs in either region.

   - The CCDs and BBDs perform the best of the four design classes, with the $k = 5, 7$ BBDs performing better than the corresponding CCDs. The FACs perform the poorest of the four design classes under the **CR** and **AR** strategies, and the SCDs perform the poorest under the **IR** strategy.

2. MSE of Slope Criteria in the Second Order Case:

- The **AR** strategy performs the best of the three strategies under the MSE of slope criteria and its performance improves as the magnitudes of the induced positive and negative correlations increase.

- The **CR** strategy performs better than the **IR** strategy but worse than the **AR** strategy, and its performance improves as the magnitude of the induced correlation increases.

- The optimal levels of the input variables for the Min-**V***| Min-**B*** designs are slightly larger than those for the Min-**V** | Min-**B** designs, but tend to be smaller than the $g = 1$ levels which are typically used in practice.

- The optimal levels of the input variables are smaller in a spherical region of interest ($g = .38 - .71$) than in a cuboidal region ($g = .65 - 1.1$).

- The optimal number of center runs for the CCDs, BBDs, and SCDs is about three or four (in both a spherical and a cuboidal region of interest). As the number of factors in the model and/or the number of design points increase, the optimal number of center runs generally increases.

- The CCDs and BBDs tend to perform the best of the four design classes under the MSE of slope criteria, and the SCDs tend to perform the poorest.

Results for the first and second order cases indicate some similarities in the relative performances of the correlation induction strategies. When taken together, the first and second order results can be summarized as follows:

1. The assignment rule blocking strategy tends to be the preferred correlation induction strategy for prediction of the response variable and for estimation of the slopes of the response function.

2. The independent streams strategy is a poor choice when estimation of the slopes of the response function is important, and the common streams strategy is a poor choice when prediction of the response is important.

3. The performance of the assignment rule blocking strategy improves as the magnitudes of the induced positive and negative correlations increase. The performance of the common streams strategy under the MSE of slope criteria improves as the magnitude of the induced positive correlation increases.

4. For protection against bias, the levels of the input variables should be reduced about 30% from the ± 1's extremes which are typically used in practice (to coded levels of $g \cong .7$). The amount of scaling required is slightly less in a

cuboidal region and slightly less when the design criteria is MSE of slope. More scaling is required as the number of factors increases.

5. In the second order case, designs which are near saturation (minimum number of design points) tend to perform poorly when estimation of the slopes of the response function is important. Large designs (about twice the size of a saturated design) tend to perform poorly in terms of prediction when correlation is induced among the responses. The ideal number of design points appears to be about 30 - 60% more than a saturated design.

In summary, the results of this research indicate that the assignment rule blocking strategy generally performs the best of the three strategies evaluated. The independent streams strategy would be a poor choice during the steepest ascent and ridge analysis stages of RSM when estimation of the slopes of the response function is important. Common streams, the strategy which is typically used in practice, would be a poor choice during the canonical and ridge analysis stages of the optimization process when prediction of the response is important. The results of this research also indicate that the ± 1 levels of the design variables which are typically used in practice should be reduced about 30% to afford protection against bias in the model coefficients. The addition of about three center runs also appears to be useful for protecting against bias due to model inadequacy.

BIBLIOGRAPHY

1) Banks, J., & Carson, J. S. (1984). *Discrete-Event System Simulation.* Prentice-Hall, New Jersey.

2) Biles, W. E., & Swain, J. J. (1980). *Optimization and Industrial Experimentation.* John Wiley & Sons, New York.

3) Box, G. E. P. (1952). Multifactor Designs of the First Order. *Biometrika,* Vol. 39, 49-57.

4) Box, G. E. P., & Behnken, D. W. (1960). Some New Three Level Designs for the Study of Quantitative Variables. *Technometrics,* Vol. 2, No. 4, 455-475.

5) Box, G. E. P., & Draper, N. R. (1959). A Basis for the Selection of a Response Surface Design. *Journal of the American Statistical Association,* Vol. 54, 622-654.

6) Box, G. E. P., & Draper, N. R. (1963). The Choice of a Second Order Rotatable Design. *Biometrika,* Vol. 50, No. 3, 335-352.

7) Box, G. E. P., & Draper, N. R. (1987). *Empirical Model-Building and Response Surfaces.* John Wiley & Sons, New York.

8) Box, G. E. P., & Hunter, J. S. (1957). Multifactor Experimental Designs for Exploring Response Surfaces. *Annals of Mathematical Statistics,* Vol. 28, 195-241.

9) Box, G. E. P., & Hunter, J. S. (1961). The 2^{k-p} Fractional Factorial Designs, I and II. *Technometrics,* Vol. 3, 311-351, 449-458.

10) Box, G. E. P., & Wilson, K. B. (1951). On the Experimental Attainment of Optimum Conditions. *Journal of the Royal Statistical Society, Series B,* Vol. 13, 1-45.

11) Brightman, H. J. (1978). Optimization Through Experimentation: Applying Response Surface Methodology. *Decision Sciences,* Vol. 9, No. 3, 481-495.

12) Cooley, B. J., & Houck, E. C. (1982). A Variance-reduction Strategy for RSM Simulation Studies. *Decision Sciences,* Vol. 13, No. 2, 303-321.

13) Draper, N. R. (1963). "Ridge Analysis" of Response Surfaces. *Technometrics,* Vol. 5, No. 4, 469-479.

14) Draper, N. R. (1985). Small Composite Designs. *Technometrics*, Vol. 27, No. 2, 173-180.

15) Draper, N. R., & Guttman, I. (1986). Response Surface Designs in Flexible Regions. *Journal of the American Statistical Association*, Vol. 81, No. 396, 1089-1094.

16) Draper, N. R., & Lawrence, W. E. (1965). Designs Which Minimize Model Inadequacies: Cuboidal Regions of Interest. *Biometrika*, Vol. 52, No. 2, 111-118.

17) Draper, N. R., & Smith H. (1981). *Applied Regression Analysis*. John Wiley & Sons, New York.

18) Emshoff, J. R., & Sisson, R. L. (1970). *Design and Use of Computer Simulation Models*. Macmillan Company, New York.

19) Fisher, R. A. (1921). Studies in Crop Variation. I. *Journal of Agricultural Science*, Vol. 11, 107-135.

20) Fishman, G. S. (1973). *Concepts and Methods in Discrete Event Simulation*. John Wiley & Sons, New York.

21) Fishman, G. S. (1974). Correlated Simulation Experiments. *Simulation*, Vol. 23, No. 6, 177-180.

22) Hammersley, J. M., & Handscomb, D. C. (1964). *Monte Carlo Methods*. Methuen & Co., London.

23) Handscomb, D. C. (1969). Monte Carlo Techniques: Theoretical, in *The Design of Computer Simulation Experiments*, Naylor, T. H. (editor), Duke University Press, Durham, NC, 252-262.

24) Hartley, H. O. (1959). Smallest Composite Designs For Quadratic Response Surfaces. *Biometrics*, Vol. 15, No. 4, 611-624.

25) Heidelberger, P., & Iglehart, D. L. (1979). Comparing Stochastic Systems Using Regenerative Simulation with Common Random Numbers. *Advanced Applied Probability*, Vol. 11, 804-819.

26) Hill, W. J., & Hunter, W. G. (1966). A Review of Response Surface Methodology: A Literature Survey. *Technometrics*, Vol. 8, No. 4, 571-590.

27) Hoerl, A. E. (1959). Optimal Solution of Many Variable Equations. *Chemical Engineering Progress*, Vol. 55, No. 11, 69-78.

28) Hussey, J. R., Myers, R. H., & Houck, E. C. (1987). Correlated Simulation Experiments in First-Order Response Surface Design. *Operations Research*, Vol. 35, No. 5, 744-758.

29) Hussey, J. R., Myers, R. H., & Houck, E. C. (1987). Pseudorandom Number Assignments in Quadratic Response Surface Designs. *IIE Transactions*, Vol. 19, No. 4, 385-403.

30) Ignall, E. (1972). On Experimental Designs for Computer Simulation Models. *Management Science*. Vol. 18, No. 3, 384-388.

31) Khuri, A. I., & Cornell, J. A. (1987). *Response Surfaces: Designs and Analyses*. Marcel Dekker, New York.

32) Kiefer, J. C. (1984). The Publications and Writings of Jack Kiefer. *Annals of Statistics*, Vol. 12, 424-430.

33) Kleijnen, J. P. C. (1974). *Statistical Techniques in Simulation, Parts I and II*. Marcel Dekker, New York.

34) Kleijnen, J. P. C. (1988) Analyzing Simulation Experiments with Common Random Numbers. *Management Science,* Vol. 34, No. 1, 65-74.

35) Koshal, R. S. (1933). Application of the Method of Maximum Likelihood to the Improvement of Curves Fitted by the Method of Moments. *Journal of the Royal Statistical Society,* Vol. A96, 303-313.

36) Krasnow, H. S. (1969). Simulation Languages, in *The Design of Computer Simulation Experiments,* Naylor, T. H. (editor), Duke University Press, Durham, NC, 320-346.

37) Law, M. L., & Kelton, W. D. (1982). *Simulation Modeling and Analysis.* McGraw Hill, New York.

38) Mead, R., & Pike, D. J. (1975). A Review of Response Surface Methodology from a Biometric Viewpoint. *Biometrics,* Vol. 31, No. 12, 803-851.

39) Mihram, G. A. (1972). *Simulation: Statistical Foundations and Methodology.* Academic Press, New York.

40) Montgomery, D. C. (1976). *Design and Analysis of Experiments.* John Wiley & Sons, New York.

41) Montgomery, D. C., & Bettencourt, V. M. (1977). Multiple Response Surface Methods in Computer Simulation. *Simulation,* Vol. 29, No. 4, 113-121.

42) Montgomery, D. C. & Evans, D. M. (1975). Second-Order Response Surface Designs in Computer Simulation. *Simulation,* Vol. 25, No. 6, 169-178.

43) Morton, R. H. (1983). Response Surface Methodology. *Mathematical Scientist,* Vol. 8, No. 1, 31-52.

44) Moy, W. A. (1969). Monte Carlo Techniques: Practical, in *The Design of Computer Simulation Experiments,* Naylor, T. H. (editor), Duke University Press, Durham, NC, 263-288.

45) Myers, R. H. (1988). Response Surface Methodology: A Literature Review. To appear in *Technometrics,* Vol. 30.

46) Myers, R. H. (1976). *Response Surface Methodology.* Allyn and Bacon, Boston.

47) Myers, R. H., & Khuri, A. I. (1979). A New Procedure for Steepest Ascent. *Communications in Statistics,* Vol. 8, 1359-1376.

48) Myers, R. H., & Lahoda, S. J. (1975). A Generalization of the Response Surface Mean Square Error Criterion with a Specific Application to the Slope. *Technometrics,* Vol. 17, No. 4, 481-486.

49) Naylor, T. H., Balintfy, J. L., Burdick, D. S., & Chu, K. (1966). *Computer Simulation Techniques.* John Wiley & Sons, New York.

50) Naylor, T. H. (1969). *The Design of Computer Simulation Experiments.* Duke University Press, Durham, NC.

51) Pazman, A. (1986). *Foundations of Optimum Experimental Design.* D. Reidel Publishing Company, Boston, MA.

52) Plackett, R. L., & Burman, J. P. (1946). The Design of Optimum Multifactorial Experiments. *Biometrika,* Vol. 33, Part 4, 305-325 and 328-332.

53) Rees, L. P., Clayton, E. R., & Taylor, B. W. III (1985). Solving Multiple Response Simulation Models Using Modified Response Surface Methodology Within a Lexicographic Goal Programming Framework. *IIE Transactions,* Vol. 17, No. 1, 47-57.

54) Rubinstein, R. Y., & Samorodnitsky, G. (1976). Variance Reduction by the Use of Common and Antithetic Random Variables. *Journal of Statistical Computation and Simulation,* Vol. 22, 161-180.

55) Schruben, L. W. (1979). Designing Correlation Induction Strategies for Simulation Experiments, in *Current Issues in Computer Simulation,* Adam, N. R., & Dogramaci, A. (editors), Academic Press, New York, 235-257.

56) Schruben, L. W., & Margolin, B. H. (1978). Pseudorandom Number Assignment in Statistically Designed Simulation and Distribution Sampling Experiments. *Journal of the American Statistical Association,* Vol. 73, No. 363, 504-520.

57) Silvey, S. D. (1980). *Optimal Designs.* Chapman and Hall, London.

58) Spendley, W., Hext, G. R., & Himsworth, F. R. (1962). Sequential Application of Simplex Designs in Optimization and EVOP. *Technometrics,* Vol. 4, 441-461.

59) Steinberg, D. M., & Hunter, W. G. (1984). Experimental Design: Review and Comment. *Technometrics,* Vol. 26, No. 2, 71-130.

60) Sullivan, R. S., Hayya, J. C., & Schaul, R. (1982). Efficiency of the Antithetic Variate Method for Simulating Stochastic Networks. *Management Science,* Vol. 28, No. 5, 563-572.

61) Tocher, K. D. (1963). *The Art of Simulation.* The English Universities Press, London.

62) Westlake, W. J. (1959). Composite Designs Based on Irregular Fractions of Factorials. *Biometrics,* Vol. 21, No. 2, 324-336.

63) Wright, R. D., & Ramsay, T. E. (1979). On the Effectiveness of Common Random Numbers. *Management Science,* Vol. 25, No. 7, 649-656.

64) Yates, F. (1935). Complex Experiments. *Journal of the Royal Statistical Society,* Vol. B2, 181-223.

APPENDICES

Appendix A

Variance-Covariance Matrix of the Response

Consider an experimental design consisting of N simulation runs, yielding a set of N sample responses, denoted as

$$\underline{y} = [\,Y_1, Y_2, \ldots\ldots, Y_N\,]' .$$

If one assumes that the sample responses have a common variance σ^2, then the variance-covariance matrix of $\underline{y}$ becomes

$$\mathrm{Var}(\underline{y}) = \sigma^2 V$$

where V is an N×N matrix of the correlation between pairs of sample responses, defined as

$$V = \begin{bmatrix} 1 & \rho_{12} & & \cdots & & \rho_{1N} \\ \rho_{21} & 1 & 1 & & & \vdots \\ \vdots & & & . & & \vdots \\ \vdots & & & & 1 & \rho_{ij} \\ \rho_{N1} & & \cdots & & \rho_{ij} & 1 \end{bmatrix}$$

where ρ_{ij} is the correlation between sample responses Y_i and Y_j ($i, j = 1, 2, \ldots, N$). Under the assumptions discussed in section 2.1.4.1 of this paper, the *signs* of the ρ_{ij} are different when independent (IRN), common (CRN), and antithetic (ARN) random number streams are used to generate the pair of responses; that is

$$
\begin{aligned}
\rho_{ij} &= 0 && \text{if IRN is used to generate } Y_i \text{ and } Y_j, \\
&= \rho_+ && \text{if CRN is used to generate } Y_i \text{ and } Y_j, \\
&= -\rho_- && \text{if ARN is used to generate } Y_i \text{ and } Y_j.
\end{aligned}
$$

Utilizing the above signs for the ρ_{ij}, the correlation matrix, V, can be defined for each correlation induction strategy examined in this research. The first strategy, **IR**, involves the use of independent pseudorandom number streams on each simulation run, thereby generating independent sample responses. The correlation matrix for the IR strategy, denoted by V_{IR}, can be written as

$$
\begin{aligned}
V_{IR} &= \begin{bmatrix} 1 & 0 & & \cdots & 0 \\ 0 & 1 & _1 & & \vdots \\ \vdots & & & \cdot & \vdots \\ \vdots & & & 1 & 0 \\ 0 & & \cdots & 0 & 1 \end{bmatrix} \\
&= I_N
\end{aligned}
$$

where I_N is the N×N identity matrix.

The second correlation induction strategy examined in this research, **CR**, involves the use of common streams for non-replicated design points and the use of independent streams for replicated design points. The correlation between sample response pairs generated with common streams is equal to ρ_+ and all other sample response pairs are uncorrelated. It is convenient to partition the vector of sample responses into two parts in order to define the correlation matrix. Suppose the number of non-replicated design points is denoted by N_1 and the number of replicated design points is denoted by N_2, then $N_1 + N_2 = N$ and the partitioned sample response vector can be written as

$$\underline{y} = \left[Y_1, \ldots\ldots, Y_{N_1} \mid Y_{N_1+1}, \ldots\ldots, Y_N \right]'.$$

The partitioned form of the correlation matrix for the **CR** strategy, denoted by V_{CR}, becomes

$$V_{CR} = \left[\begin{array}{ccccc|cccc}
1 & \rho_+ & & \cdots & \rho_+ & 0 & 0 & \cdots & 0 \\
\rho_+ & 1 & & & \vdots & 0 & \cdot & & \vdots \\
\vdots & & 1 & & \vdots & \vdots & & & \vdots \\
\vdots & & & \cdot & \rho_+ & \vdots & & \cdot & \vdots \\
\rho_+ & & \cdots & \rho_+ & 1 & 0 & \cdots & 0 & 0 \\
\hline
0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
0 & \cdot & & & \vdots & 0 & 1 & & \vdots \\
\vdots & & & \cdot & \vdots & \vdots & & \cdot & 0 \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1
\end{array}\right]$$

$$= \left[\begin{array}{c|c}
I_{N_1}(1-\rho_+) + \underline{1}_{N_1}\underline{1}'_{N_1}\rho_+ & \underline{0}_{N_1}\underline{0}'_{N_2} \\
\hline
\underline{0}_{N_2}\underline{0}'_{N_1} & I_{N_2}
\end{array}\right]$$

where I_{N_1} is an $(N_1 \times N_1)$ identity matrix ,
I_{N_2} is an $(N_2 \times N_2)$ identity matrix ,
$\underline{1}_{N_1}$ is an $(N_1 \times 1)$ column vector of 1's ,
$\underline{1}_{N_2}$ is an $(N_2 \times 1)$ column vector of 1's ,
$\underline{0}_{N_1}$ is an $(N_1 \times 1)$ column vector of 0's , and
$\underline{0}_{N_2}$ is an $(N_2 \times 1)$ column vector of 0's .

The correlation matrix, V_{CR}, can also be written in equation form as

$$V_{CR} = I_N + \rho_+ \underline{u}\underline{u}' - \rho_+ U$$

where U is an $N \times N$ diagonal matrix of the vector $\underline{u}$, and $\underline{u}$ is an $N \times 1$ vector whose i^{th} element is

$$u_i = 1 \quad \text{if a common } \textit{or} \text{ antithetic random number stream is used for design point } i$$
$$= 0 \quad \text{if an independent random number stream is used for design point } i\,.$$

For the special case in which there are no replicated design points; that is, the *pure* CR strategy, the equation for V_{CR} can be written as

$$V_{CR} = (1-\rho_+)\, I_N + \rho_+\, \underline{1}_N \underline{1}'_N\,.$$

The third correlation induction strategy examined this research, AR, involves the use of common streams for non-replicated design points in one orthogonal block, the use of antithetic streams for non-replicated design points in the opposite block, and the use independent streams for replications of design points within blocks and for design points in the third block (if one exists). The correlation between sample response pairs within blocks is equal to ρ_+, the correlation between responses in opposite blocks is equal to $-\rho_-$, and all other sample response pairs are uncorrelated. It is convenient to partition the vector of sample responses into three parts in order to define the correlation matrix. Suppose the number of design points assigned common streams is denoted by N_{1a}, the number of design points assigned antithetic streams is denoted by N_{1b}, and the number of design points assigned independent streams is denoted by N_2, then $N_{1a} + N_{1b} + N_2 = N$ and the partitioned sample response vector can be written as

$$\underline{y} = \left[\, Y_1, \ldots, Y_{N_{1a}} \;\middle|\; Y_{N_{1a}+1}, \ldots, Y_{N_{1a}+N_{1b}} \;\middle|\; Y_{N_{1a}+N_{1b}+1}, \ldots, Y_N \,\right]',$$

and the partitioned form of the correlation matrix for the AR strategy, denoted by V_{AR}, becomes

$$
V_{AR} = \left[\begin{array}{ccccc|cccc|cccc}
1 & \rho_+ & & \cdots & \rho_+ & -\rho_- & -\rho_- & \cdots & -\rho_- & 0 & 0 & \cdots & 0 \\
\rho_+ & 1 & & & \vdots & -\rho_- & \cdot & & \vdots & 0 & \cdot & & \vdots \\
\vdots & & 1 & & \vdots & \vdots & & -\rho_- & \vdots & \vdots & & 0 & \vdots \\
\vdots & & & \cdot & \rho_+ & \vdots & & \cdot & \vdots & \vdots & & \cdot & \vdots \\
\rho_+ & & \cdots & \rho_+ & 1 & -\rho_- & \cdots & -\rho_- & -\rho_- & 0 & \cdots & 0 & 0 \\
\hline
-\rho_- & -\rho_- & -\rho_- & \cdots & -\rho_- & 1 & \rho_+ & \cdots & \rho_+ & 0 & 0 & \cdots & 0 \\
-\rho_- & \cdot & & & \vdots & \rho_+ & 1 & & \vdots & 0 & 0 & & \vdots \\
\vdots & & & \cdot & \vdots & \vdots & & \cdot & \rho_+ & \vdots & & \cdot & 0 \\
-\rho_- & \cdots & -\rho_- & -\rho_- & -\rho_- & \rho_+ & \cdots & \rho_+ & 1 & 0 & \cdots & 0 & 0 \\
\hline
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
0 & \cdot & & & \vdots & 0 & \cdot & & \vdots & 0 & 1 & & \vdots \\
\vdots & & & \cdot & \vdots & \vdots & & \cdot & \vdots & \vdots & & \cdot & 0 \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1
\end{array}\right]
$$

$$
= \left[\begin{array}{c|c|c}
\mathrm{I}_{N_{1a}}(1-\rho_+) + \underline{1}_{N_{1a}}\underline{1}'_{N_{1a}}\rho_+ & -\underline{1}_{N_{1a}}\underline{1}'_{N_{1b}}\rho_- & \underline{0}_{N_{1a}}\underline{0}'_{N_2} \\
\hline
-\underline{1}_{N_{1b}}\underline{1}'_{N_{1a}}\rho_- & \mathrm{I}_{N_{1b}}(1-\rho_+) + \underline{1}_{N_{1b}}\underline{1}'_{N_{1b}}\rho_+ & \underline{0}_{N_{1b}}\underline{0}'_{N_2} \\
\hline
\underline{0}_{N_2}\underline{0}'_{N_{1a}} & \underline{0}_{N_2}\underline{0}'_{N_{1b}} & \mathrm{I}_{N_2}
\end{array}\right].
$$

The correlation matrix, V_{AR}, can now be written in equation form as

$$
V_{AR} = \mathrm{I}_N + \tfrac{1}{2}(\rho_+ - \rho_-)\,\underline{u}\,\underline{u}' + \tfrac{1}{2}(\rho_+ + \rho_-)\,\underline{v}\,\underline{v}' - \rho_+\,\mathrm{U}
$$

where $\underline{v}$ is an $N\times 1$ vector whose i^{th} element is

$$
\begin{aligned}
v_i &= 1 && \text{if a common random number stream is used for design point } i \\
&= 0 && \text{if an independent random number stream is used for design point } i \\
&= -1 && \text{if an antithetic random number stream is used for design point } i\,.
\end{aligned}
$$

For the special case in which there are no replicated design points within blocks and the design partitions into two blocks; that is, the *pure* AR strategy, the equation for V_{AR} can be written as

$$
V_{AR} = (1-\rho_+)\,\mathrm{I}_N + \tfrac{1}{2}(\rho_+ - \rho_-)\,\underline{1}_N\underline{1}'_N + \tfrac{1}{2}(\rho_+ + \rho_-)\,\underline{v}\,\underline{v}'\,.
$$

Appendix B

First Order Response Model

For the case in which an experimenter fits a first order polynomial model to the experimental data, the equation for the *fitted* response model is

$$\hat{y} = b_0 + \sum_{i=1}^{k} b_i x_i .$$

If the experimenter is uncertain as to whether or not the fitted first order model can adequately describe the response surface, then an appropriate experimental plan would *protect* against biases if, in fact, a second order polynomial model was needed to describe the response surface curvature. The equation for the second order *protection* model is

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i<j}\sum \beta_{ij} x_i x_j + \varepsilon .$$

When model uncertainty exists, it is convenient to partition the polynomial model into two parts; a fitted and an unfitted part. The matrix form of the partitioned model is

$$\underline{y} = X_1 \underline{\beta}_1 + X_2 \underline{\beta}_2 + \underline{\varepsilon} .$$

For the case fitting a first order polynomial model ($d_1 = 1$), while protecting against a true second order polynomial model ($d_2 = 2$), the $\underline{\beta}_1$ and $\underline{\beta}_2$ vectors of the partitioned model are

$$\underline{\beta}_1 = [\beta_0, \beta_1, \ldots, \beta_k]'$$

$$\underline{\beta}_2 = [\beta_{11}, \beta_{22}, \ldots, \beta_{kk}, \quad \beta_{12}, \beta_{13}, \ldots, \beta_{k-1,k}]'$$

where $\underline{\beta}_1$ has dimensions $(p_1 \times 1)$, with $p_1 = k+1$ coefficients, and $\underline{\beta}_2$ has dimensions $(p_2 \times 1)$, with $p_2 = k + \binom{k}{2} = \frac{1}{2}k(k+1)$ coefficients. The corresponding X_1 and X_2 matrices are

$$X_1 = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{1N} & x_{2N} & \cdots & x_{kN} \end{bmatrix} \quad (N \times p_1)$$

$$X_2 = \begin{bmatrix} x_{11}^2 & \cdots & x_{k1}^2 & x_{11}\,x_{21} & \cdots & x_{k-1,1}\,x_{k1} \\ x_{12}^2 & \cdots & x_{k2}^2 & x_{12}\,x_{22} & \cdots & x_{k-1,2}\,x_{k2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{1N}^2 & \cdots & x_{kN}^2 & x_{1N}\,x_{2N} & \cdots & x_{k-1,N}\,x_{kN} \end{bmatrix} \quad (N \times p_2)$$

where k is the number of design variables and N is the number of design points.

The moment matrices of a design are useful in describing the average variance and bias properties of specific design classes. The elements of the design moment matrices are defined as

$$\text{Design Moment} = [\,1^{\delta_1}\ 2^{\delta_2} \ldots\ldots k^{\delta_k}\,] = \frac{1}{N} \sum_{u=1}^{N} x_{1u}^{\delta_1}\, x_{2u}^{\delta_2} \ldots\ldots x_{ku}^{\delta_k}$$

where $\sum_{i=1}^{k} \delta_i = \delta$ is the order of the design moment. The design moment matrices, which contain the design moments of both the fitted and unfitted model terms, are defined as

$$M_{11} = N^{-1} X'_1 X_1$$

$$M_{21} = N^{-1} X'_2 X_1$$

$$M_{22} = N^{-1} X'_2 X_2 .$$

For the fit-protection situation of $d_1 = 1$ and $d_2 = 2$, the M_{11} design moment matrix contains moments through order $2d_1 = 2$, the M_{21} matrix contains moments through order $d_1 + d_2 = 3$, and the M_{22} matrix contains moments through order $2d_2 = 4$. Utilizing the design moment notation, the three moment matrices can be written as

$$M_{11} = \left[\begin{array}{c|ccccc} 1 & [1] & [2] & \cdots & [k] \\ \hline [1] & [11] & [12] & \cdots & [1k] \\ \cdot & \cdot & [22] & \cdot & [2k] \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ [k] & [1k] & \cdots & \cdots & [kk] \end{array}\right] \quad (p_1 \times p_1)$$

$$M_{21} = \left[\begin{array}{c|ccccc} [11] & [111] & [112] & \cdots & [11k] \\ \cdot & \cdot & [222] & \cdot & [22k] \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ [kk] & [1kk] & \cdots & \cdots & [kkk] \\ \hline [12] & [112] & [122] & \cdots & [12k] \\ \cdot & \cdot & [123] & \cdot & [13k] \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ [k-1,k] & [1,k-1,k] & \cdots & \cdots & [k-1,k,k] \end{array}\right] \quad (p_2 \times p_1)$$

$$
M_{22} = \left[\begin{array}{cccc|cccc}
[1111] & [1122] & \cdots & [11kk] & [1112] & [1113] & \cdots & [11,k\text{-}1,k] \\
. & [2222] & . & [22kk] & . & [1223] & . & [22,k\text{-}1,k] \\
. & . & . & . & . & . & . & . \\
[11kk] & \cdots & \cdots & [kkkk] & [12kk] & \cdots & \cdots & [k\text{-}1,kkk] \\
\hline
[1112] & [1222] & \cdots & [12kk] & [1122] & [1123] & \cdots & [12,k\text{-}1,k] \\
. & [1223] & . & [13kk] & . & [1133] & . & [13,k\text{-}1,k] \\
. & . & . & . & . & . & . & . \\
[11,k\text{-}1,k] & \cdots & \cdots & [k\text{-}1,kkk] & [12,k\text{-}1,k] & \cdots & \cdots & [k\text{-}1,k\text{-}1,kk]
\end{array}\right] (p_2 \times p_2)
$$

The region moment matrices of the design are useful in describing the *average* variance and bias properties of a design within a specified region of interest. The elements of the region moment matrices are defined as

$$
\text{Region Moment} = w_{1^{\delta_1}2^{\delta_2}\ldots k^{\delta_k}} = \Omega_r \int_R x_1^{\delta_1} x_2^{\delta_2} \ldots\ldots x_k^{\delta_k} \, d\mathbf{x}
$$

where $\sum_{i=1}^{k} \delta_i = \delta$ is the order of the region moment and Ω_r^{-1} is the volume of the region of interest. The region moment matrices, which contain the region moments of both the fitted and unfitted model terms, are defined as

$$
\mu_{11} = \Omega_r \int_R \mathbf{x}'_1 \mathbf{x}_1 \, d\mathbf{x}
$$

$$
\mu_{21} = \Omega_r \int_R \mathbf{x}'_2 \mathbf{x}_1 \, d\mathbf{x}
$$

$$
\mu_{22} = \Omega_r \int_R \mathbf{x}'_2 \mathbf{x}_2 \, d\mathbf{x}
$$

where $\mathbf{x}_1$ is a $p_1 \times 1$ vector of the fitted model terms in X_1, and $\mathbf{x}_2$ is a $p_2 \times 1$ vector of the unfitted model terms in X_2.

For the case of fitting a first order model with protection against a true second order model, the region moment matrices, which are similar in form to the design moment matrices, become

$$\mu_{11} = \left[\begin{array}{c|cccc} 1 & w_1 & w_2 & \cdots & w_k \\ \hline w_1 & w_{11} & w_{12} & \cdots & w_{1k} \\ \cdot & \cdot & w_{22} & \cdot & w_{2k} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ w_k & w_{1k} & \cdots & \cdots & w_{kk} \end{array}\right] (p_1 \times p_1)$$

$$\mu_{21} = \left[\begin{array}{c|cccc} w_{11} & w_{111} & w_{112} & \cdots & w_{11k} \\ \cdot & \cdot & w_{222} & \cdot & w_{22k} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ w_{kk} & w_{1kk} & \cdots & \cdots & w_{kkk} \\ \hline w_{12} & w_{112} & w_{122} & \cdots & w_{12k} \\ \cdot & \cdot & w_{123} & \cdot & w_{13k} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ w_{k-1,k} & w_{1,k-1,k} & \cdots & \cdots & w_{k-1,kk} \end{array}\right] (p_2 \times p_1)$$

$$\mu_{22} = \left[\begin{array}{cccc|cccc} w_{1111} & w_{1122} & \cdots & w_{11kk} & w_{1112} & w_{1113} & \cdots & w_{11,k-1,k} \\ \cdot & w_{2222} & \cdot & w_{22kk} & \cdot & w_{1223} & \cdot & w_{22,k-1,k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ w_{11kk} & \cdots & \cdots & w_{kkkk} & w_{12,k-1,k} & \cdots & \cdots & w_{k-1,kkk} \\ \hline w_{1112} & w_{1222} & \cdots & w_{12kk} & w_{1122} & w_{1123} & \cdots & w_{12,k-1,k} \\ \cdot & w_{1223} & \cdot & w_{13kk} & \cdot & w_{1133} & \cdot & w_{13,k-1,k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ w_{11,k-1,k} & \cdots & \cdots & w_{k-1,kkk} & w_{12,k-1,k} & \cdots & \cdots & w_{k-1,k-1,kk} \end{array}\right] (p_2 \times p_2).$$

Appendix C

First Order Model For Estimation of Slopes

For the case in which an experimenter fits a first order polynomial model to the experimental data, with the purpose of predicting the rate of change of the response variable with respect to the input variables, the equation for the i^{th} partial derivative of the fitted response function is

$$\hat{\gamma}_{(x_i)} = \frac{\partial \hat{y}}{\partial x_i} = b_i \ .$$

If the experimenter suspects that a second order polynomial model may be needed to describe the true response surface and wants to protect against biases due to unfitted second order terms, then the equation for the i^{th} partial derivative of the protection model is

$$\gamma_{(x_i)} = \frac{\partial y}{\partial x_i} = \beta_i + 2\beta_{ii} x_i + \sum_{j \neq i} \beta_{ij} x_j \ .$$

The matrix form of the partial derivative of the response function can be partitioned into two parts; a fitted part and an unfitted part, written as

$$\underline{y}_{(x)} = \Lambda'_{1\,(x)}\,\underline{\beta}_1 + \Lambda'_{2\,(x)}\,\underline{\beta}_2 \,.$$

For the fit-protection situation of $d_1 = 1$ and $d_2 = 2$, the $\underline{\beta}$ vectors of the partitioned model are

$$\underline{\beta}_1 = [\beta_0, \beta_1, \ldots, \beta_k]'$$

$$\underline{\beta}_2 = [\beta_{11}, \beta_{22}, \ldots, \beta_{kk}, \quad \beta_{12}, \beta_{13}, \ldots, \beta_{k-1,k}]'$$

where $\underline{\beta}_1$ has dimensions $(p_1 \times 1)$, with $p_1 = k + 1$ coefficients, and $\underline{\beta}_2$ has dimensions $(p_2 \times 1)$, with $p_2 = k + \binom{k}{2} = \frac{1}{2}k(k+1)$ coefficients. The corresponding $\Lambda'_{1(x)}$ and $\Lambda'_{2(x)}$ matrices can be written as

$$\Lambda'_{1\,(x)} = \left[\begin{array}{c|ccccc}
0 & 1 & 0 & . & . & 0 \\
0 & 0 & 1 & . & . & 0 \\
0 & 0 & 0 & . & . & 0 \\
. & . & . & . & . & . \\
. & . & . & . & . & . \\
0 & 0 & 0 & . & . & 1
\end{array}\right]$$

$$\Lambda'_{2\,(x)} = \left[\begin{array}{ccccc|cccccc|cccccccc}
2x_1 & 0 & . & . & 0 & x_2 & x_3 & x_4 & . & . & x_k & 0 & 0 & 0 & . & . & 0 & . & 0 \\
0 & 2x_2 & . & . & 0 & x_1 & 0 & 0 & . & . & 0 & x_3 & x_4 & x_5 & . & . & x_k & . & 0 \\
0 & 0 & . & . & 0 & 0 & x_1 & 0 & . & . & 0 & x_2 & 0 & 0 & . & . & 0 & . & 0 \\
0 & 0 & . & . & 0 & 0 & 0 & x_1 & . & . & 0 & 0 & x_2 & 0 & . & . & 0 & . & 0 \\
. & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
0 & 0 & . & . & 2x_k & 0 & 0 & 0 & . & . & x_1 & 0 & 0 & 0 & . & . & x_2 & . & x_{k-1}
\end{array}\right]$$

where $\Lambda'_{1(x)}$ is a $k \times p_1$ matrix and $\Lambda'_{2(x)}$ is a $k \times p_2$ matrix.

When integrated over the region of interest, the matrices of the partial derivatives of the response function are useful in describing the average variance and bias properties of designs used in the estimation of the slope coefficients. These region matrices for the partial derivatives of the response function are defined as

$$\mu_{11}^{*} = \Omega_r \int_R \Lambda_{1\,(\underline{x})}\, \Lambda'_{1\,(\underline{x})}\, d\underline{x}$$

$$\mu_{21}^{*} = \Omega_r \int_R \Lambda_{2\,(\underline{x})}\, \Lambda'_{1\,(\underline{x})}\, d\underline{x}$$

$$\mu_{22}^{*} = \Omega_r \int_R \Lambda_{2\,(\underline{x})}\, \Lambda'_{2\,(\underline{x})}\, d\underline{x}\ .$$

For the case of fitting a first order model with protection against a true second order model, the region matrices of the partial derivatives of the response function are

$$\mu_{11}^{*} = \left[\begin{array}{c|cccc}
0 & 0 & 0 & \cdots & 0 \\ \hline
0 & 1 & 0 & \cdots & 0 \\
. & 0 & 1 & . & 0 \\
. & . & 0 & . & . \\
0 & 0 & . & \cdots & 1
\end{array}\right]_{(p_1 \times p_1)}$$

$$\mu_{21}^{*} = \left[\begin{array}{c|ccccc}
0 & 2w_1 & 0 & \cdots & . & 0 \\
. & 0 & 2w_2 & 0 & \cdots & 0 \\
. & . & 0 & . & . & . \\
. & . & . & . & . & 0 \\
0 & 0 & \cdots & \cdots & 0 & 2w_k \\ \hline
0 & w_2 & w_1 & w_1 & \cdots & w_1 \\
. & . & 0 & 0 & \cdots & 0 \\
. & . & . & . & \cdots & 0 \\
. & w_k & 0 & 0 & \cdots & 0 \\
. & 0 & w_3 & w_2 & \cdots & w_2 \\
. & 0 & . & 0 & \cdots & 0 \\
. & 0 & w_k & 0 & . & . \\
. & . & 0 & . & . & 0 \\
0 & 0 & . & 0 & . & w_{k-1}
\end{array}\right]_{(p_2 \times p_1)}$$

$$\mu_{22}^{*} = \left[\begin{array}{cccc|cccc}
4w_{11} & 0 & \cdots & 0 & 2w_{12} & 2w_{13} & \cdots & 0 \\
0 & 4w_{22} & . & 0 & 2w_{12} & 0 & . & 0 \\
0 & 0 & . & . & 0 & 2w_{13} & . & . \\
. & . & . & . & . & 0 & . & . \\
0 & \cdots & \cdots & 4w_{kk} & 0 & \cdots & \cdots & 2w_{k-1,k} \\
\hline
2w_{12} & 2w_{12} & \cdots & 0 & w_{11}+w_{22} & w_{23} & \cdots & 0 \\
2w_{13} & 0 & \cdots & 0 & w_{23} & w_{11}+w_{33} & . & 0 \\
. & 0 & \cdots & 0 & . & . & . & . \\
2w_{1k} & 0 & \cdots & 0 & w_{2k} & w_{3k} & . & w_{1,k-1} \\
0 & 2w_{23} & 2w_{23} & 0 & w_{13} & w_{12} & . & 0 \\
0 & . & \cdots & 0 & . & 0 & . & . \\
0 & 2w_{2k} & 0 & . & w_{1k} & 0 & . & w_{2,k-1} \\
. & . & . & . & 0 & . & . & 0 \\
. & 0 & . & 0 & . & . & . & . \\
0 & \cdots & \cdots & 2w_{k-1,k} & 0 & \cdots & . & w_{k-1,k-1}+w_{kk}
\end{array}\right] (p_2 \times p_2)$$

where the w_{ii} elements of the region matrices are defined in Appendix D (pages 287-291).

Appendix D

Region Moments for Spherical and Cuboidal Regions

An experimental design is a plan which specifies numerical values for each input variable on a set of experimental runs. A collection of *N* experimental runs, for specific combinations of *k* input variables, has the design matrix

$$D = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{1N} & x_{2N} & \cdots & x_{kN} \end{bmatrix}$$

where x_{iu} $(i = 1, \ldots, k;\ u = 1, \ldots, N)$ represents the u^{th} level of the i^{th} input variable. Each row of *D* corresponds to an experimental run (or design *point*) and each column of *D* corresponds to an input variable (or design *variable*). For convenience, the design variables are generally centered and scaled such that the following two conditions are met:

$$\bar{x}_i = [i] = \frac{1}{N} \sum_{u=1}^{N} x_{iu} = 0$$

$$(i = 1, 2, \ldots, k)$$

$$\overline{x_i^2} = [ii] = \frac{1}{N} \sum_{u=1}^{N} x_{iu}^2 = 1 \ .$$

In addition to the coding of individual design *variables*, restrictions placed on the design *points* effect the shape of the *k*-dimensional experimental region. The two most frequently used regions shapes are the spherical and cuboidal regions. A spherical region is attained by requiring that each of the coded design points satisfy the inequality

$$\sum_{i=1}^{k} x_{iu}^2 \leq 1 \qquad (u = 1, 2, \ldots, N)$$

and a cuboidal region is attained by restricting the levels of the coded variables to satisfy the inequality

$$|x_{iu}| \leq 1 \qquad (i = 1, 2, \ldots, k), \qquad (u = 1, 2, \ldots, N).$$

Coded design points which satisfy either of these requirements as an equality fall on the outer boundary of the experimental region.

The region moment matrices, μ_{11}, μ_{21}, μ_{22}, and the region matrices of the partial derivatives of the response function, μ_{11}^*, μ_{21}^*, μ_{22}^*, are a function of the shape of the region of interest. The general form of the elements of the region matrices are defined as

$$w_{1^{\delta_1} 2^{\delta_2} \ldots\ldots k^{\delta_k}} = \Omega_r \int_R x_1^{\delta_1} \, x_2^{\delta_2} \, \ldots\ldots \, x_k^{\delta_k} \, d\underline{x}$$

where $\sum_{i=1}^{k} \delta_i = \delta$ is the order of the region moment and Ω_r^{-1} is the volume of the region.

Spherical Region

The elements of the region matrices for a spherical region of interest can be obtained by making use of the integral

$$\int_R x_1^{\delta_1} x_2^{\delta_2} \ldots x_k^{\delta_k} \, dx_1 \, dx_2 \ldots dx_k = \frac{\Gamma\left[\frac{(\delta_1+1)}{2}\right] \Gamma\left[\frac{(\delta_2+1)}{2}\right] \cdots \Gamma\left[\frac{(\delta_k+1)}{2}\right]}{\Gamma\left[\frac{\sum_{i=1}^{k}(\delta_i+1)}{2}+1\right]}$$

where $\Gamma[\ldots]$ refers to the gamma function. If any of the δ_i are equal to an odd number, then the value of the integral is zero. Therefore, all of the spherical region moments involving *odd* powers of x_i are equal to zero. Only region moments involving only *even* powers of x_i will take on values other than zero.

Utilizing the above integral, the volume of a spherical region in the coded design variables becomes

$$\Omega_r^{-1} = \int_R d\underline{x} = \frac{\Gamma\left[\frac{1}{2}\right]^k}{\Gamma\left[\frac{(k+2)}{2}\right]} = \frac{\pi^{k/2}}{\Gamma\left[\frac{(k+2)}{2}\right]} .$$

The nonzero, even-order, spherical region moments through order six become

$$w_{ii} = \Omega_r \int_R x_i^2 \, d\underline{x} = \frac{1}{k+2}$$

$$w_{iijj} = \Omega_r \int_R x_i^2 \, x_j^2 \, d\underline{x} = \frac{1}{(k+2)\,(k+4)}$$

$$w_{iiii} = \Omega_r \int_R x_i^4 \, d\underline{x} = \frac{3}{(k+2)(k+4)}$$

$$w_{iijjkk} = \Omega_r \int_R x_i^2 x_j^2 x_k^2 \, d\underline{x} = \frac{1}{(k+2)(k+4)(k+6)}$$

$$w_{iiiijj} = \Omega_r \int_R x_i^4 x_j^2 \, d\underline{x} = \frac{3}{(k+2)(k+4)(k+6)}$$

$$w_{iiiiii} = \Omega_r \int_R x_i^6 \, d\underline{x} = \frac{15}{(k+2)(k+4)(k+6)} \quad .$$

Cuboidal Region

The elements of the region matrices for a cuboidal region in the coded design variables can be obtained by integrating the region moments with respect to each x_i, between -1 and $+1$, as follows

$$\int_{-1}^{1} \cdots \int_{-1}^{1} x_1^{\delta_1} x_2^{\delta_2} \cdots x_k^{\delta_k} \, dx_1 \, dx_2 \ldots dx_k = \left[\frac{x_1^{\delta_1+1}}{\delta_1+1} \right|_{-1}^{1} \left[\frac{x_2^{\delta_2+1}}{\delta_2+1} \right|_{-1}^{1} \cdots \left[\frac{x_k^{\delta_k+1}}{\delta_k+1} \right|_{-1}^{1} \quad .$$

If any of the δ_i are odd, then the value of the integral is zero. Therefore, similar to a spherical region, all of the cuboidal region moments involving *odd* powers of x_i are equal to zero.

Utilizing the above integral, the volume of a cuboidal region in the coded design variables becomes

$$\Omega_r^{-1} = \int_{-1}^{1} \cdots \int_{-1}^{1} d\underline{x} = 2^k \, .$$

The nonzero, even-order, cuboidal region moments through order six become

$$w_{ii} = \Omega_r \int_{-1}^{1} \cdots \int_{-1}^{1} x_i^2 \, d\underline{x} = \frac{1}{3}$$

$$w_{iijj} = \Omega_r \int_{-1}^{1} \cdots \int_{-1}^{1} x_i^2 \, x_j^2 \, d\underline{x} = \frac{1}{9}$$

$$w_{iiii} = \Omega_r \int_{-1}^{1} \cdots \int_{-1}^{1} x_i^4 \, d\underline{x} = \frac{1}{5}$$

$$w_{iijjkk} = \Omega_r \int_{-1}^{1} \cdots \int_{-1}^{1} x_i^2 \, x_j^2 \, x_k^2 \, d\underline{x} = \frac{1}{27}$$

$$w_{iiiijj} = \Omega_r \int_{-1}^{1} \cdots \int_{-1}^{1} x_i^4 \, x_j^2 \, d\underline{x} = \frac{1}{15}$$

$$w_{iiiiii} = \Omega_r \int_{-1}^{1} \cdots \int_{-1}^{1} x_i^6 \, d\underline{x} = \frac{1}{7} \quad .$$

Appendix E

Alias Matrices in the First Order Case

Consider the situation in which an experimenter is fitting a first order polynomial model using a two-level factorial design and desires protection against a true second order model; that is, the fit-protection situation of $d_1 = 1$, $d_2 = 2$. The fitted model and the partitioned protection model, respectively, can be written in general linear model form as

$$\hat{\underline{y}} = X_1 \underline{b}_1$$

$$\underline{y} = X_1 \underline{\beta}_1 + X_2 \underline{\beta}_2 + \underline{\varepsilon}$$

where the subscript 1 indicates that the matrices include the linear terms in the fitted model and the subscript 2 indicates that the matrices include the unfitted interaction and quadratic terms in the protection model.

The ordinary least squares (OLS) estimator of $\underline{\beta}_1$ results in a minimum sum of squares of the errors ($\underline{\varepsilon}' \underline{\varepsilon}$) and is denoted as $\underline{b}_{1,OLS}$. The weighted least squares (WLS) estimator of $\underline{\beta}_1$ results in a minimum sum of squares of the weighted errors ($\underline{\varepsilon}' V^{-1} \underline{\varepsilon}$) and is denoted as $\underline{b}_{1,WLS}$. The equations for the OLS and WLS estimators of $\underline{\beta}_1$ are

$$\underline{b}_{1,OLS} = (X'_1 X_1)^{-1} X'_1 \underline{y}$$
$$\underline{b}_{1,WLS} = (X'_1 V^{-1} X_1)^{-1} X'_1 V^{-1} \underline{y}$$

where V is the correlation matrix of the sample responses, illustrated in Appendix A (pages 273-277) for the **IR**, **CR**, and **AR** correlation induction strategies.

If the experimenter has underfit the response relationship, and the unfitted terms in the protection model are needed to describe the true response surface, then the least squares estimator of $\underline{\beta}_1$ may be biased. The bias in $\underline{b}_1$ resulting from the unfitted terms in X_2 can be written as

$$\begin{aligned} \text{Bias}[\underline{b}_1] &= E[\underline{b}_1] - \underline{\beta}_1 \\ &= A\,\underline{\beta}_2 \end{aligned}$$

where A is the *alias*, or *bias*, matrix whose form depends on the least squares estimation technique. The alias matrices under OLS and WLS, respectively, are

$$A_{OLS} = (X'_1 X_1)^{-1} X'_1 X_2$$
$$A_{WLS} = (X'_1 V^{-1} X_1)^{-1} X'_1 V^{-1} X_2 \ .$$

Therefore, the expected value of the least squares estimators can be written as

$$E[\underline{b}_{1,OLS}] = \underline{\beta}_1 + A_{OLS}\,\underline{\beta}_2$$
$$E[\underline{b}_{1,WLS}] = \underline{\beta}_1 + A_{WLS}\,\underline{\beta}_2 \ .$$

The OLS alias matrix, which is not a function of the correlation matrix, V, can be written in terms of the moment matrices of the design as

$$A_{OLS} = M_{11}^{-1} M_{12}$$

where the moment matrices, M, are shown in Appendix B (pages 278-282) for the $d_1 = 1$, $d_2 = 2$ fit-protection situation. Because the two-level factorial designs are first order orthogonal designs, the design moments involving odd powers of x_i are equal to zero, and the M_{11}^{-1} and M_{12} matrices simplify to

$$M_{11}^{-1} = \left[\begin{array}{c|c} 1 & \underline{0}'_k \\ \hline \underline{0}_k & \frac{1}{[ii]} I_k \end{array}\right]_{(p_1 \times p_1)}$$

$$M_{12} = \left[\begin{array}{c|c} [ii]\,\underline{1}'_k & \underline{0}'_m \\ \hline \underline{0}_k\,\underline{0}'_k & \underline{0}_k\,\underline{0}'_m \end{array}\right]_{(p_1 \times p_2)}$$

where k = number of design variables

m = number of pairs of design variables = $\frac{1}{2} k (k - 1)$

p_1 = number of parameters in the fitted first order model = $k + 1$

p_2 = number of parameters in the protection model, but not in the fitted model = $k + m$.

Incorporating the M_{11}^{-1} and M_{12} matrices into the equation for A_{OLS}, the OLS alias matrix for a two-level factorial design becomes

$$A_{OLS} = M_{12} \ .$$

The form of the M_{12} matrix indicates that only the b_0 intercept term is biased, and the remaining *k* linear coefficients are unbiased. The expected value of the OLS estimator of the intercept term becomes

$$E\,[\,b_{0,OLS}] = \beta_0 + [ii] \sum_{i=1}^{k} \beta_{ii}$$

and the remaining coefficients of the first order model are unbiased; that is,

$$E[b_{i,OLS}] = \beta_i .$$

For the situation in which the responses are correlated and the experimenter is using WLS estimation, the alias matrix is a function of the inverted correlation matrix, V^{-1}. The correlation matrices for the IR, CR, and AR correlation induction strategies (V_{IR}, V_{CR}, and V_{AR}) are shown in Appendix A (pages 273-277). Incorporating these matrices into the A_{WLS} equation reveals that the WLS alias matrix for a two-level factorial design is simply a multiple of the OLS alias matrix. Denoting the scalar quantity which converts the OLS alias matrix into the WLS alias matrix as z, the relationship can be written as

$$A_{WLS} = z\, A_{OLS}$$

where z is a term whose form depends on the correlation induction strategy. For the IR strategy, in which $V_{IR} = I_N$, it follows that $A_{WLS} = A_{OLS}$, and therefore $z_{IR} = 1$.

Before defining z under the CR and AR correlation induction strategies it is useful to define some notation for the number of design points utilizing independent, common, and antithetic random number streams. In developing the notation, it was assumed that the designs may be augmented with any number of center runs and the factorial points of the designs are *not* replicated. Under these assumptions, the notation used to designate the number of design points in the 2^k designs is as follows:

$$\begin{aligned} N &= \text{total number of design points} \\ &= F + N_c \\ &= N_1 + N_2 \end{aligned}$$

where F = number of factorial points (2^k or 2^{k-p})
N_c = number of center points
N_1 = number of points utilizing *common* or *antithetic* random number streams
N_2 = number of points utilizing *independent* random number streams.

The number of design points using independent random number streams (N_2) for the **IR**, **CR**, and **AR** correlation induction strategies, respectively, become

IR Strategy: $N_2 = N$

CR Strategy:
$$N_2 = N_c - 1 \quad \text{if } N_c > 1$$
$$= 0 \quad \text{if } N_c \leq 1$$

AR Strategy:
$$N_2 = N_c - 2 \quad \text{if } N_c > 1$$
$$= N_c \quad \text{if } N_c \leq 1 .$$

For the **CR** strategy, the OLS and WLS alias matrices are equivalent only if there are no replicated center runs necessitating the use of independent random number streams. Otherwise, the following z_{CR} term is needed to convert the OLS alias matrix into the WLS alias matrix for the **CR** strategy:

$$z_{CR} = \frac{N}{N - N_2\,\rho_+ + N_1 N_2\,\rho_+} .$$

For the **AR** strategy, the OLS and WLS alias matrices are equivalent only if the design partitions into two orthogonal blocks and there are no replicated center runs within blocks. Otherwise, the following z_{AR} term is needed to convert the OLS alias matrix into the WLS alias matrix for the **AR** strategy:

$$z_{AR} = \frac{N}{N - N_2\,\rho_+ + \tfrac{1}{2} N_1 N_2\,(\rho_+ - \rho_-)} .$$

Thus, for an experimenter using a two-level factorial design and OLS estimation, the alias matrix is independent of the correlation induction strategy. However, for an experimenter using WLS, the alias matrix is a function of the correlation matrix, V, and therefore the scalar z term is needed to transform the OLS alias matrix into the WLS alias matrix.

Appendix F

Variances of First Order Coefficients

Consider the situation in which an experimenter is fitting a first order polynomial model using a two-level factorial design. The first order model can be written in general linear model form as

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon} \,.$$

The ordinary least squares (OLS) and the weighted least squares (WLS) estimators of $\underline{\beta}$, which result in minimum values of $(\underline{\varepsilon}'\,\underline{\varepsilon})$ and $(\underline{\varepsilon}'\,V^{-1}\,\underline{\varepsilon})$, respectively, are defined as

$$\underline{b}_{OLS} = (X'X)^{-1}X'\underline{y}$$

$$\underline{b}_{WLS} = (X'V^{-1}X)^{-1}X'V^{-1}\underline{y}$$

where V is the correlation matrix of the sample responses, illustrated in Appendix A (pages 273-277) for the **IR**, **CR**, and **AR** correlation induction strategies.

Under the assumptions that $E[\,\underline{\varepsilon}\,] = \underline{0}$ and $Var[\,\underline{\varepsilon}\,] = \sigma^2 V$, the variance-covariance matrices of the OLS and WLS estimators of $\underline{\beta}$ become

$$\mathrm{Var}[\underline{b}_{OLS}] = (X'X)^{-1} X' \; V \; X(X'X)^{-1} \sigma^2$$

$$\mathrm{Var}[\underline{b}_{WLS}] = (X' V^{-1} X)^{-1} \sigma^2 \; .$$

The above equations indicate that the variance-covariance matrix of $\underline{b}$ differs for the OLS and WLS estimators. In addition, both matrices are a function of V, and therefore depend on the correlation induction strategy. For the two-level factorial designs, the variance-covariance matrices are diagonal matrices and the diagonal elements (the variances of the first order model coefficients) are equivalent (with the exception of the variance of the intercept term).

In the remainder of this Appendix, the variance-covariance matrix of $\underline{b}$ is derived for each of the correlation induction strategies under both OLS and WLS estimation. The variances of the OLS estimators, which are presented first, are less complicated because inversion of the V matrix is not required.

For the IR correlation induction strategy, the variance-covariance matrix of the OLS estimator of $\underline{\beta}$ is denoted as $\mathrm{Var}[\underline{b}_{OLS}]_{IR}$. In the case of a two-level factorial design, this variance-covariance matrix becomes

$$\begin{aligned}
\mathrm{Var}[\underline{b}_{OLS}]_{IR} &= (X'X)^{-1} X' \quad V_{IR} \quad X(X'X)^{-1} \quad \sigma^2 \\
&= (X'X)^{-1} X' \quad I_N \quad X(X'X)^{-1} \quad \sigma^2 \\
&= (X'X)^{-1} \quad \sigma^2 \\
&= \mathrm{Diag}\left(\frac{1}{N}, \frac{1}{N[ii]}, \ldots\ldots, \frac{1}{N[ii]} \right) \sigma^2 \\
&= \mathrm{Diag}\left(\mathrm{Var}(b_0), \mathrm{Var}(b_i), \ldots\ldots, \mathrm{Var}(b_i) \right) .
\end{aligned}$$

For the CR correlation induction strategy, the variance-covariance matrix of the OLS estimator of $\underline{\beta}$ becomes

$$\begin{aligned}
\mathrm{Var}[\underline{b}_{OLS}]_{CR} &= (X'X)^{-1}X' \quad V_{CR} \quad X(X'X)^{-1}\sigma^2 \\
&= (X'X)^{-1}X' \quad I_N \quad X(X'X)^{-1} \; \sigma^2 \\
&\quad + (X'X)^{-1}X' \quad \rho_+ \, \underline{u}\,\underline{u}' \quad X(X'X)^{-1} \; \sigma^2 \\
&\quad - (X'X)^{-1}X' \quad \rho_+ \, U \quad X(X'X)^{-1} \; \sigma^2 \\
&= \mathrm{Diag}\left(\frac{1}{N}, \frac{1}{N[ii]}, \ldots\ldots, \frac{1}{N[ii]} \right) \sigma^2 \\
&\quad + \mathrm{Diag}\left(\frac{N_1^{\,2}\rho_+}{N^2}, 0, \ldots\ldots, 0 \right) \sigma^2 \\
&\quad - \mathrm{Diag}\left(\frac{N_1\rho_+}{N^2}, \frac{\rho_+}{N[ii]}, \ldots\ldots, \frac{\rho_+}{N[ii]} \right) \sigma^2 \\
&= \mathrm{Diag}\left(\frac{N + N_1^{\,2}\rho_+ - N_1\rho_+}{N^2}, \frac{1-\rho_+}{N[ii]}, \ldots\ldots, \frac{1-\rho_+}{N[ii]} \right) \sigma^2 \\
&= \mathrm{Diag}\left(\mathrm{Var}(b_0), \mathrm{Var}(b_i), \ldots\ldots, \mathrm{Var}(b_i) \right)
\end{aligned}$$

where N is the total number of design points, N_1 is the number of design points using *common* random number streams, and $N - N_1$ is the number of design points using *independent* random number streams. For designs with no center runs, $N_1 = F$, where F is the number of factorial points; for designs with center runs, $N_1 = F + 1$.

For the AR correlation induction strategy, the variance-covariance matrix of the OLS estimator of $\underline{\beta}$ becomes

$$\begin{aligned}
\mathrm{Var}[\underline{b}_{OLS}]_{AR} &= (X'X)^{-1}X' \quad V_{AR} \quad X(X'X)^{-1}\sigma^2 \\
&= (X'X)^{-1}X' \quad I_N \quad X(X'X)^{-1} \quad \sigma^2 \\
&\quad + (X'X)^{-1}X' \quad \tfrac{1}{2}(\rho_+ - \rho_-)\,\underline{u}\,\underline{u}' \quad X(X'X)^{-1} \quad \sigma^2 \\
&\quad + (X'X)^{-1}X' \quad \tfrac{1}{2}(\rho_+ + \rho_-)\,\underline{v}\,\underline{v}' \quad X(X'X)^{-1} \quad \sigma^2 \\
&\quad - (X'X)^{-1}X' \quad \rho_+ U \quad X(X'X)^{-1} \quad \sigma^2 \\
&= \mathrm{Diag}\left(\frac{1}{N}, \frac{1}{N[ii]}, \ldots\ldots, \frac{1}{N[ii]} \right) \sigma^2 \\
&\quad + \mathrm{Diag}\left(\frac{\tfrac{1}{2}N_1^2(\rho_+ - \rho_-)}{N^2}, 0, \ldots\ldots, 0 \right) \sigma^2 \\
&\quad + \mathrm{Diag}(0, 0, \ldots\ldots, 0)\ \sigma^2 \\
&\quad - \mathrm{Diag}\left(\frac{N_1\rho_+}{N^2}, \frac{\rho_+}{N[ii]}, \ldots\ldots, \frac{\rho_+}{N[ii]} \right) \sigma^2 \\
&= \mathrm{Diag}\left(\frac{N + \tfrac{1}{2}N_1^2(\rho_+ - \rho_-) - N_1\rho_+}{N^2}, \frac{1-\rho_+}{N[ii]}, \ldots\ldots, \frac{1-\rho_+}{N[ii]} \right) \sigma^2 \\
&= \mathrm{Diag}\left(\mathrm{Var}(b_0), \mathrm{Var}(b_i), \ldots\ldots, \mathrm{Var}(b_i) \right)
\end{aligned}$$

where N is the total number of design points, N_1 is the number of design points using *common* or *antithetic* random number streams, and $N - N_1$ is the number of design points using *independent* random number streams. For designs in two blocks having less than three center runs, $N_1 = F$; for designs in two blocks having three or more center runs, $N_1 = F + 2$.

The variance-covariance matrices of $\underline{b}_{OLS}$ for the three correlation induction strategies indicate that the covariance between any two linear coefficients is zero in the case of a two-level factorial design. Also, the variances of the first order coefficients (except the intercept term) are equal and can be written as

$$\mathrm{Var}(b_i) = \frac{1-\rho_+}{N\,[ii]}\,\sigma^2$$

where ρ_+ is equal to zero for the IR correlation induction strategy. Thus, for a specified magnitude of the induced positive correlation, ρ_+, the variances of the least squares estimators are the same for the CR and AR strategies.

The variance of the intercept term is different for each of the correlation induction strategies. The variances of the OLS estimators of β_0 for the IR, CR, and AR strategies, respectively, become

$$\mathrm{Var}(b_0)_{OLS,IR} = \frac{1}{N}\,\sigma^2$$

$$\mathrm{Var}(b_0)_{OLS,CR} = \frac{N + N_1^2\,\rho_+ - N_1\,\rho_+}{N^2}\,\sigma^2$$

$$\mathrm{Var}(b_0)_{OLS,AR} = \frac{N + \tfrac{1}{2}\,N_1^2\,(\rho_+ - \rho_-) - N_1\,\rho_+}{N^2}\,\sigma^2 \;.$$

Having completed the derivations of the variances of the OLS estimators of $\underline{\beta}$, the variances of the WLS estimators are presented next. The variances of the k first order coefficients are the same under OLS and WLS, but the variance of the intercept term is the same for the IR, *pure* CR, and *pure* AR correlation induction strategies only. When the CR and AR correlation induction strategies are *modified* to allow for replicated center runs, the variance of the intercept term is different under OLS and WLS estimation. The variances of the WLS estimators of the intercept term, for IR, CR, and AR correlation induction strategies, respectively, can be written as

$$\operatorname{Var}(b_0)_{\text{WLS,IR}} = \frac{1}{N}\,\sigma^2$$

$$\operatorname{Var}(b_0)_{\text{WLS,CR}} = \frac{1-\rho_+ + N_1\rho_+}{N - N_2\rho_+ + N_1N_2\rho_+}\,\sigma^2$$

$$\operatorname{Var}(b_0)_{\text{WLS,AR}} = \frac{1-\rho_+ + \tfrac{1}{2}N_1(\rho_+ - \rho_-)}{N - N_2\rho_+ + \tfrac{1}{2}N_1N_2(\rho_+ - \rho_-)}\,\sigma^2 .$$

For the special case in which the CR and AR correlation induction strategies do not necessitate the use of independent random number streams; that is, $N_2 = 0$ and $N_1 = N$, the variances of the intercept term are the same under OLS and WLS estimation. The variances of the intercept term for the *pure* CR and *pure* AR strategies, respectively, become

$$\operatorname{Var}(b_0)_{\text{CR}} = \frac{1-\rho_+ + N\rho_+}{N}\,\sigma^2$$

$$\operatorname{Var}(b_0)_{\text{AR}} = \frac{1-\rho_+ + \tfrac{1}{2}N(\rho_+ - \rho_-)}{N}\,\sigma^2 .$$

Appendix G

B-Optimal First Order Designs

The bias component of the MSE of slope criteria, **B**, was considered in section 3.2.1 of this paper. The Min-**B** value of the pure second order design moment for a two-level factorial design was shown to be the following:

$$[ii] = \frac{w_{ii}}{z}$$

where z is the scalar term that converts the OLS alias matrix into the WLS alias matrix. Therefore, $z = 1$ for an OLS analysis, and the WLS equations for z under the **IR**, **CR**, and **AR** correlation induction strategies (as shown in Appendix E on page 296) become

$$z_{IR} = 1$$

$$z_{CR} = \frac{N}{N - N_2\,\rho_+ + N_1 N_2\,\rho_+}$$

$$z_{AR} = \frac{N}{N - N_2\,\rho_+ + \tfrac{1}{2} N_1 N_2\,(\rho_+ - \rho_-)}$$

where N_1 is the number of design points assigned *common* or *antithetic* random number streams and N_2 is the number of design points assigned *independent* streams.

Under OLS estimation, $z = 1$, and the **B**-optimal value of the second order design moment is $[ii] = w_{ii}$. Therefore, a Min-**B** design has pure second order design moments of $[ii] = 1/3$ if the region is cuboidal, or $[ii] = 1/(k+2)$ if the region is spherical. Under WLS estimation, the **B**-optimal value of [ii] is a function of N, N_1, N_2, ρ_+, ρ_-, and k, or equivalently, [ii] is a function of ρ_+, ρ_-, k, and N_c (the number of center runs).

The figures in the remainder of this appendix are plots of the the **B**-optimal values of [ii] under WLS estimation. The optimal values of [ii] are plotted versus ρ_+, ρ_-, k, and N_c for spherical and cuboidal regions of interest. There are plots for both the CR and AR correlation induction strategies, and the results for the IR strategy are indicated whenever $\rho_+ = \rho_- = 0$. The values of k correspond to the number of factors in the following two-level factorial designs:

1. Full 2^k factorial designs with $k = 1, 2, 3, 4$ factors.
2. One-half fractions of the 2^k factorial designs with $k = 5, 6$ factors.
3. One-quarter fractions of the 2^k designs with $k = 7$ factors.

Figures G-1 through G-8 present the results for the CR strategy and Figures G-9 through G-14 present the results for the AR strategy. Each *odd*-numbered figure corresponds to a cuboidal region of interest and the immediately following *even*-numbered figure corresponds to a spherical region of interest.

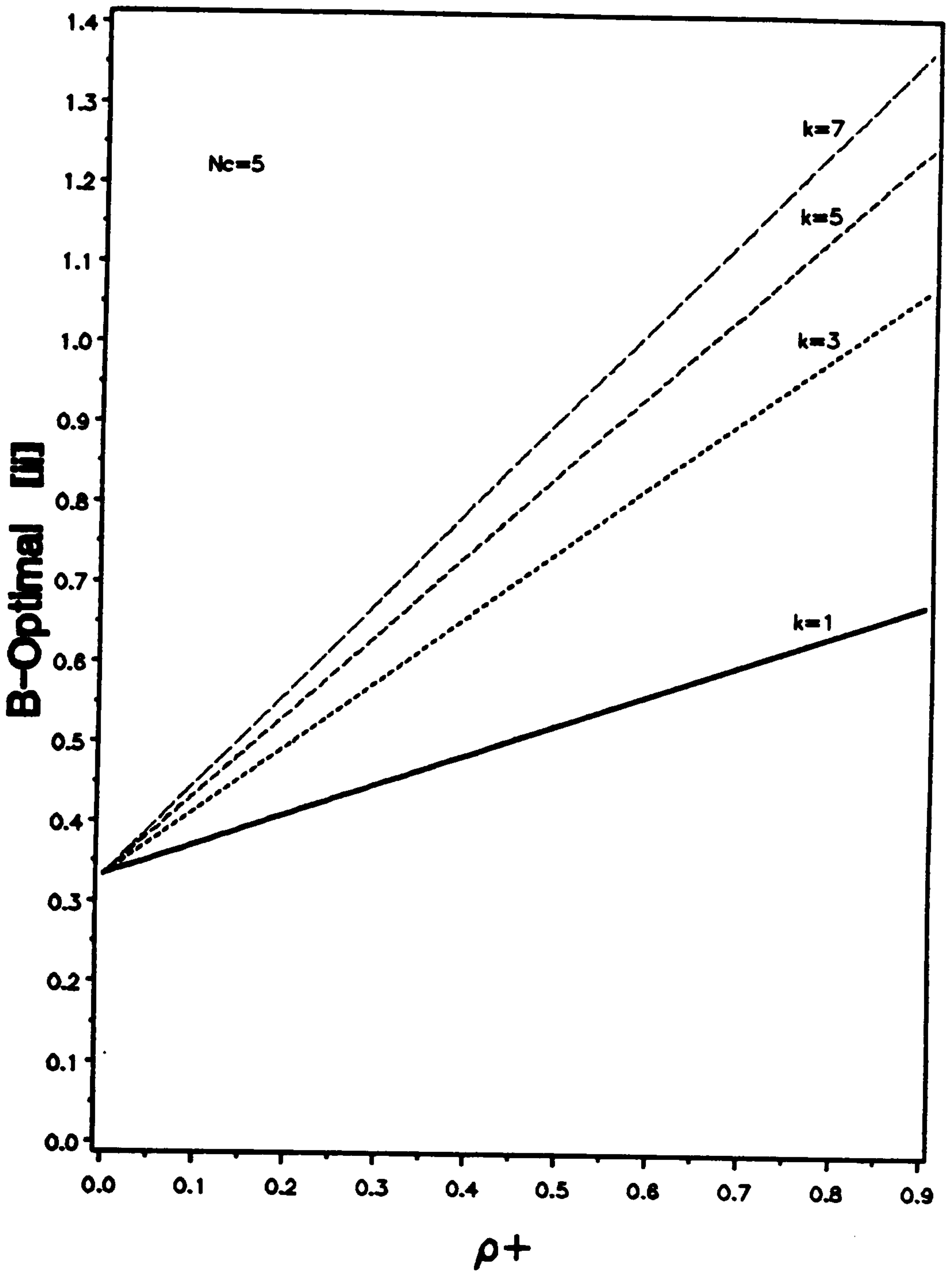


Figure G-1 First Order B-Optimal [ii] versus ρ_+ for the CR Strategy in a Cuboidal Region with $N_c = 5$.

CR Strategy under Weighted Least Squares
Two-level Factorial Design with $N_c = 5$ center runs.
Values of $k = 1, 3, 5, 7$.

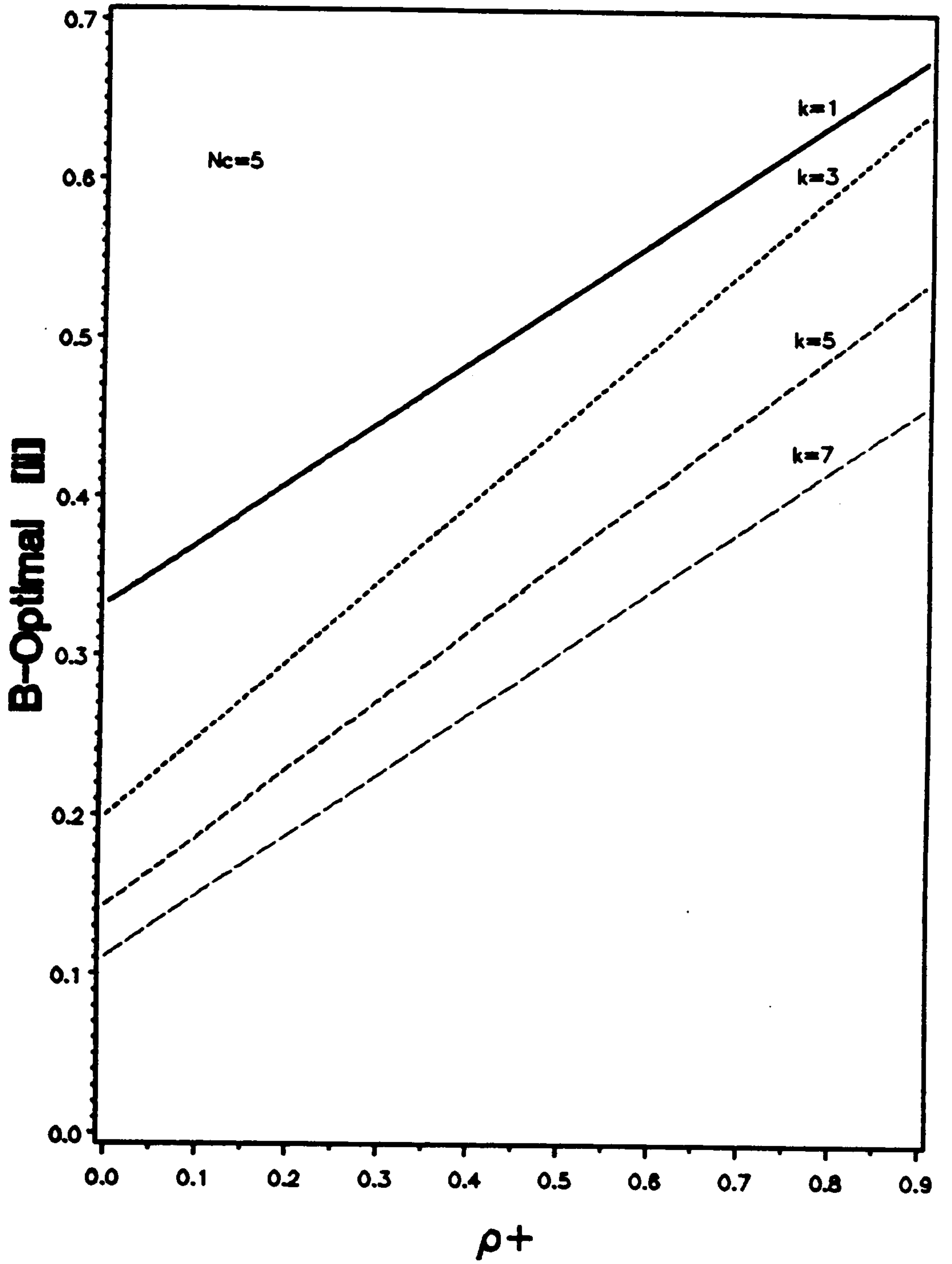


Figure G-2 First Order B-Optimal [ii] versus ρ_+ for the CR Strategy in a Spherical Region with $N_c = 5$.

CR Strategy under Weighted Least Squares
Two-level Factorial Design with $N_c = 5$ center runs.
Values of $k = 1, 3, 5, 7$.

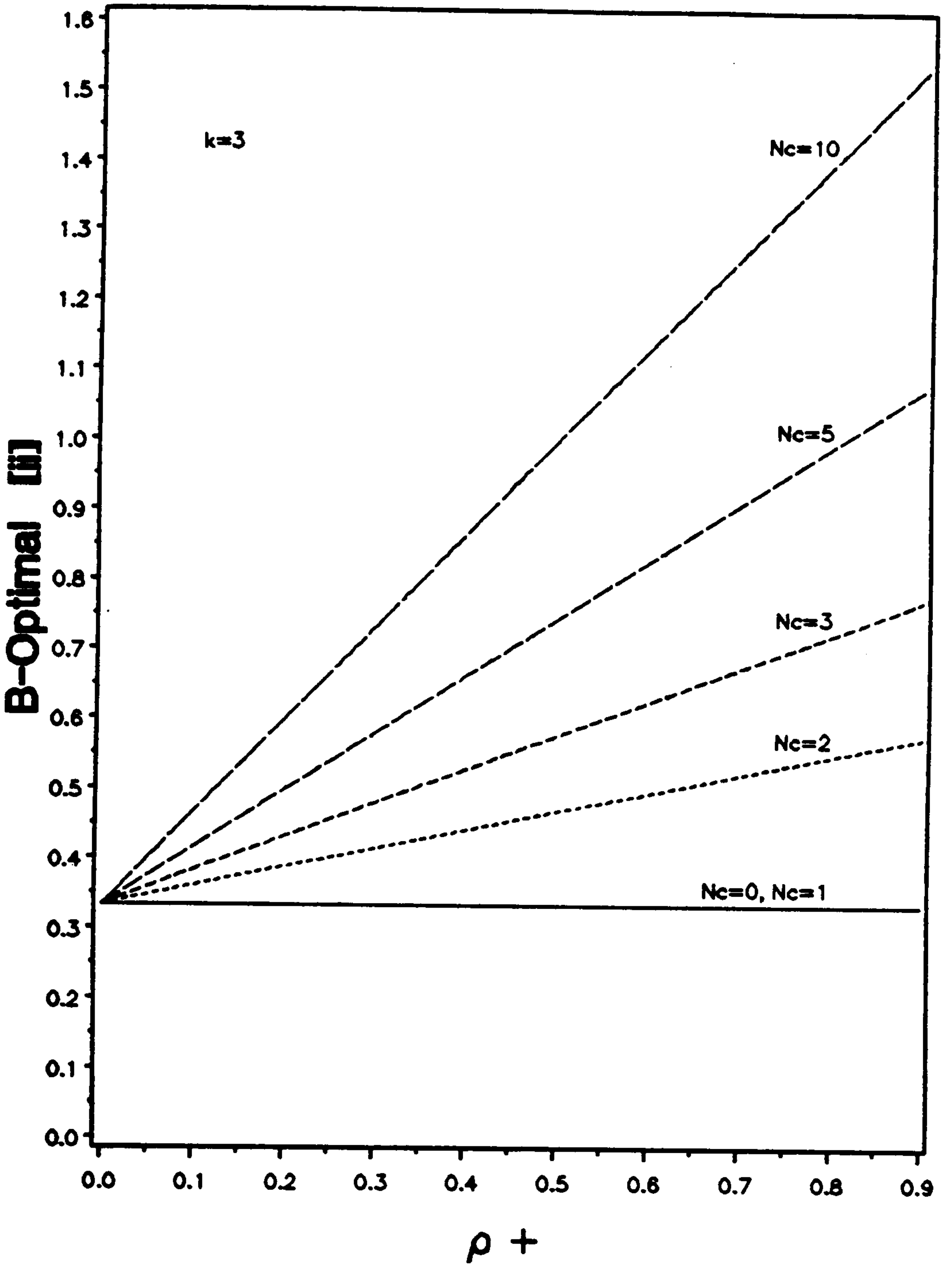


Figure G-3 First Order B-Optimal [ii] versus ρ_+ for the CR Strategy in a Cuboidal Region with $k=3$.

CR Strategy under Weighted Least Squares
Two-level Factorial Design with $k = 3$ factors.
Values of $N_c = 0, 1, 2, 3, 5, 10$.

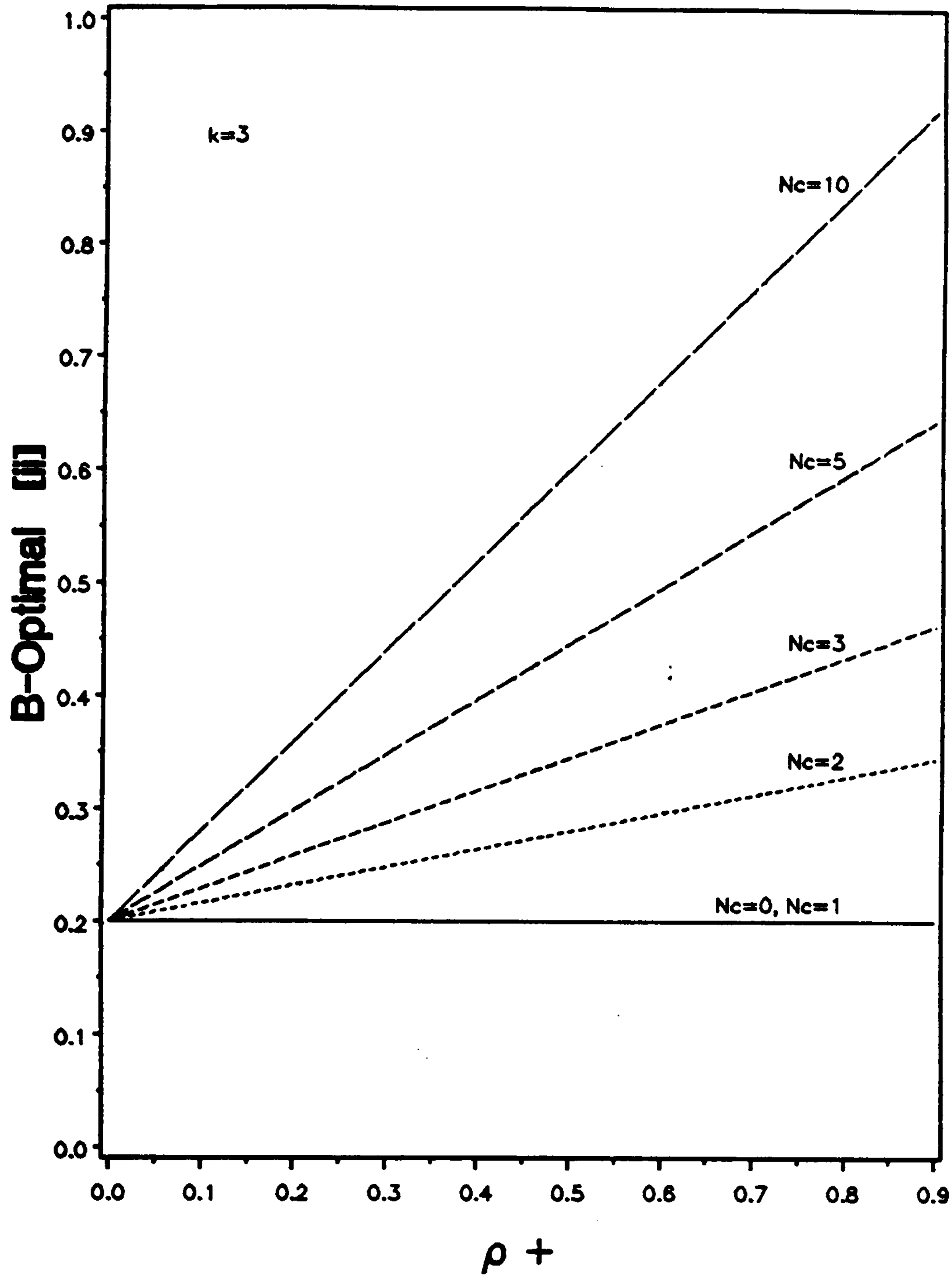


Figure G-4 **First Order B-Optimal [ii] versus ρ_+ for the CR Strategy in a Spherical Region with k = 3.**

CR Strategy under Weighted Least Squares
Two-level Factorial Design with k = 3 factors.
Values of N_c = 0, 1, 2, 3, 5, 10.

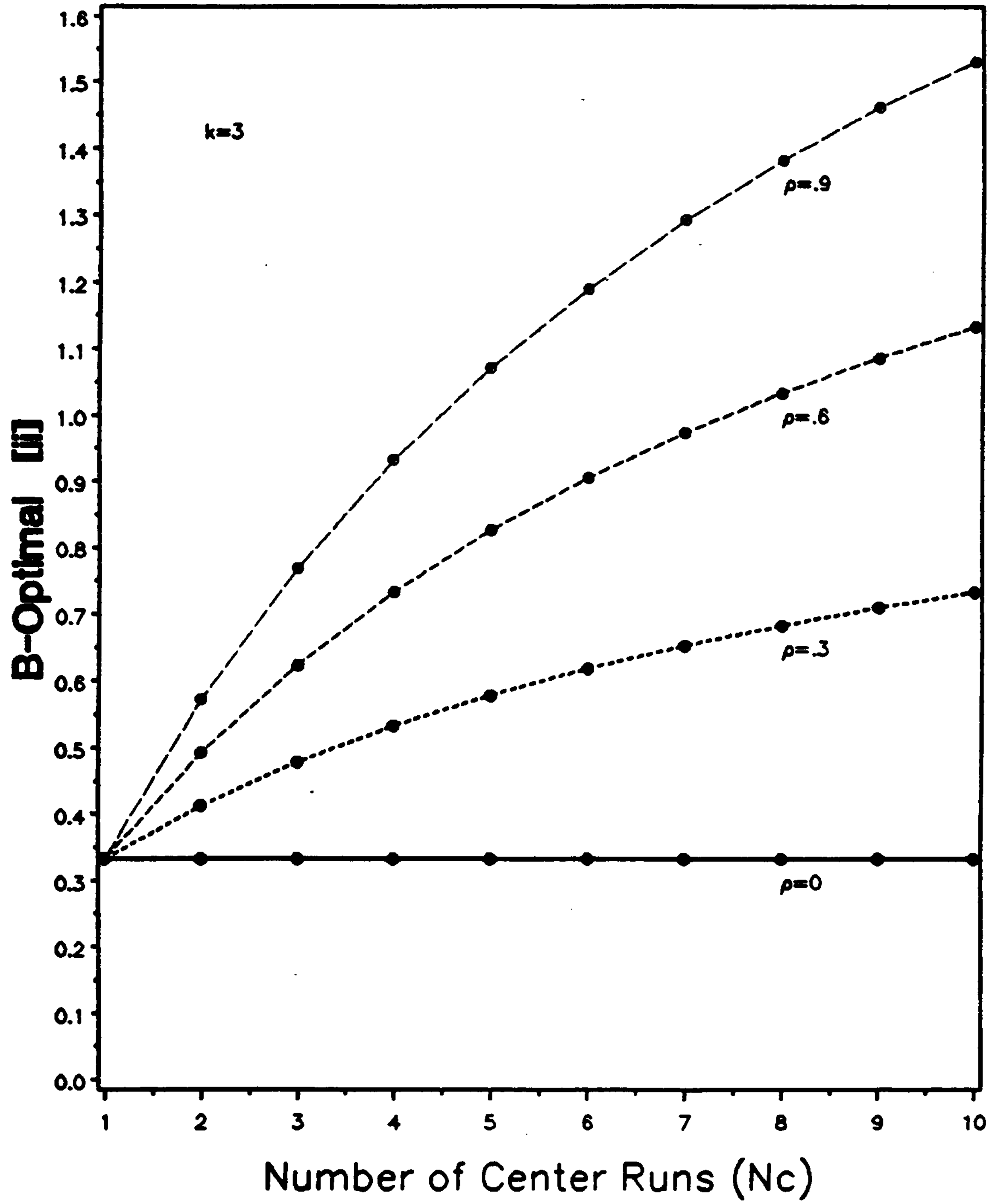


Figure G-5 **First Order B-Optimal [ii] versus N_c for the CR Strategy in a Cuboidal Region with $k = 3$.**

CR Strategy under Weighted Least Squares
Two-level Factorial Design with $k = 3$ factors.
Values of $\rho_+ = 0, .3, .6, .9$.

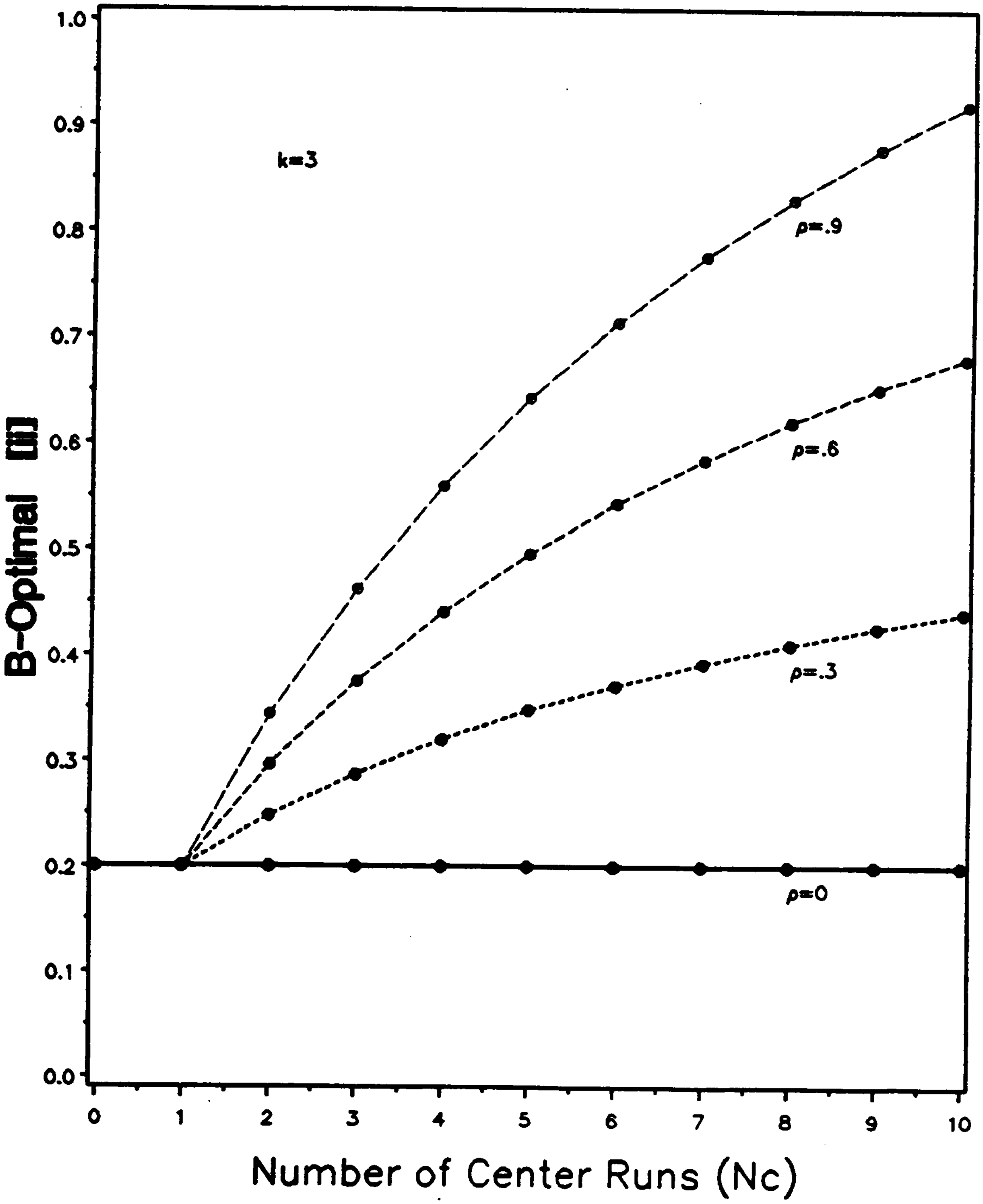


Figure G-6 **First Order B-Optimal [ii] versus N_c for the CR Strategy in a Spherical Region with k = 3.**

CR Strategy under Weighted Least Squares
Two-level Factorial Design with k = 3 factors.
Values of ρ_{+} = 0, .3, .6, .9.

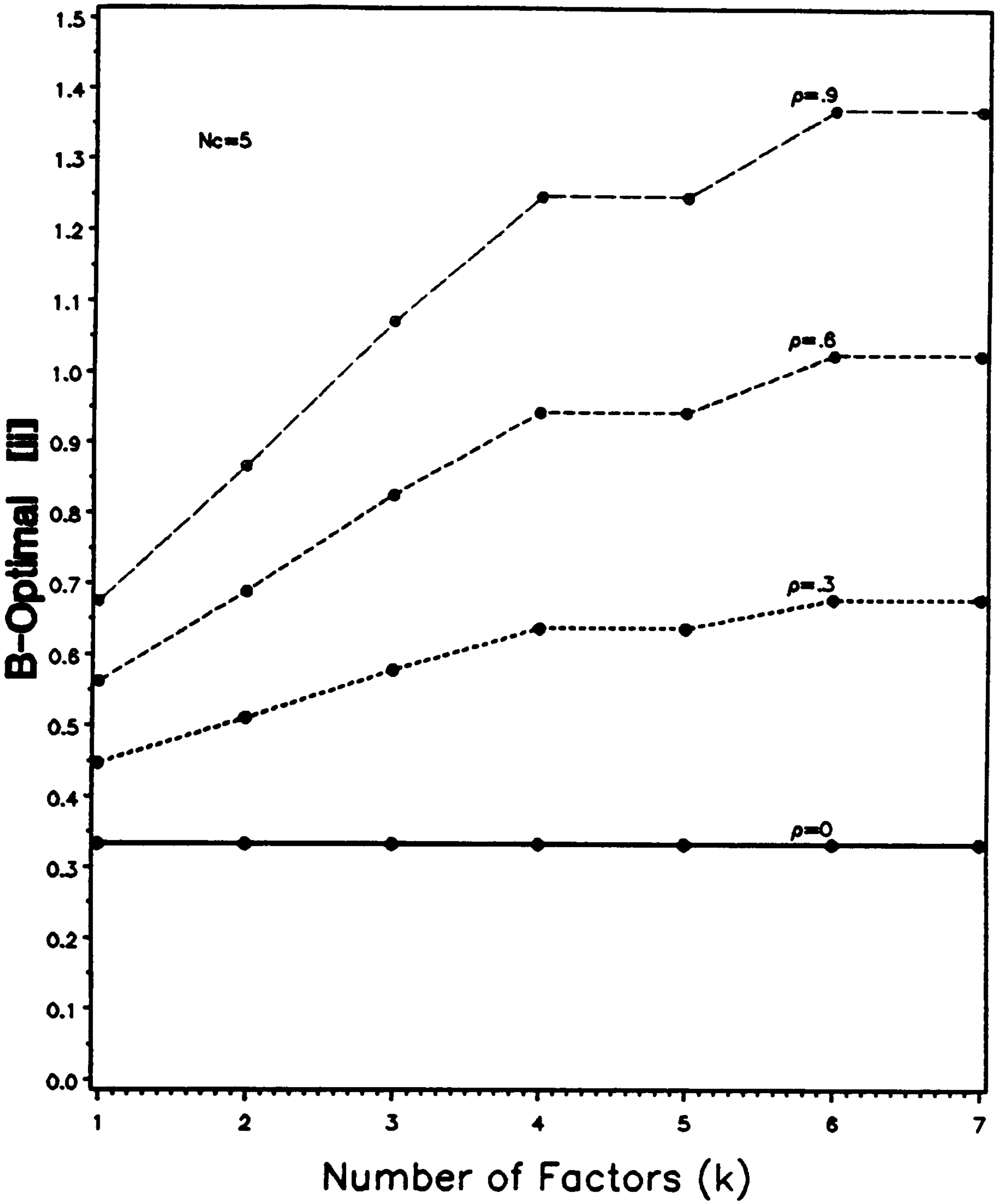


Figure G-7 First Order B-Optimal [ii] versus k for the CR Strategy in a Cuboidal Region with $N_c = 5$.

CR Strategy under Weighted Least Squares
Two-level Factorial Design with $N_c = 5$ center runs.
Values of $\rho_+ = 0, .3, .6, .9$.

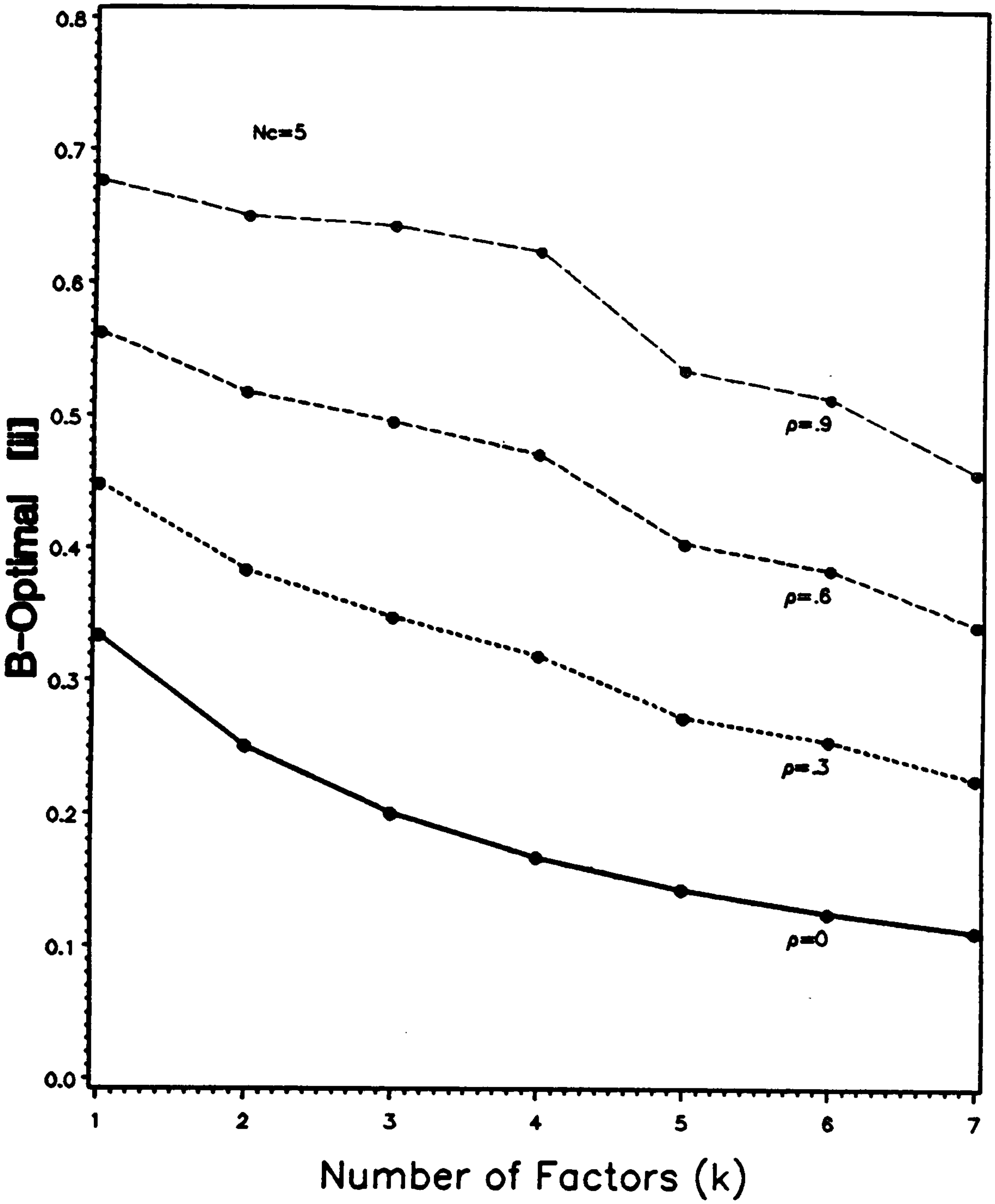


Figure G-8 First Order B-Optimal [ii] versus *k* for the CR Strategy in a Spherical Region with $N_c = 5$.

CR Strategy under Weighted Least Squares
Two-level Factorial Design with $N_c = 5$ center runs.
Values of $\rho_+ = 0, .3, .6, .9$.

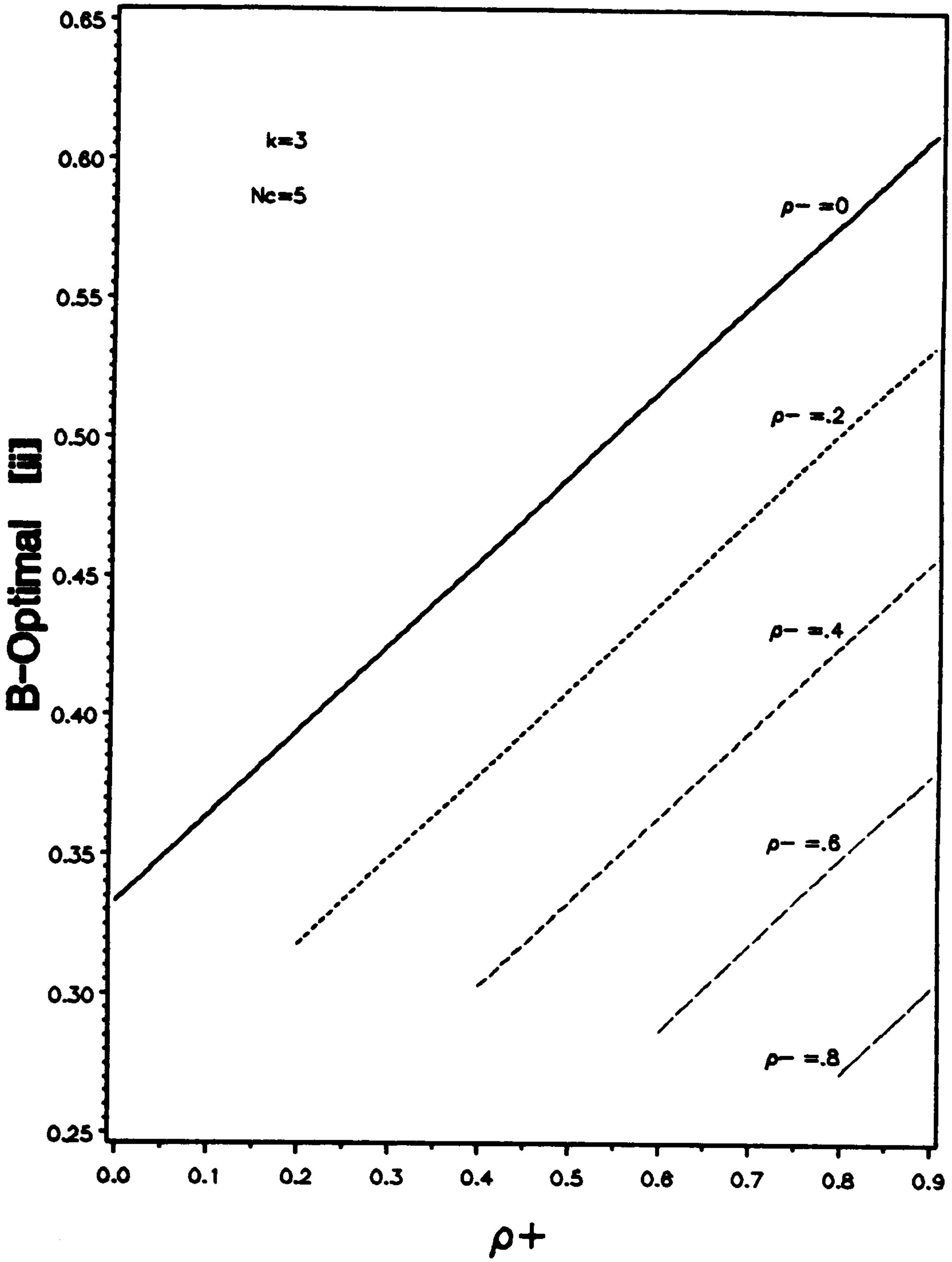


Figure G-9 **First Order B-Optimal [ii] versus ρ_+ for the AR Strategy in a Cuboidal Region with $N_c = 5$.**

AR Strategy under Weighted Least Squares
Two-level Factorial Design with $k = 3$ factors and $N_c = 5$ center runs.
Values of $\rho_- = 0, .2, .4\ .6, .8$ and $\rho_- \leq \rho_+$.

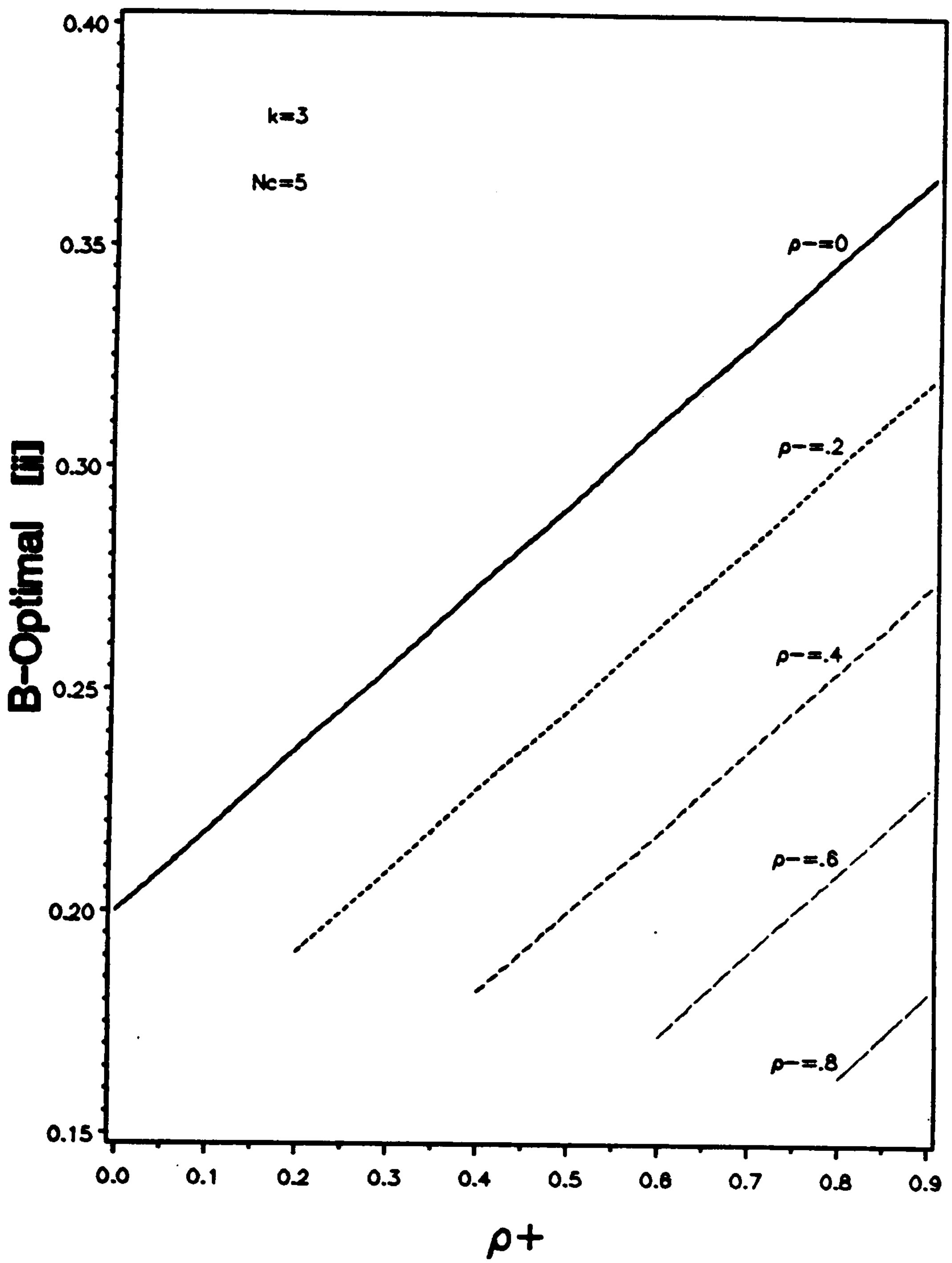


Figure G-10 First Order B-Optimal [ii] versus ρ_+ for the AR Strategy in a Spherical Region with $N_c = 5$.

AR Strategy under Weighted Least Squares

Two-level Factorial Design with $k = 3$ factors and $N_c = 5$ center runs.

Values of $\rho_- = 0, .2, .4 .6, .8$ and $\rho_- \le \rho_+$.

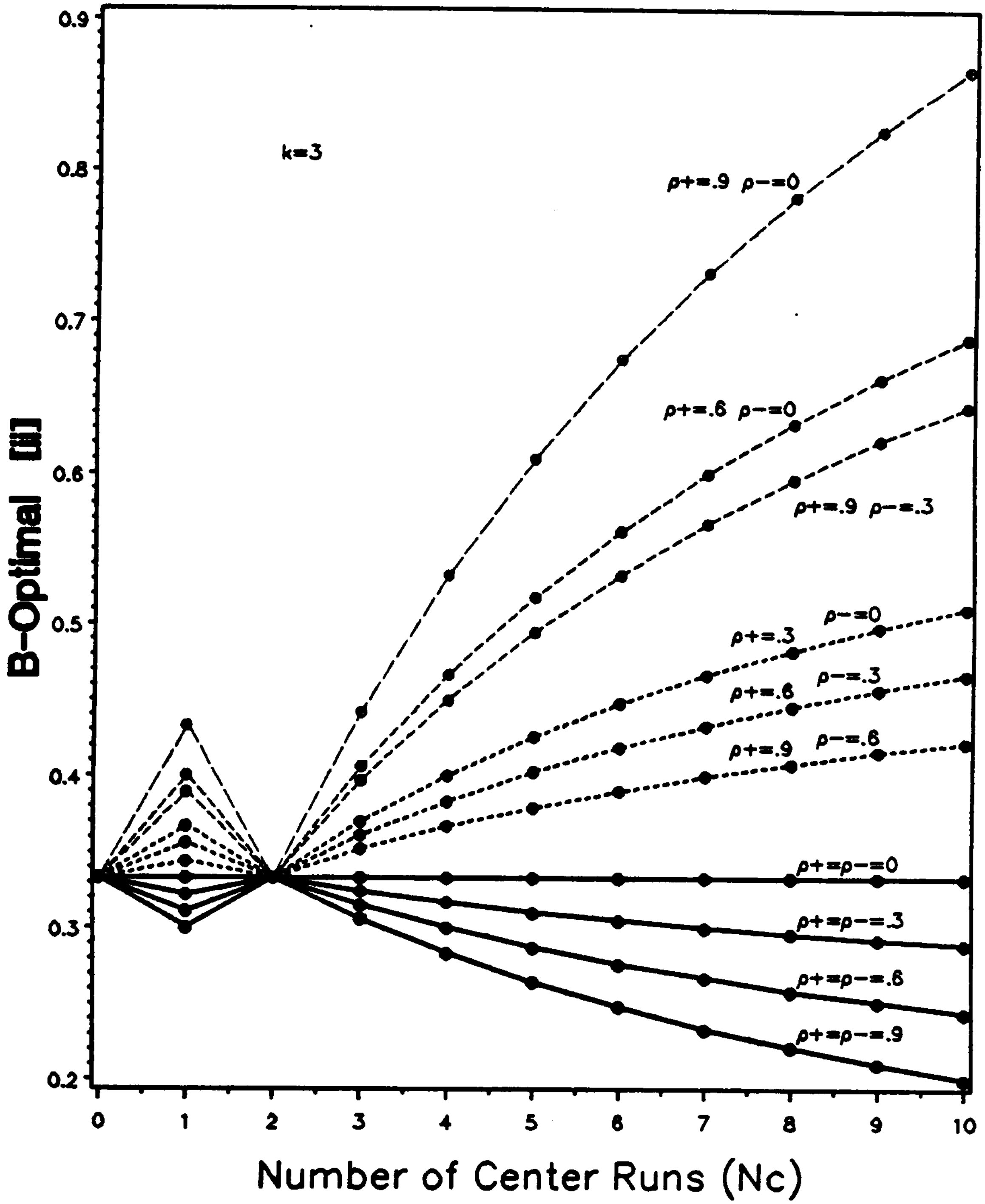


Figure G-11 **First Order B-Optimal [ii] versus N_c for the AR Strategy in a Cuboidal Region with k = 3.**

AR Strategy under Weighted Least Squares
Two-level Factorial Design with k = 3 factors.
Values of $\rho_- \leq \rho_+ = 0, .3, .6, .9$.

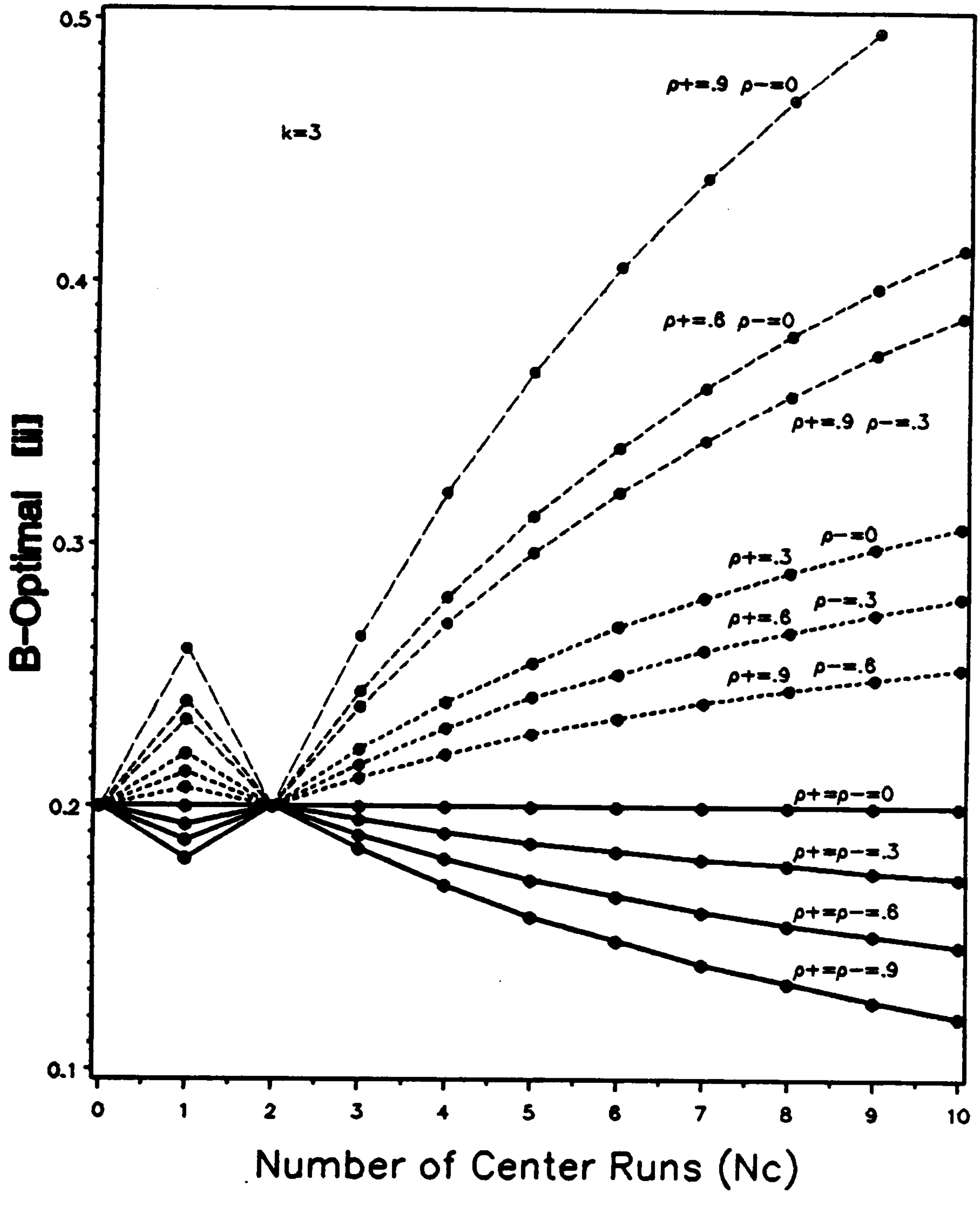


Figure G-12 First Order B-Optimal [ii] versus N_c for the AR Strategy in a Spherical Region with k = 3.

AR Strategy under Weighted Least Squares
Two-level Factorial Design with k = 3 factors.
Values of $\rho_- \leq \rho_+ = 0, .3, .6, .9$.

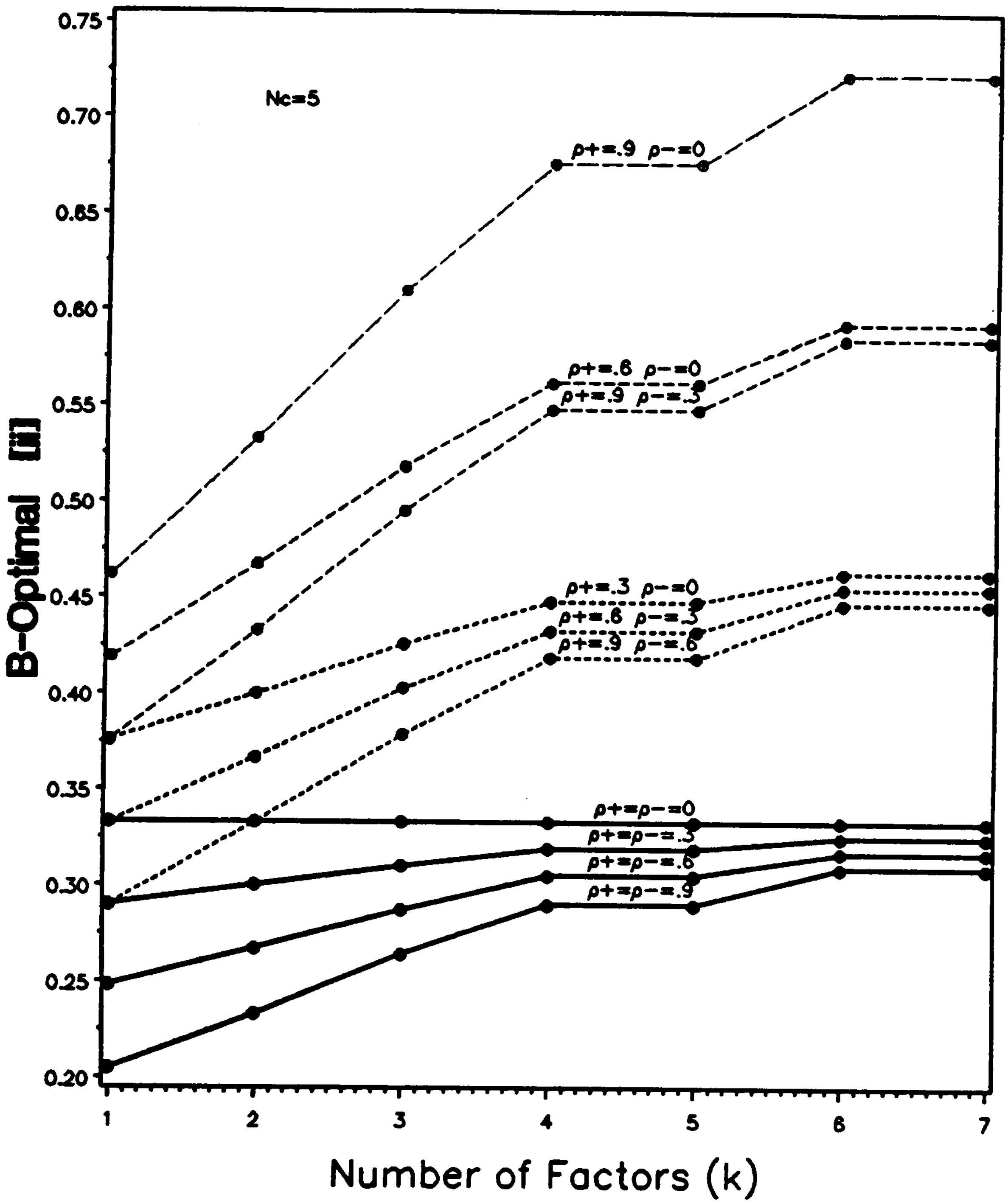


Figure G-13 First Order B-Optimal [ii] versus k for the AR Strategy in a Cuboidal Region with $N_c = 5$.

AR Strategy under Weighted Least Squares
Two-level Factorial Design with $N_c = 5$ center runs.
Values of $\rho_- \leq \rho_+ = 0, .3, .6, .9$.

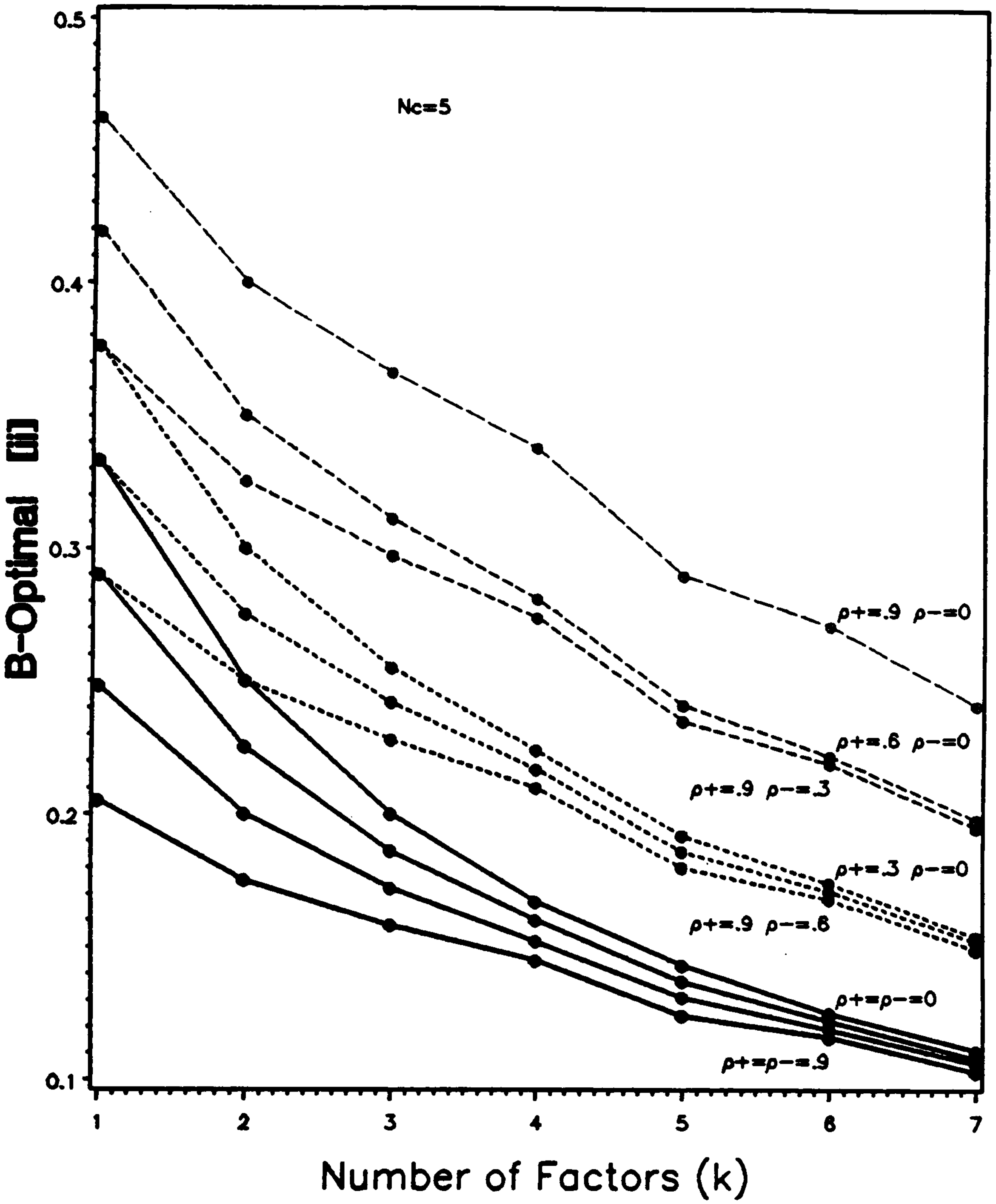


Figure G-14 First Order B-Optimal [ii] versus k for the AR Strategy in a Spherical Region with $N_c = 5$.

AR Strategy under Weighted Least Squares
Two-level Factorial Design with $N_c = 5$ center runs.
Values of $\rho_- \leq \rho_+ = 0, .3, .6, .9$.

Appendix H

J-Optimal First Order Designs in a Cuboidal Region

The figures presented in Chapter 3 illustrated the J-optimal two-level factorial designs for the correlation induction strategies in a spherical region of interest. Due to the similarity of the results for spherical and cuboidal regions, the results for cuboidal regions were omitted from Chapter 3 and are presented in this appendix. Figures H-1 through H-21 are the plots for cuboidal regions that correspond to the plots for spherical regions presented in Figures 1 through 21.

The *line* convention used to denote the values of ρ_+ and ρ_- in each of the figures is as follows:

Type of Line	CR Strategy	AR Strategy
1. ————————	$\rho_+ = 0$	$\rho_+ - \rho_- = 0$
2. ------------------	$\rho_+ = .3$	$\rho_+ - \rho_- = .3$
3. — — — — — —	$\rho_+ = .6$	$\rho_+ - \rho_- = .6$
4. —— —— —— ——	$\rho_+ = .9$	$\rho_+ - \rho_- = .9$.

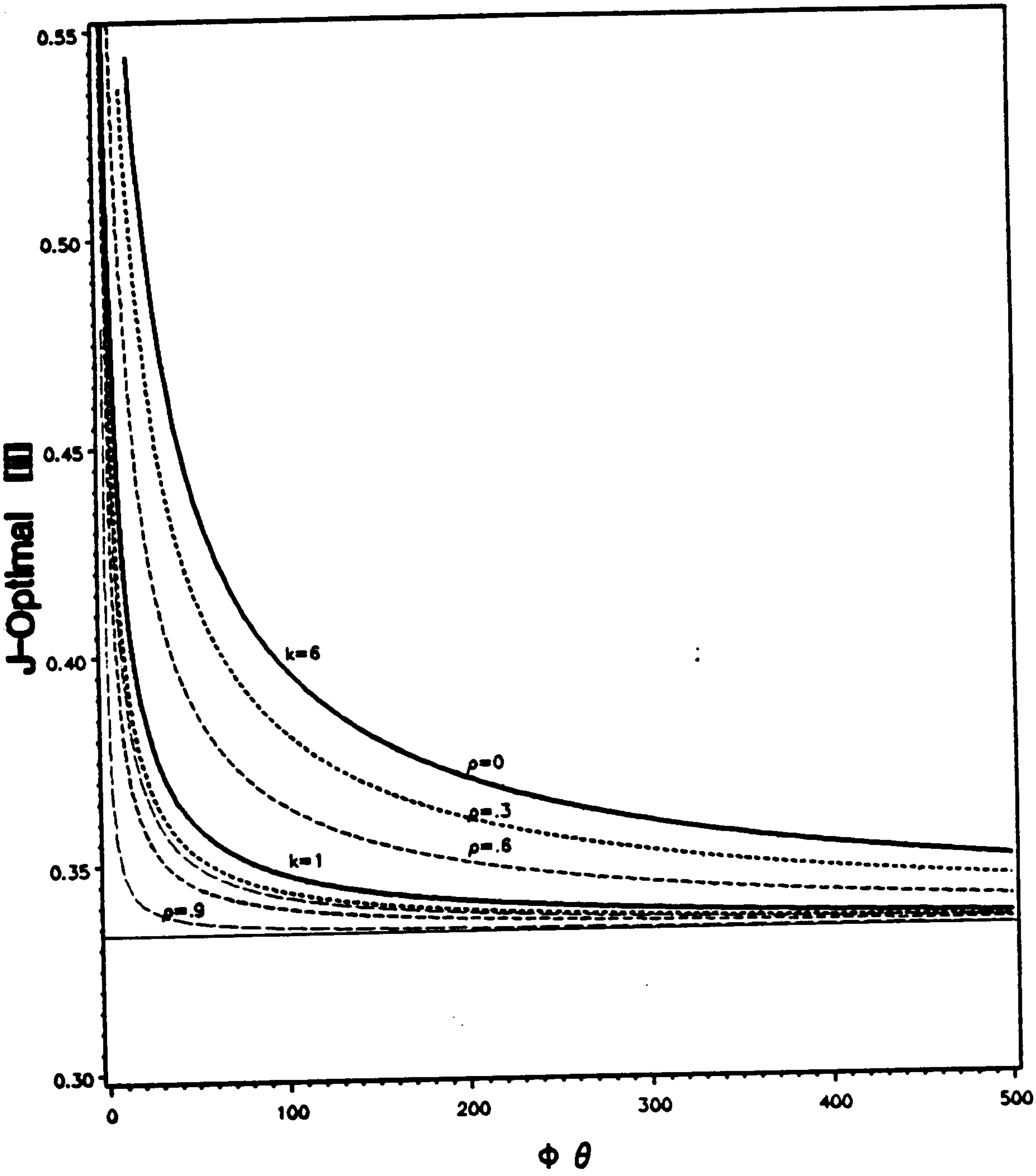


Figure H-1 First Order J-optimal [ii] versus $\phi\theta$ under OLS.
Region of interest is Cuboidal.
Values of $k = 1, 6$ and values of $\rho_{*} = 0, .3, .6, .9$.
The V-optimal value of the second moment is [ii] = 1.0.
The horizontal line at [ii] = .33 indicates the B-optimal value.

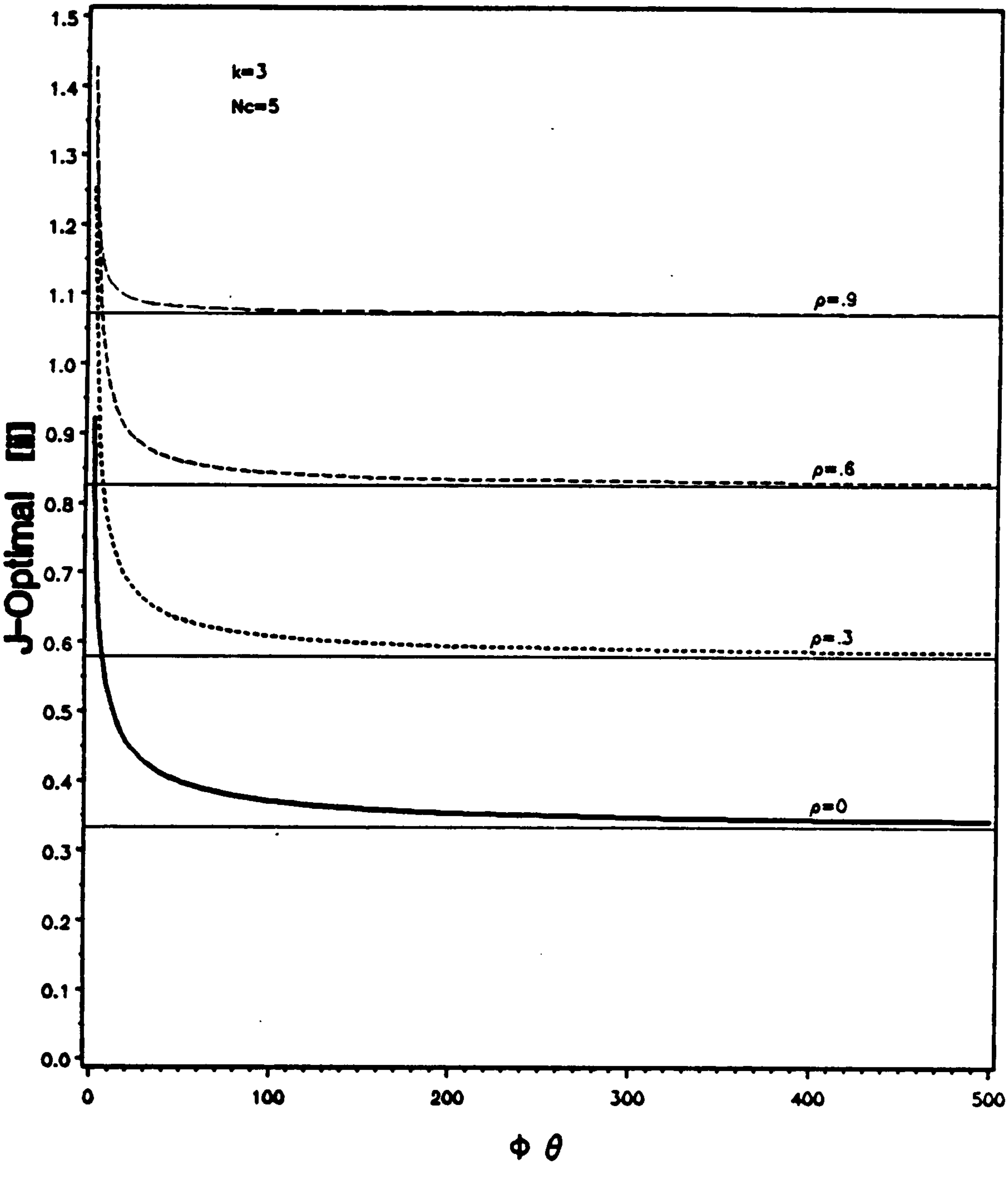


Figure H-2 **J-optimal [ii] versus $\phi\theta$ for the Modified CR Strategy under WLS.**

Region of interest is Cuboidal.

Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.

Values of $\rho_* = 0, .3, .6, .9$.

The V-optimal value of the second moment is $[ii] = 1.0$.

The horizontal lines indicate the B-optimal values of [ii].

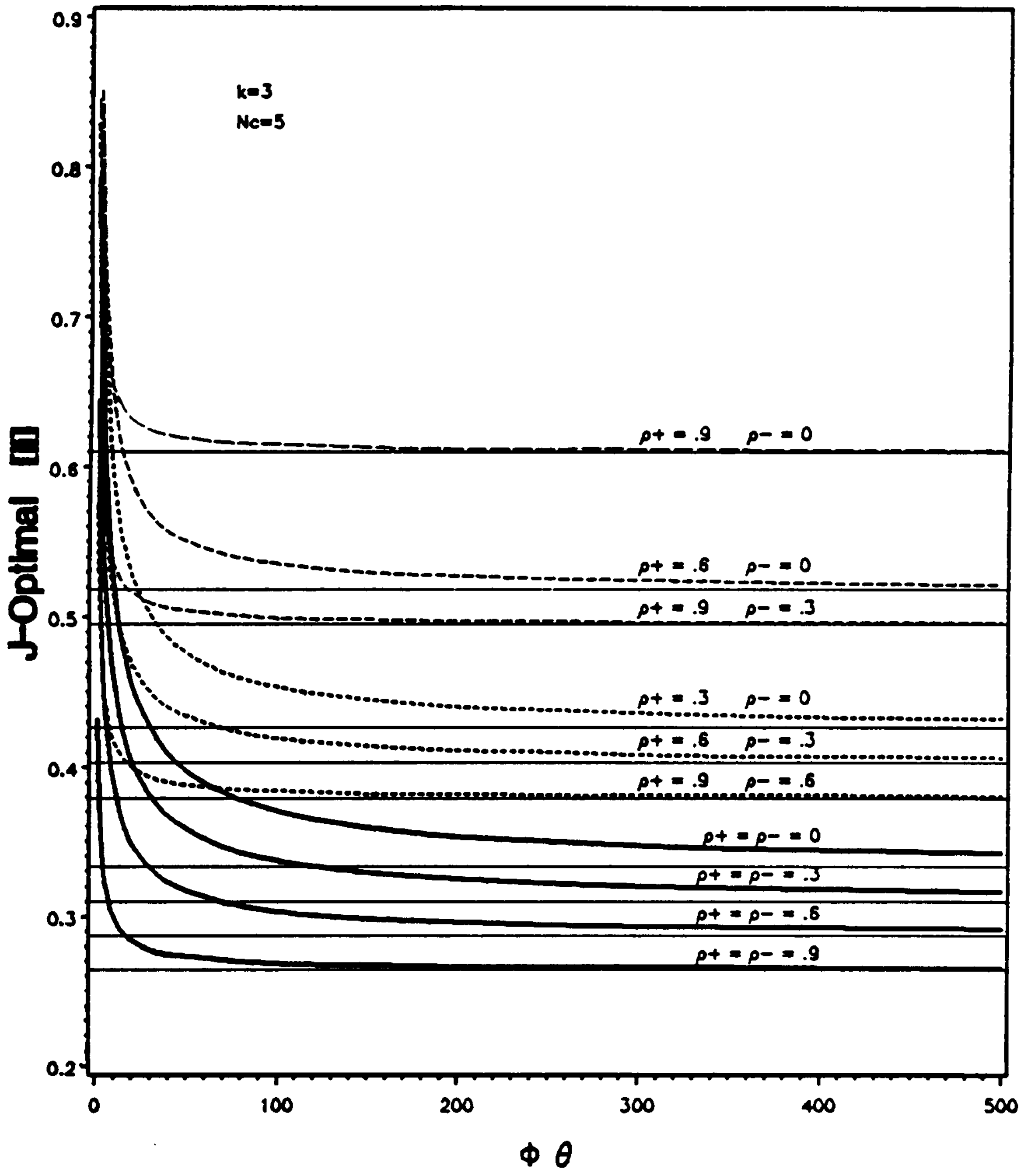


Figure H-3 **J-optimal [ii] versus $\phi\theta$ for the Modified AR Strategy under WLS.**
Region of interest is Cuboidal.
Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.
Values of $\rho_- \leq \rho_+ = 0, .3, .6, .9$.
The V-optimal value of the second moment is [ii] = 1.0.
The horizontal lines indicate the B-optimal values of [ii].

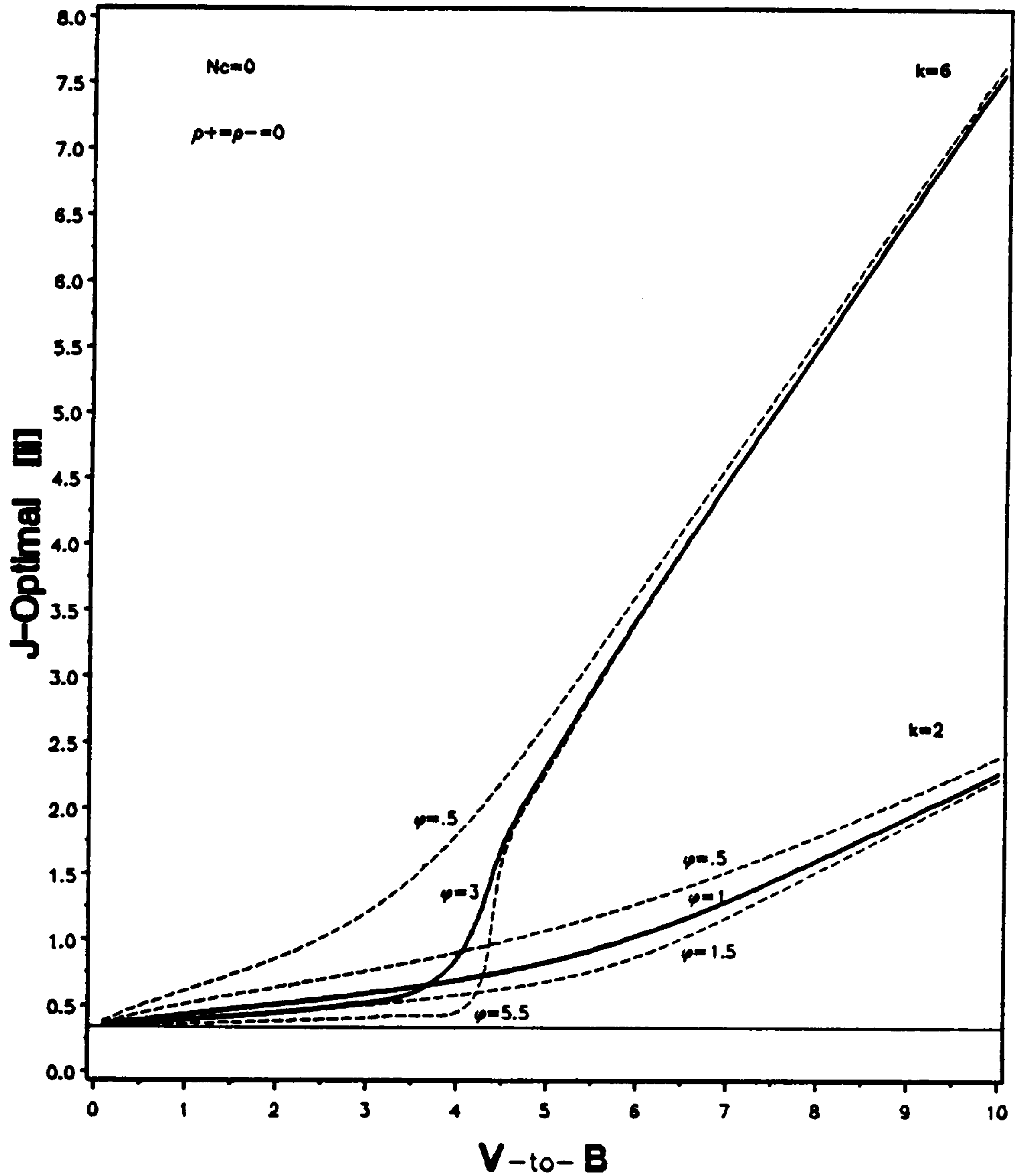


Figure H-4 J-optimal [ii] versus V/B for various ϕ.

IR Strategy in a Cuboidal region.
Two-level factorial design with $k = 2$ and $k = 6$ factors.
Values of ϕ = .5, k, k - .5 .
The V-optimal value of the second moment is [ii] = 1.0.
The horizontal line at [ii] = .33 indicates the B-optimal value.

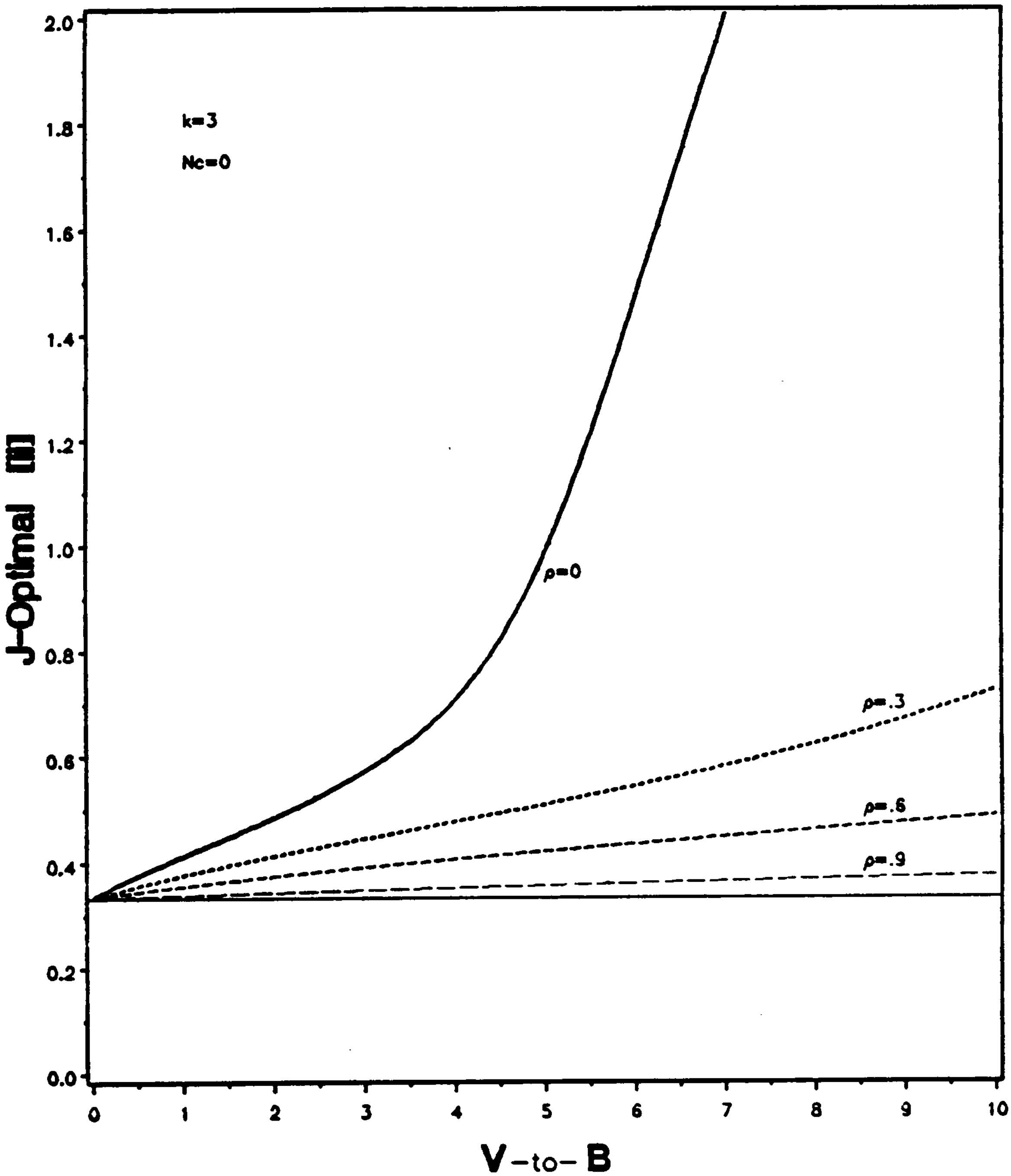


Figure H-5 J-optimal [ii] versus V/B for the Pure CR Strategy.

Region of interest is Cuboidal.
Two-level factorial design with $k = 3$ factors and $N_c = 0$ center runs.
Values of $\rho_* = 0, .3, .6, .9$ and the value of $\phi = k/2$.
The V-optimal value of the second moment is $[ii] = 1.0$.
The horizontal line at $[ii] = .33$ indicates the B-optimal value.

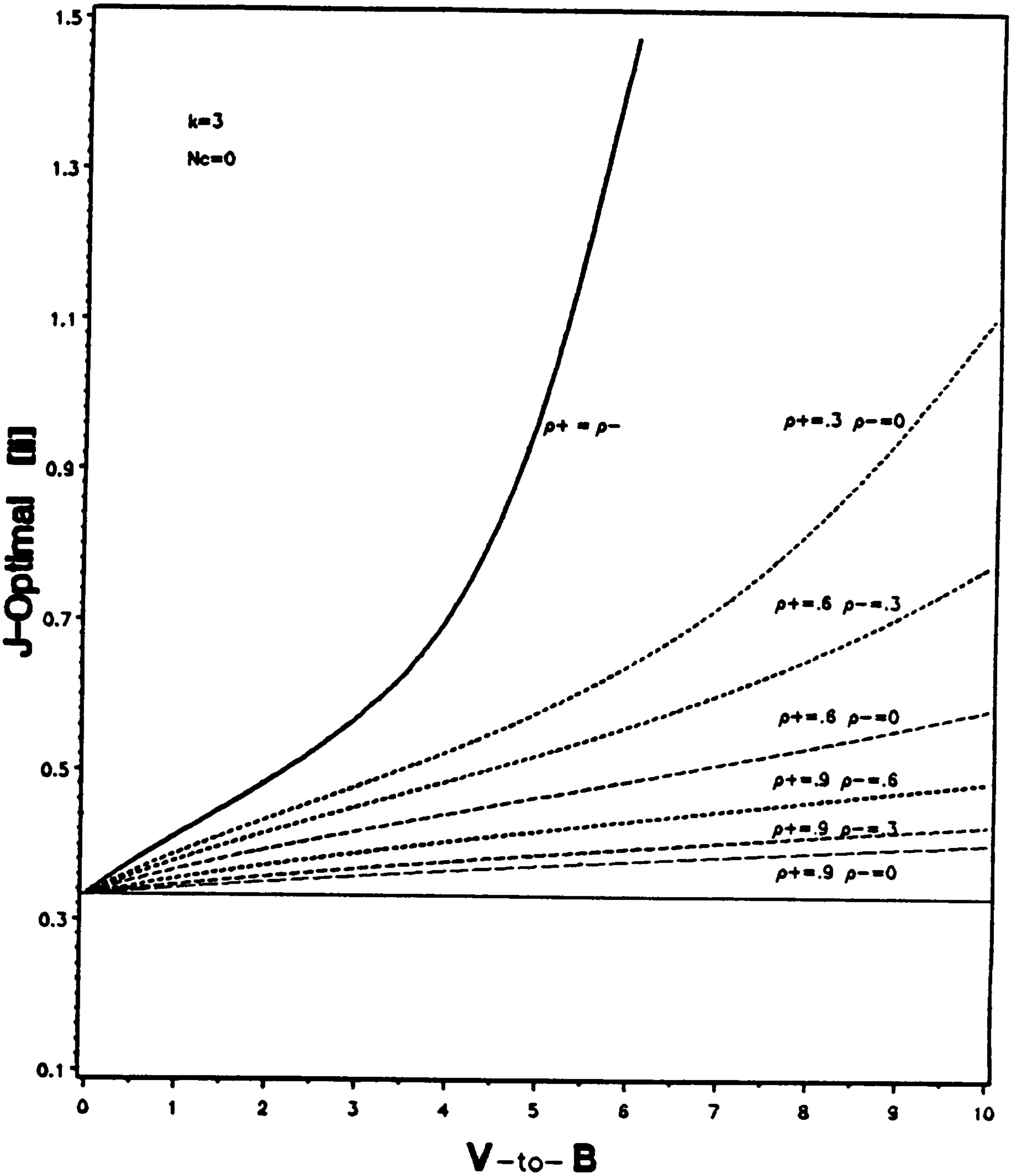


Figure H-6 J-optimal [ii] versus V/B for the Pure AR Strategy.

Region of interest is Cuboidal.
Two-level factorial design with $k = 3$ factors and $N_c = 0$ center runs.
Values of $\rho_- \leq \rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.
The V-optimal value of the second moment is $[ii] = 1.0$.
The horizontal line at $[ii] = .33$ indicates the B-optimal value.

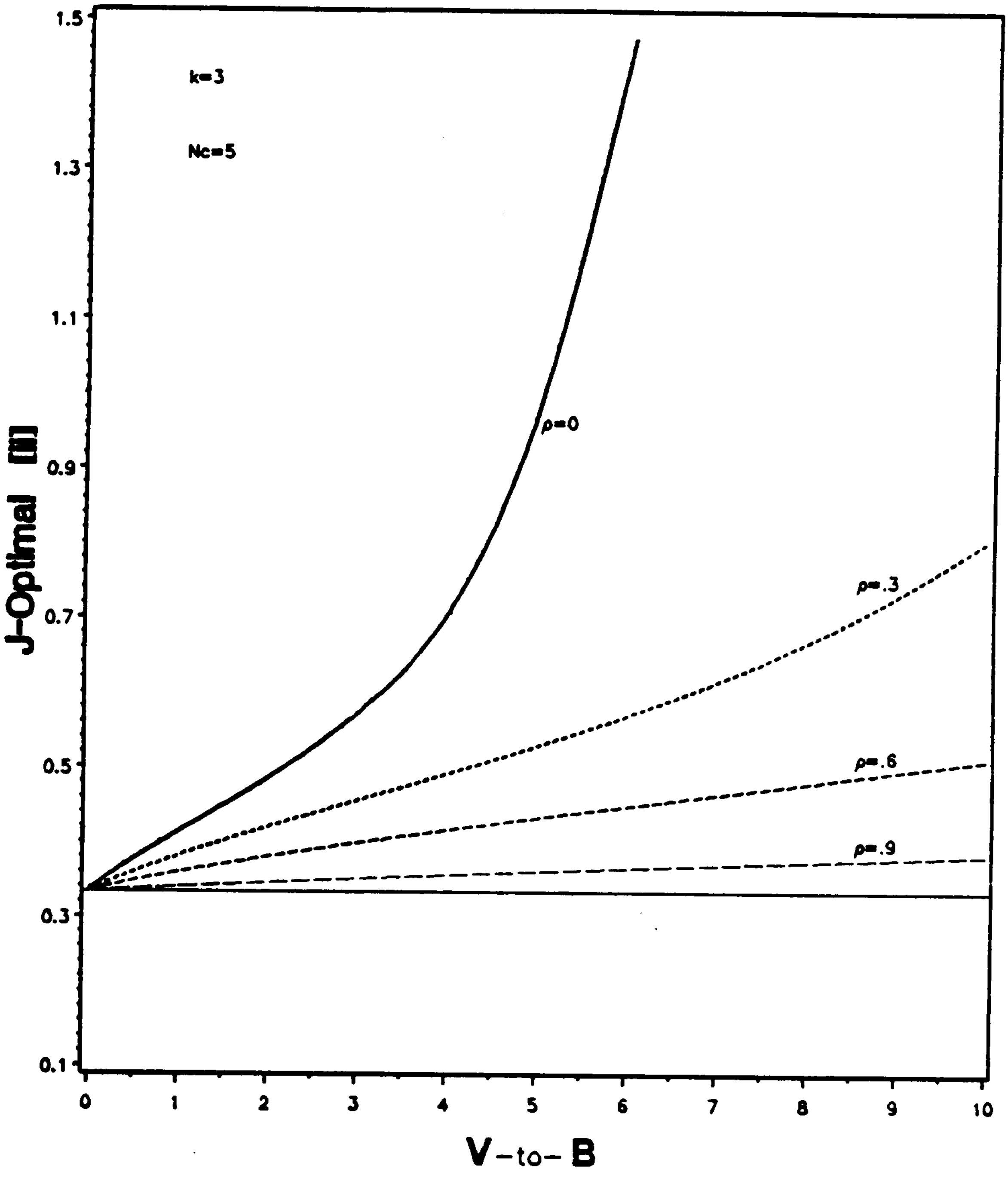


Figure H-7 **J-optimal [ii] versus V/B for the Modified CR Strategy under OLS.**
Region of interest is Cuboidal.
Modified CR Strategy under Ordinary Least Squares.
Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.
Values of $\rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.
The V-optimal value of the second moment is $[ii] = 1.0$.
The horizontal line at $[ii] = .33$ indicates the B-optimal value.

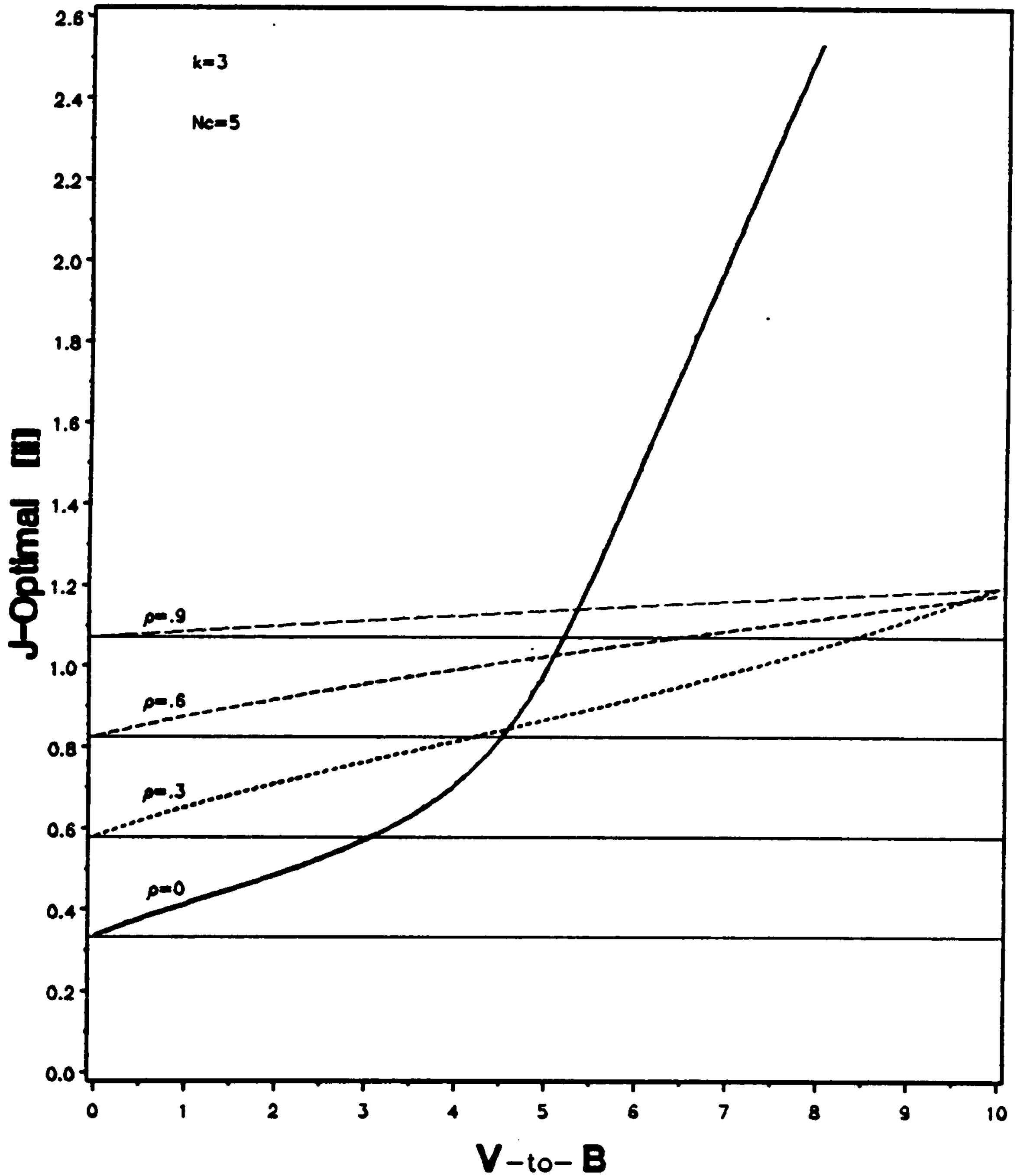


Figure H-8 J-optimal [ii] versus V/B for the Modified CR Strategy under WLS.

Region of interest is Cuboidal.

Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.

Values of $\rho_* = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The V-optimal value of the second moment is $[ii] = 1.0$.

The horizontal lines indicate the B-optimal values.

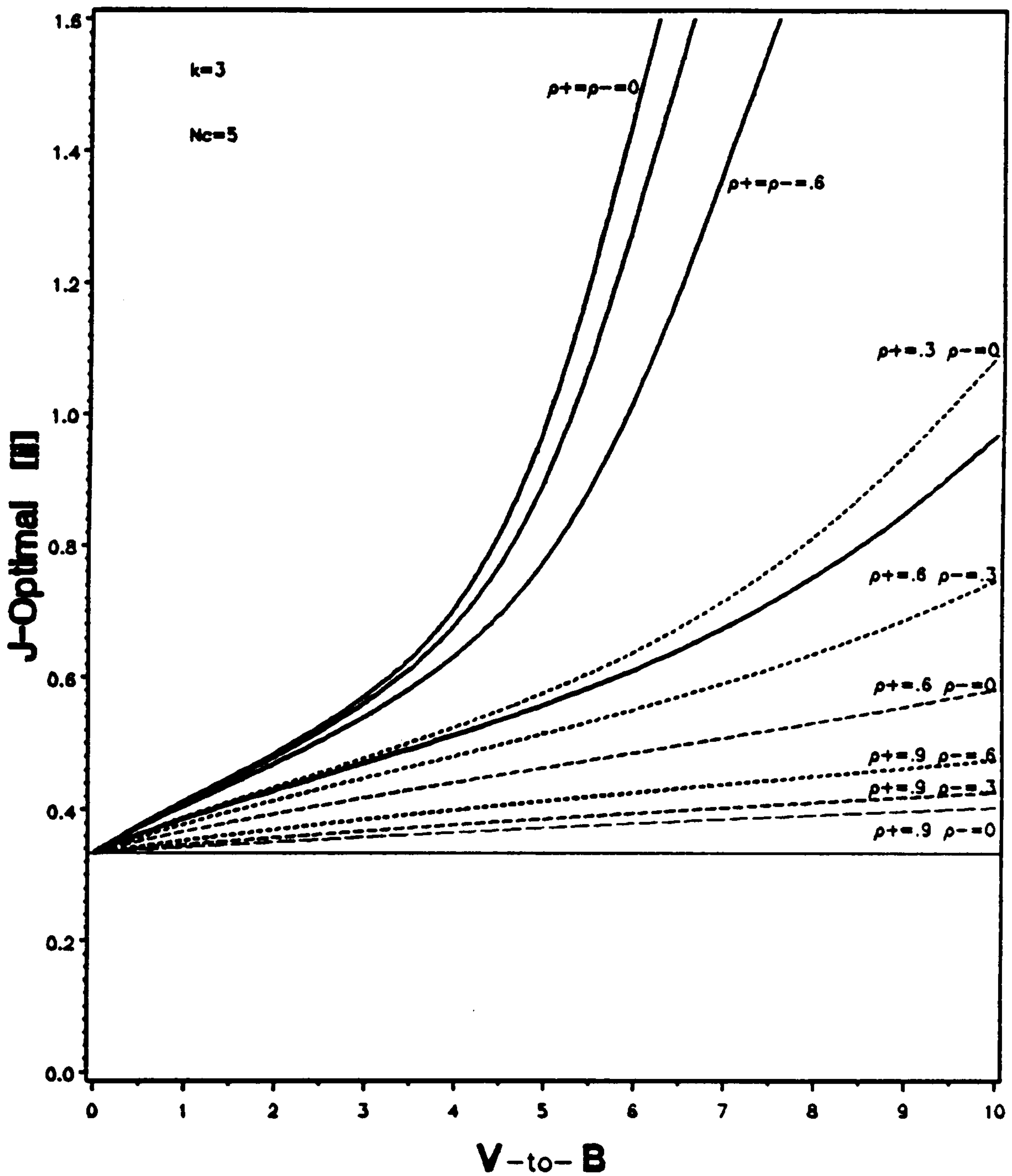


Figure H-9 **J-optimal [ii] versus V/B for the Modified AR Strategy under OLS.**
Region of interest is Cuboidal.
Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.
Values of $\rho_- \leq \rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.
The V-optimal value of the second moment is [ii] = 1.0.
The horizontal line at [ii] = .33 indicates the B-optimal value.

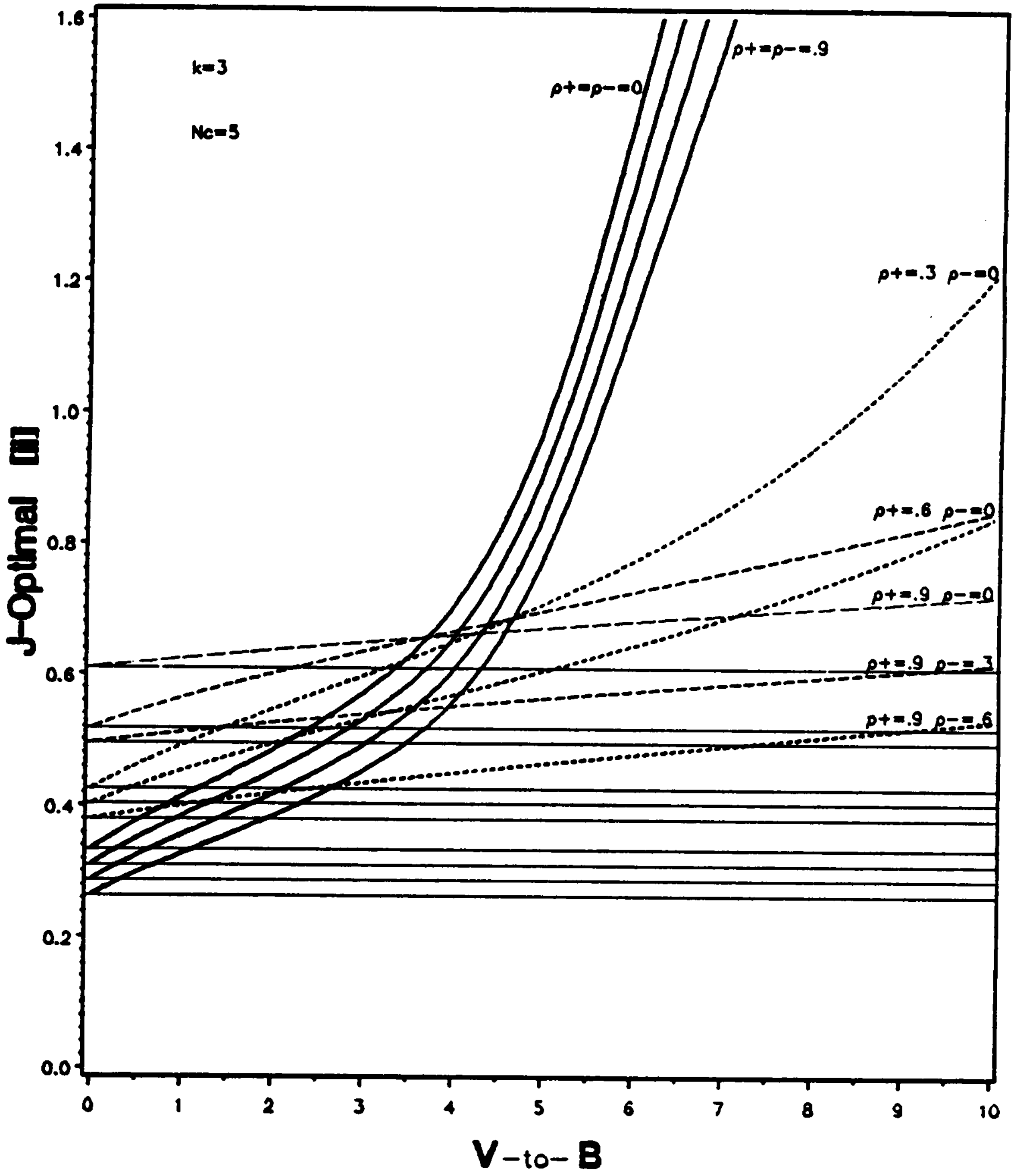


Figure H-10 J-optimal [ii] versus V/B for the Modified AR Strategy under WLS.

Region of interest is Cuboidal.

Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.

Values of $\rho_- \leq \rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The V-optimal value of the second moment is $[ii] = 1.0$.

The horizontal lines indicate the B-optimal values.

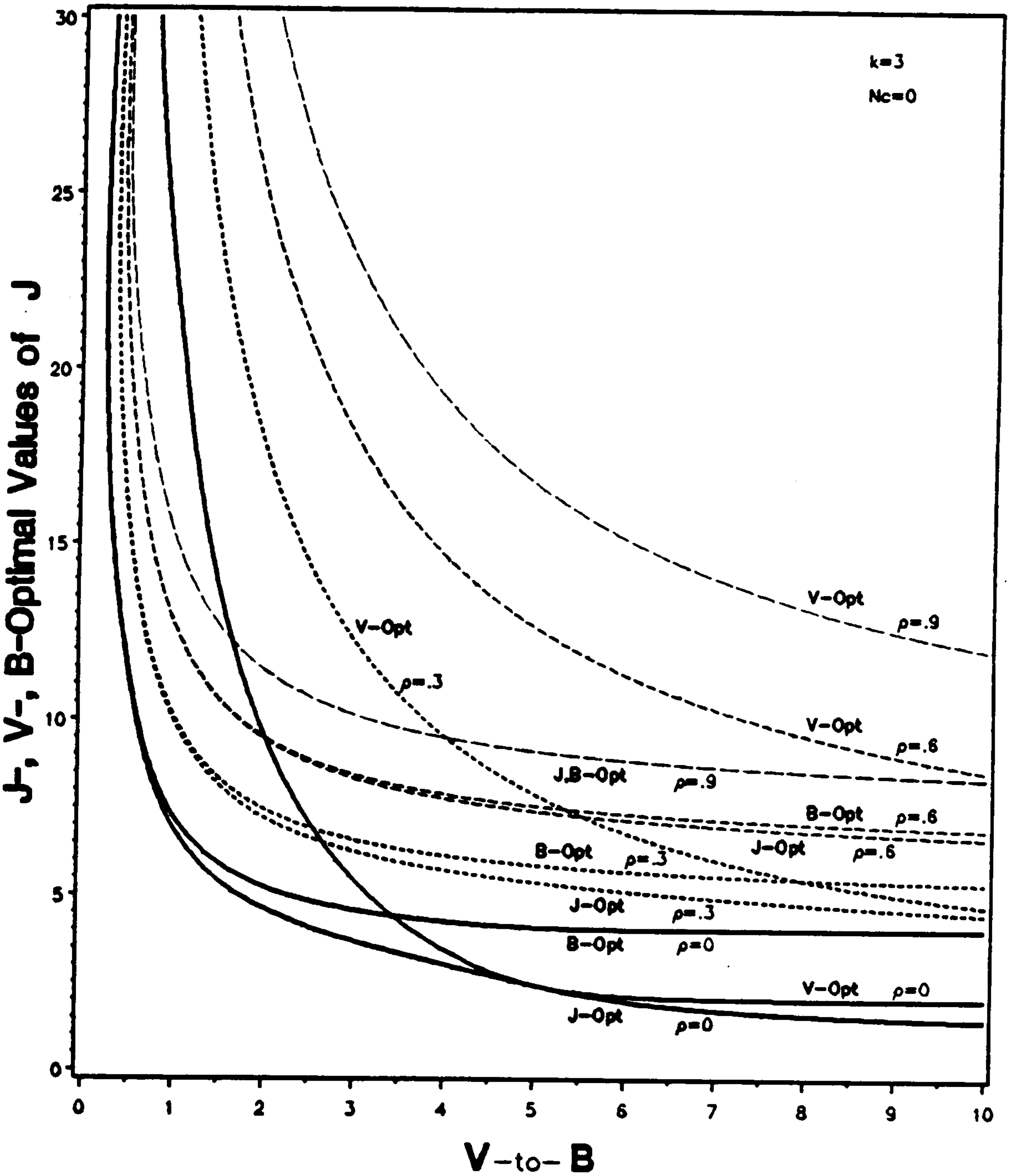


Figure H-11 J-, V-, B-optimal values of J versus V/B for the Pure CR Strategy.

Region of interest is Cuboidal.

Two-level factorial design with $k = 3$ factors and $N_c = 0$ center runs.

Values of $\rho_* = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

The B-optimal values of J are computed using $[ii] = w_{ii} = 1/(k+2)$.

The V-optimal values of J are computed using $[ii] = 1$.

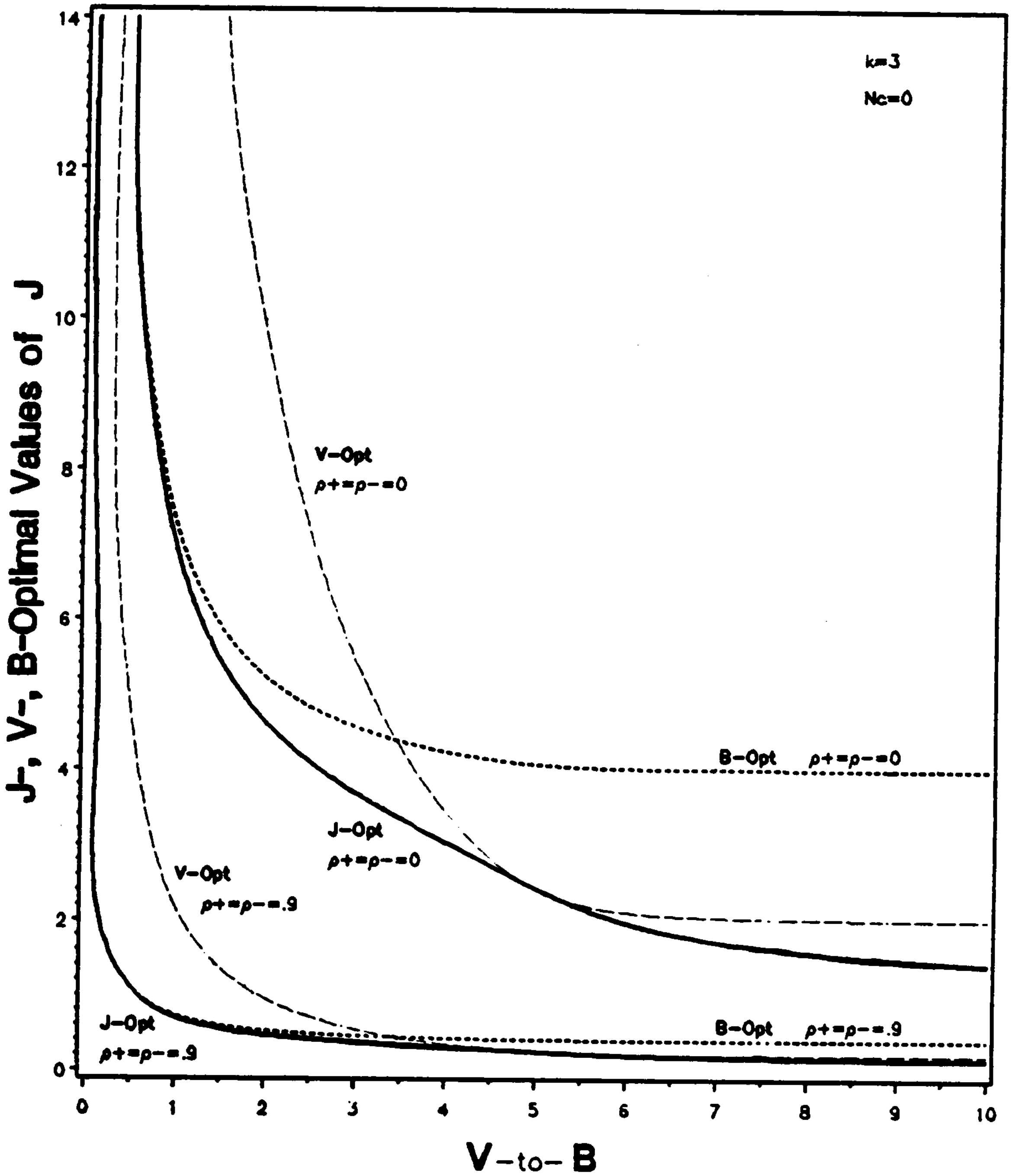


Figure H-12 **J-, V-, B-optimal values of J versus V/B for the Pure AR Strategy.**

Region of interest is Cuboidal.

Two-level factorial design with $k = 3$ factors and $N_c = 0$ center runs.

Value of $\rho_- = \rho_+ = .9$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

The B-optimal values of J are computed using $[ii] = w_{ii} = 1/(k+2)$.

The V-optimal values of J are computed using $[ii] = 1$.

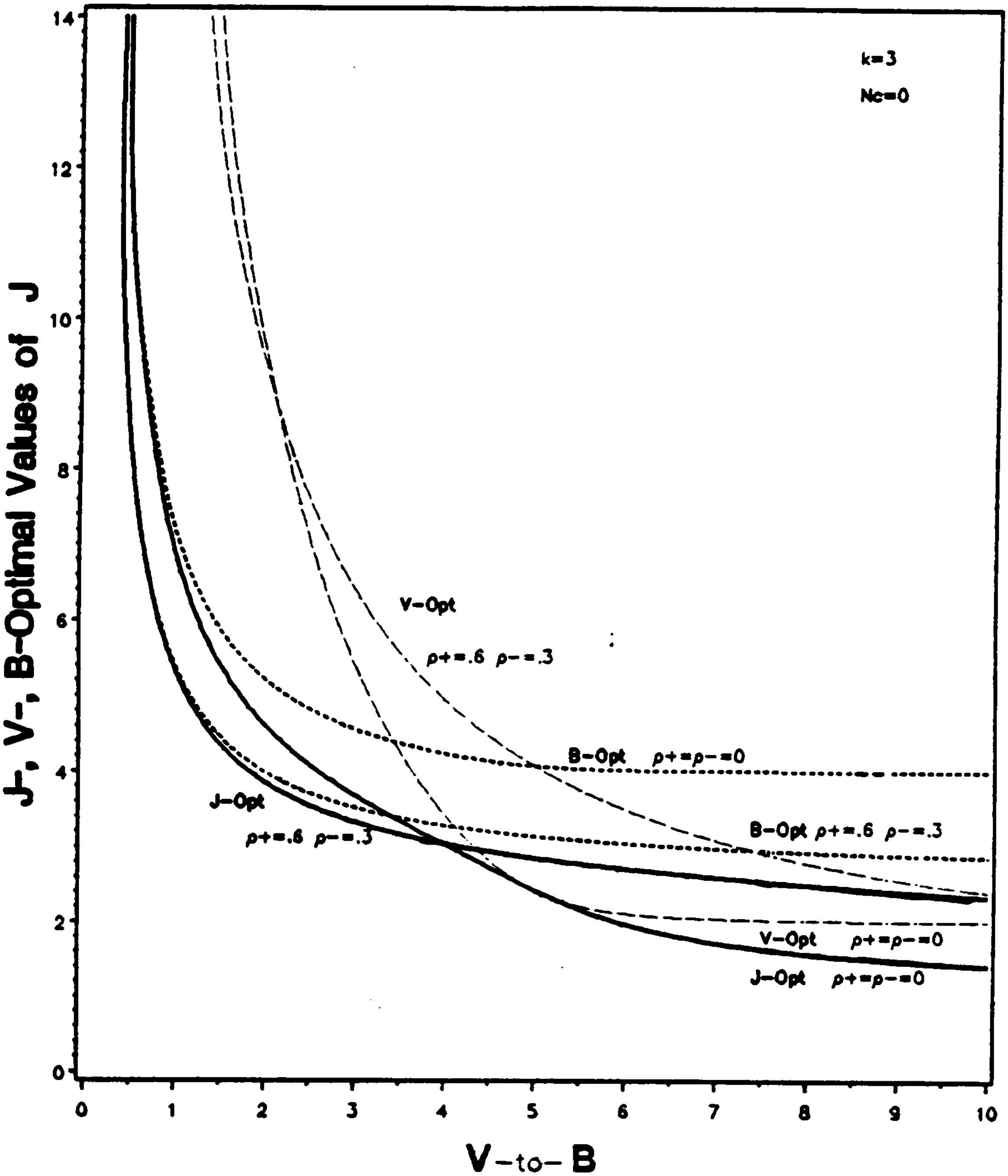


Figure H-13 J-, V-, B-optimal values of J versus V/B for the Pure AR Strategy.

Region of interest is Cuboidal.

Two-level factorial design with $k = 3$ factors and $N_c = 0$ center runs.

Value of $\rho_+ = .6$ and $\rho_- = .9$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

The B-optimal values of J are computed using $[ii] = w_{ii} = 1/(k+2)$.

The V-optimal values of J are computed using $[ii] = 1$.

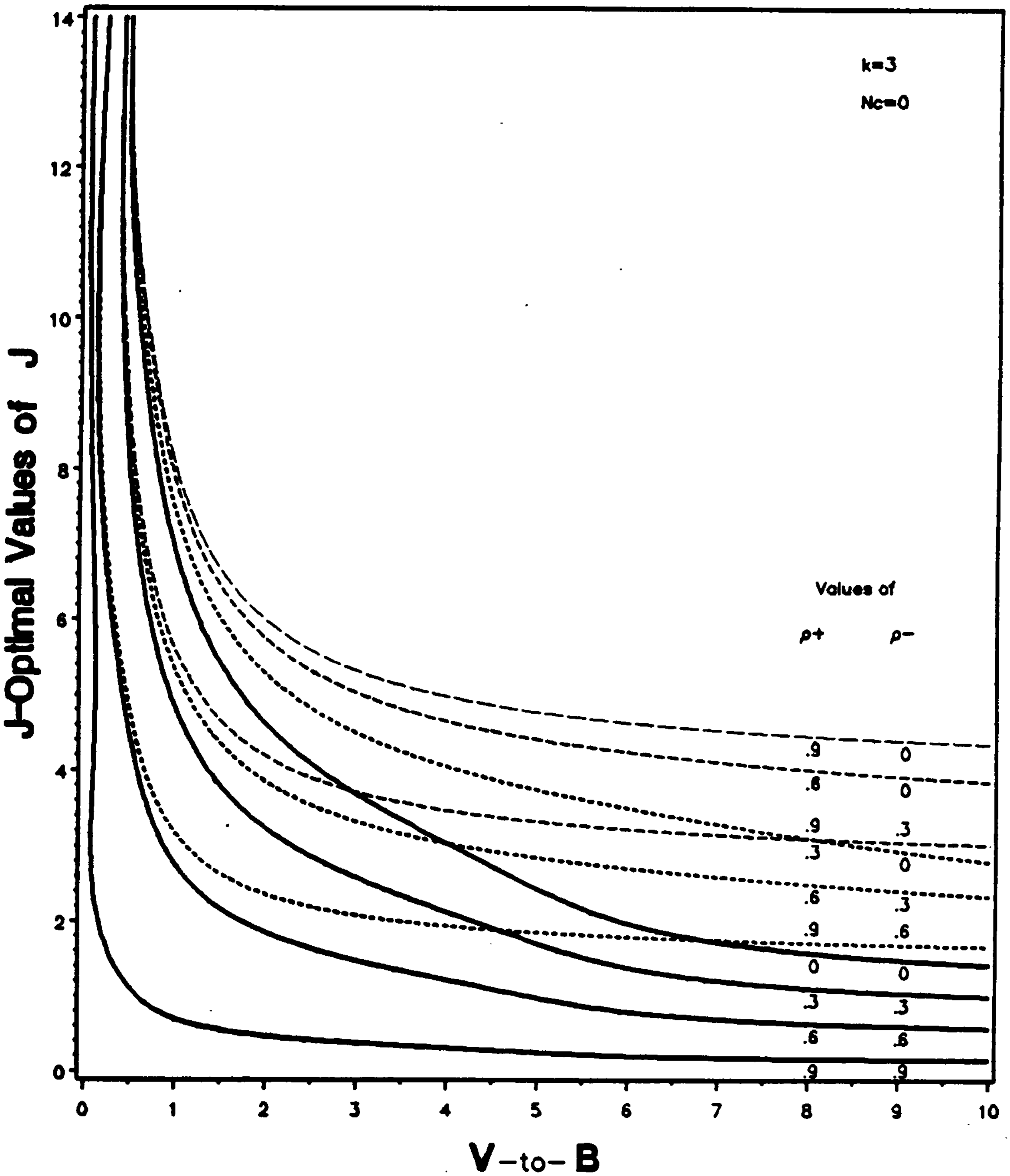


Figure H-14 J-optimal values of J versus V/B for the Pure AR Strategy.

Region of interest is Cuboidal.

Two-level factorial design with $k=3$ factors and $N_c = 0$ center runs.

Value of $\rho_- \le \rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.

The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

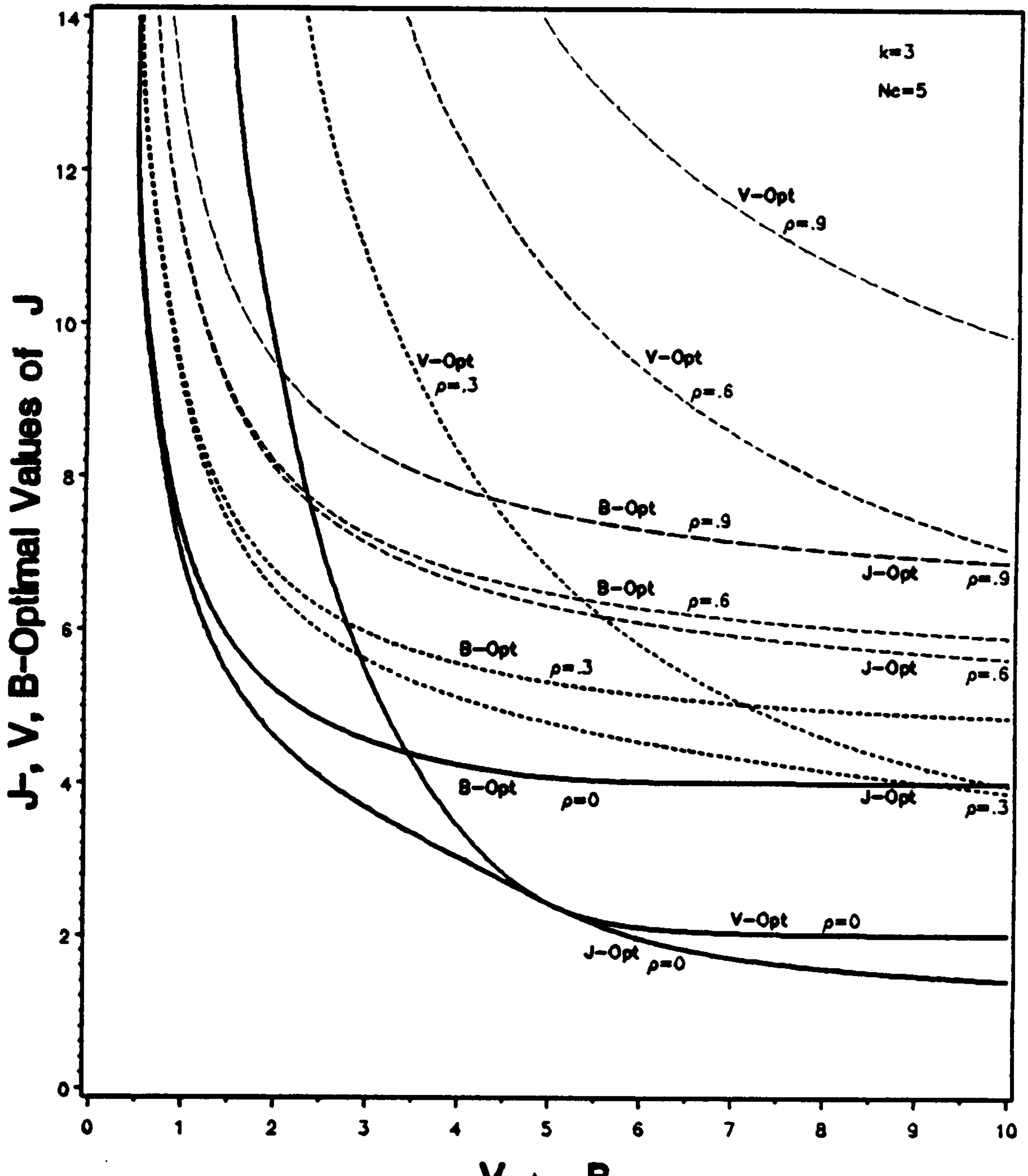


Figure H-15 J-, V-, B-optimal values of J versus V/B for the Modified CR Strategy under OLS.
Region of interest is Cuboidal.
Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.
Values of $\rho_{+} = 0, .3, .6, .9$ and the value of $\phi = k/2$.
The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.
The B-optimal values of J are computed using $[ii] = w_{ii} = 1/(k+2)$.
The V-optimal values of J are computed using $[ii] = 1$.

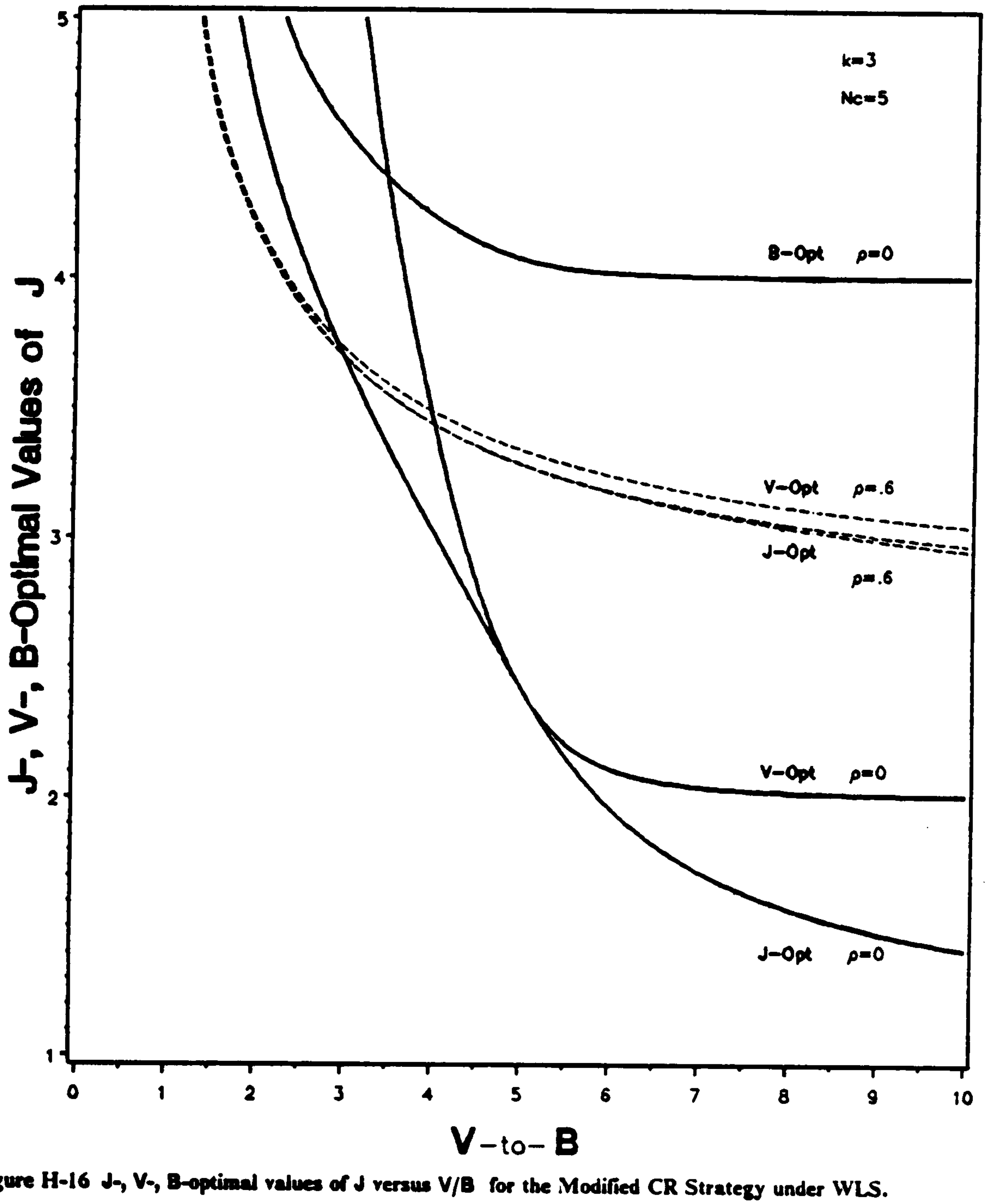


Figure H-16 J-, V-, B-optimal values of J versus V/B for the Modified CR Strategy under WLS.
Region of interest is Cuboidal.
Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.
Values of $\rho_* = 0, .6$ and the value of $\phi = k/2$.
The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.
The B-optimal values of J are computed using $[ii] = w_{ii}/z = 1/z\,(k+2)$.
The V-optimal values of J are computed using $[ii] = 1$.

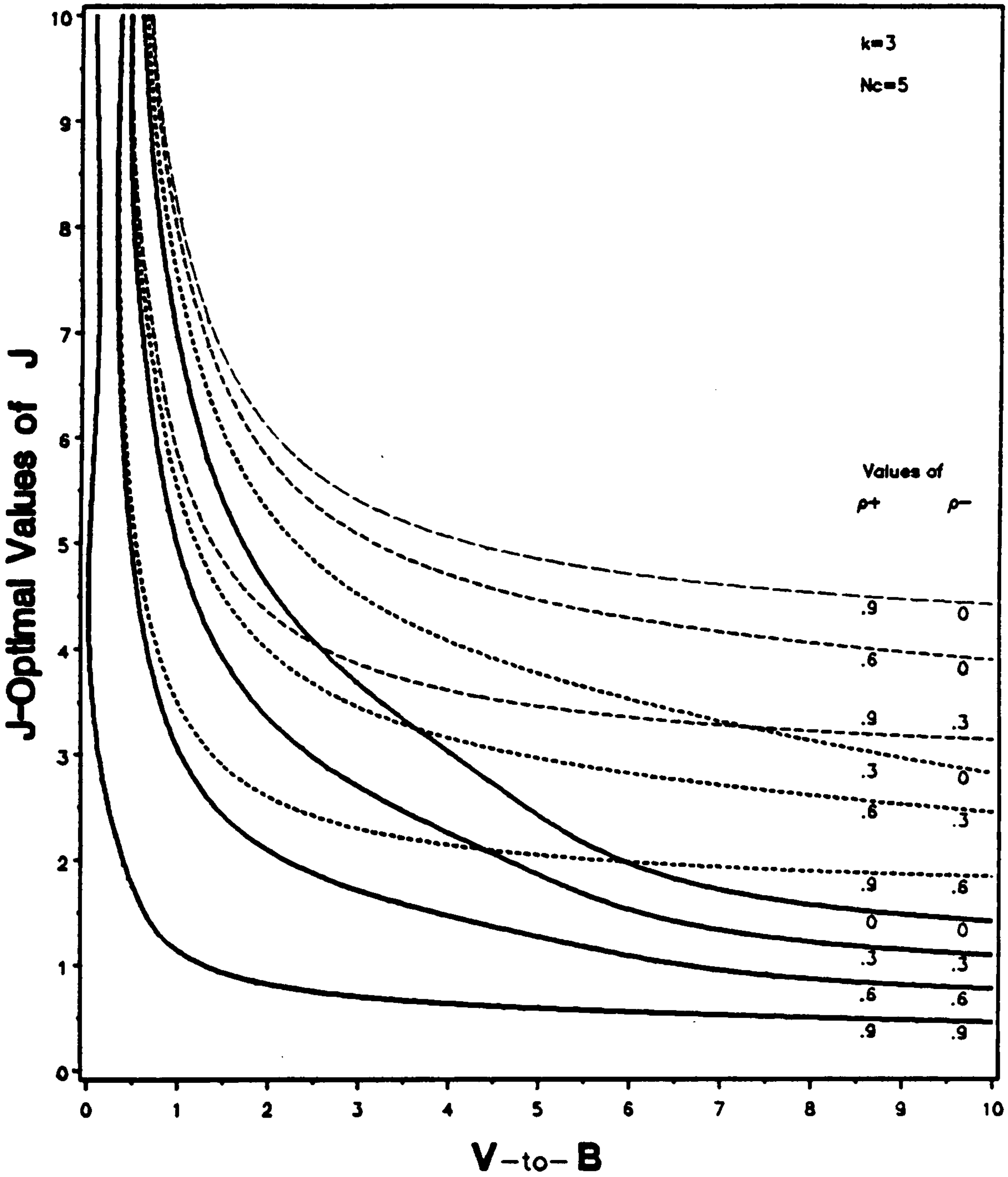


Figure H-17 J-optimal values of J versus V/B for the Modified AR Strategy under OLS.
Region of interest is Cuboidal.
Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.
Value of $\rho_- \le \rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.
The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

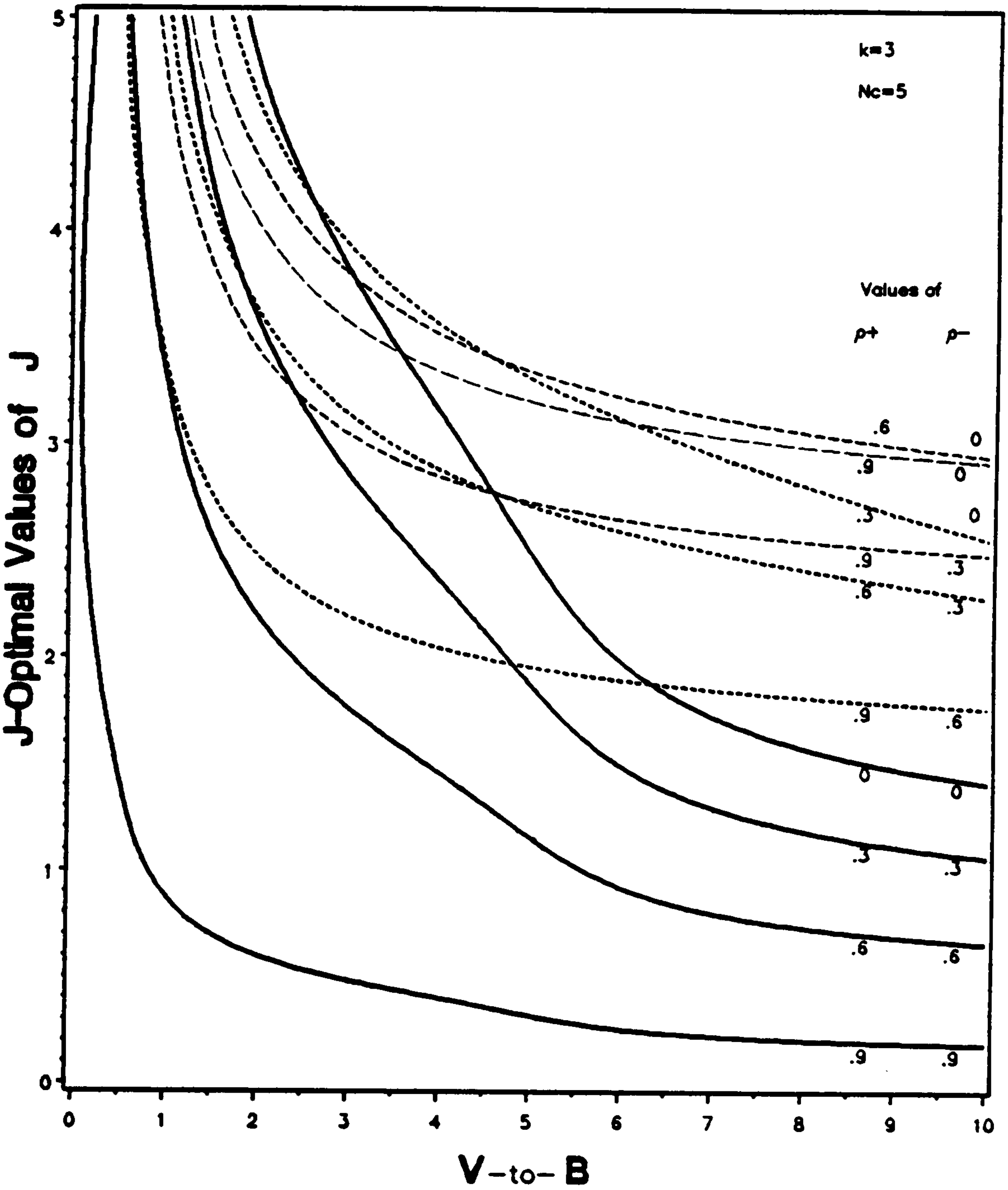


Figure H-18 J-optimal values of J versus V/B for the Modified AR Strategy under WLS.
Region of interest is Cuboidal.
Two-level factorial design with $k = 3$ factors and $N_c = 5$ center runs.
Value of $\rho_- \leq \rho_+ = 0, .3, .6, .9$ and the value of $\phi = k/2$.
The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

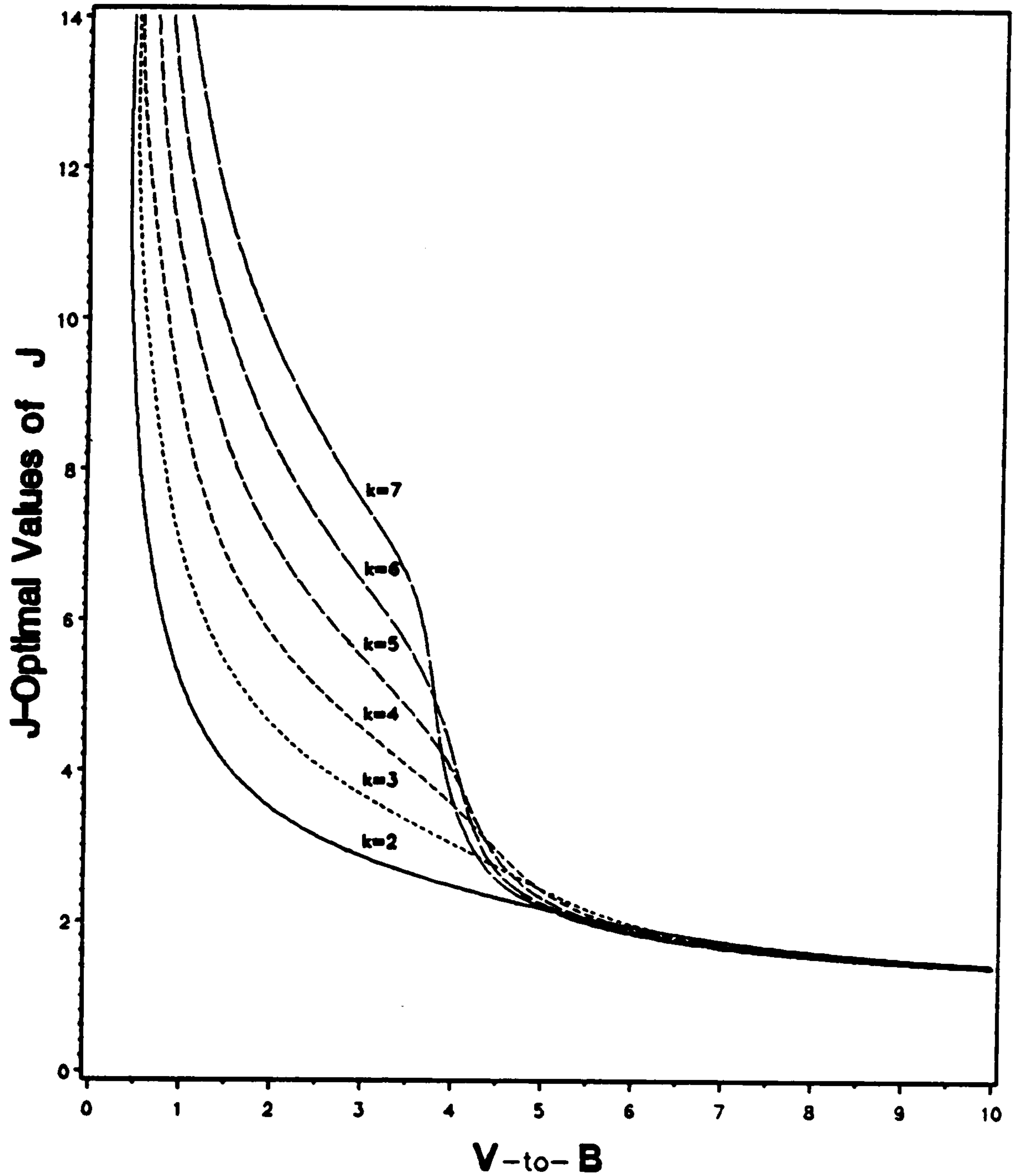


Figure H-19 J-optimal values of J versus V/B for various values of k.

IR Strategy in a Cuboidal region.
Two-level factorial design with $k = 2, \ldots, 7$ factors.
Value of $N_c = 0$ and the value of $\phi = k/2$.
The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

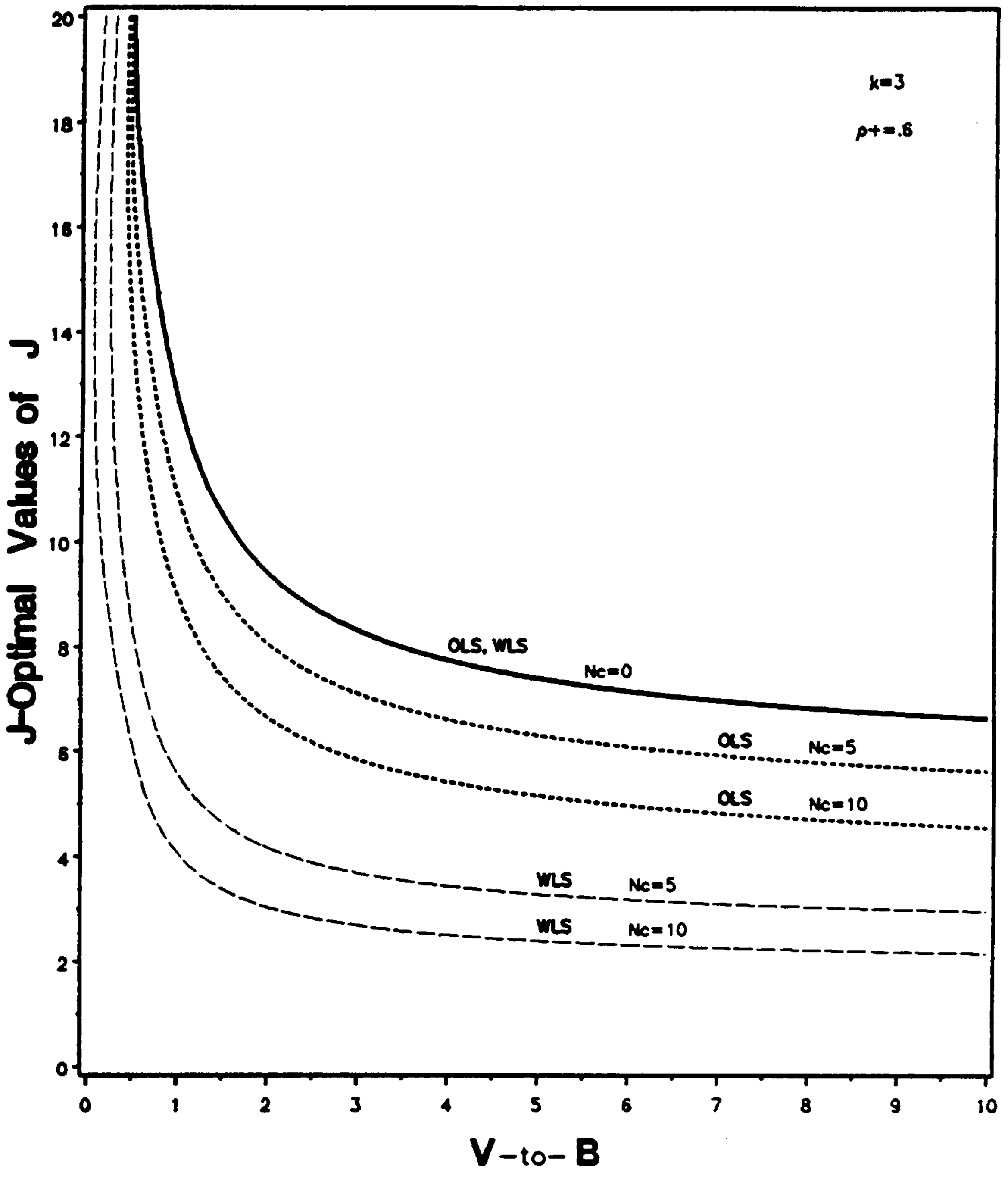


Figure H-20 J-optimal values of J versus V/B for various values of N_c.

CR Strategy with $\rho_+ = .6$ in a Cuboidal region.
Least squares estimation techniques are OLS and WLS.
Two-level factorial design with $k = 3$ factors.
Values of $N_c = 0, 5, 10$ and the value of $\phi = k/2$.
The J-optimal values of J utilize the [ii] which minimize J for a given V/B ratio.

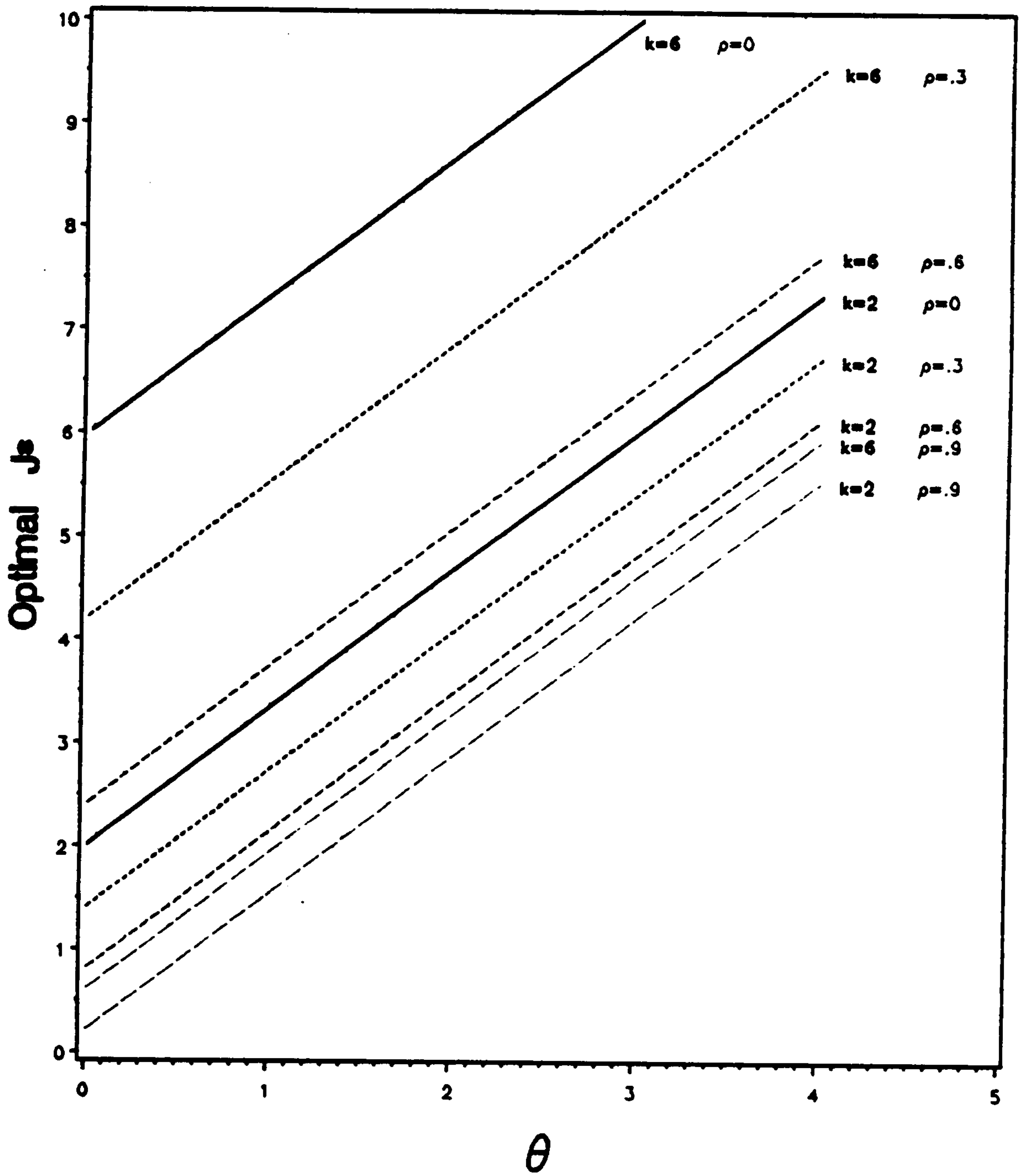


Figure H-21 First order optimal values of J^* versus θ.

Region of interest is Cuboidal.

Two-level factorial design with k = 2 and k = 6 factors.

Values of $\rho_+ = 0, .3, .6, .9$.

The optimal values of J are computed using $[ii] = 1$.

Appendix I

Computer Programs for Determination of J and J*

in the First Order Case

```
//A0392JMD JOB 33F02,JMD,REGION=1536K
//*
//*  This program computes the J-optimal values of [ii], J, B, and V,
//*            as well as the B-optimal values of [ii], J, B, and V,
//*                   and the V-optimal values of J, B, and V.
//*
//*  This program uses a FORTRAN subroutine, which calls an
//*       IMSL subroutine MYPOLY from within SAS PROC MATRIX.
//*
/*PRIORITY IDLE
/*JOBPARM LINES=1,CARDS=100,ACCTPG
//STEP1 EXEC VSF2C
//FORT.SYSIN DD *
C---- INITIALIZATIONS:
      INTEGER FUNCTION MATSUB(NARG,ARGS)
      INTEGER*4 NARG, ARGS(1)
      INTEGER*4 MIN, MAX, ROW, COL, ILOC, OLOC, IER
      REAL*8 IBASE(1), OBASE(1)
C---- ENSURE THAT ONLY ONE ARGUMENT IS PASSED TO THIS PROCEDURE:
      IF(NARG.NE.1) THEN
      MATSUB = 5
         RETURN
      ENDIF
C---- ENSURE THAT THE ONE ARGUEMENT IS A 1 BY NDEG+1 OR NDEG+1 VECTOR
C     WHERE 0 , NDEG , 101
      CALL ARG(ARGS(1),ROW,COL,ILOC,IBASE)
      MIN = MIN0(ROW,COL)
      MAX = MAX0(ROW,COL)
      IF(MIN.NE.1.OR.MAX.LE.1.OR.MAX.GE.102) THEN
         MATSUB = 6
         RETURN
```

```
      ENDIF
C---- DEFINE THE MATRIX IN WHICH THE ZEROS OF THE POLYNOMIAL WILL BE
C     PLACED
      CALL SETUP(1,MAX-1,2)
      CALL ARG(1,ROW,COL,OLOC,OBASE)
      IF(ROW.EQ.0.OR.COL.EQ.0) THEN
         MATSUB = 1
         RETURN
      ENDIF
C---- FIND THE ZEROS OF A POLYNOMIAL USING THE IMSL ROUTINE ZRPOLY
      CALL ZRPOLY(IBASE(ILOC),MAX-1,OBASE(OLOC),IER)
      MATSUB = IER
      RETURN
      END
/*
//STEP2 EXEC PGM=IEWL,PARM='MAP,LIST'
//SYSPRINT   DD SYSOUT=A
//SYSUT1     DD UNIT=SYSDA,SPACE=(TRK,(40,40))
//SYSLIB     DD DSN=SYS2.SAS.SUBLIB,DISP=SHR
//           DD DSN=SYS2.SAS.LIBRARY,DISP=SHR
//           DD DSN=SYS2.SAS.LIBRARY,DISP=SHR
//           DD DSN=SYS2.PLIBASE,DISP=SHR
//           DD DSN=SYS2.VSF2FORT,DISP=SHR   R3.VFORTLIB,DISP=SHR
//           DD DSN=VPI.IMSLDP,DISP=SHR IMSL77.DP,DISP=SHR
//           DD DSN=VPI.UTILITY.NEW,DISP=SHR
//SYSLIN     DD DSN=&&LOADSET,DISP=(OLD,DELETE,DELETE)
//           DD *
 INCLUDE SYSLIB(MATMAIN)
 ENTRY MATMAIN
 SETSSI BF110000
 NAME MYPOLY(R)
/*
//SYSLMOD DD DSN=&LIBRARY,DISP=(NEW,PASS,DELETE),UNIT=SYSDA,
//           SPACE=(CYL,(10,20,20),,CONTIG)
//STEP1 EXEC SAS
//SYSIN DD *
OPTIONS NOCENTER NODATE NONUMBER LS=120;
TITLE 'Correlation Induction Schemes as a Function of VB Ratio';
PROC MATRIX FUZZ FW=5;

*Input Desired Output:  1=[ii]  2=[ii],J  3=J  4=J,J-Bias,J-Var  5=All;
OUTPUT=3;
*Input Region Shape: 1=Cube 2=Sphere;
REGION=2;
*Input Least Squares Technique: 1=OLS  2=WLS;
LST=1;
*Input Correlation Strategy: 1=IR   2=CR   3=AR;
STR=3;
*Input Number of Center Runs;
NC=5;
*Input Number of factors, k;
KDATA=(1 3 5);
*Input Values of rho plus;
RPDATA=(0 .3 .6 .9);
*Input Values of Variance to Bias Ratio;
VBDATA=(.01 .05 .1 .5 .75 1
         1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 8 8.5 9 9.5 10);

* IR and CR Strategies;
IF STR<3 THEN RMDATA=(0);
* AR Strategy;
IF STR=3 THEN RMDATA=RPDATA;

NK=NCOL(KDATA);
DO KNO=1 TO NK BY 1;
K=KDATA(1,KNO);
PHI=K#/2;
```

```
*Cuboidal Region;
IF REGION=1 THEN W=1#/3;
IF REGION=1 THEN Y=(4+5#K-3#PHI)#/(15#K+30);
*Spherical Region;
IF REGION=2 THEN W=1#/(K+2);
IF REGION=2 THEN Y=(K+2-PHI)#/((K+2)#(K+4));

IF K>0 THEN FR=1;
IF K>4 THEN FR=.5;
IF K>6 THEN FR=.25;
N=FR#(2##K)+NC;
IF NC>0 THEN N2C=NC-1;
IF NC=0 THEN N2C=0;
N1C=N-N2C;
IF NC>1 THEN N2A=NC-2;
IF NC<=1 THEN N2A=NC;
N1A=N-N2A;

NVB=NCOL(VBDATA);
NRP=NCOL(RPDATA);
NRM=NCOL(RMDATA);
NRHO=NRP#(NRP+1)#/2;
DATAI=J(NVB+2,NRHO+1,0);
DATAW=J(NVB+2,NRHO+1,0);
DATAJ=J(NVB+2,NRHO+1,0);
DATAB=J(NVB+2,NRHO+1,0);
DATAV=J(NVB+2,NRHO+1,0);

DO VBNO=1 TO NVB BY 1;
VB=VBDATA(1,VBNO);
COUNT=0;

DO RPNO=1 TO NRP BY 1;
RP=RPDATA(1,RPNO);

DO RMNO=1 TO NRM BY 1 WHILE(RMNO<=RPNO);
RM=RMDATA(1,RMNO);

*Ir&OLS; IF LST=1            THEN Z=1;
*Cr&WLS; IF LST=2 AND STR=2 THEN Z=N#/(N-N2C#RP+N1C#N2C#RP);
*Ar&WLS; IF LST=2 AND STR=3 THEN Z=N#/(N-N2A#RP+.5#N1A#N2A#(RP-RM));

IIV=1;
IIB=W#/Z;

*1 IR;  IF             STR=1 THEN V=1;
*2 CrO; IF LST=1 AND STR=2 THEN V=1-((N1C#RP)#/N)+(((N1C#N1C)#RP)#/N);
*3 ArO; IF LST=1 AND STR=3 THEN V=1-((N1A#RP)#/N)
                                  +(((N1A#N1A)#(RP-RM))#/(2#N));
*4 CrW; IF LST=2 AND STR=2 THEN V=(N-N#RP+N#N1C#RP)#/
                                  (N-N2C#RP+N1C#N2C#RP);
*5 ArW; IF LST=2 AND STR=3 THEN V=(N-N#RP+.5#N#N1A#(RP-RM))#/
                                  (N-N2A#RP+.5#N1A#N2A#(RP-RM));
*6 CrP; IF N2C=0 AND STR=2 THEN V=1-RP+N#RP;
*7 ArP; IF N2A=0 AND STR=3 THEN V=1-RP+.5#N#(RP-RM);

A=((W#K#(1-RP)#(2-VB))#/(2#V))-(W#/Z);
B=(W#W#K#(1-RP)#(VB-1))#/(Z#V);
C=-(W#W#K#(1-RP)#(W#PHI+2#Y)#VB)#/(2#PHI#Z#Z#V);
S=B-(A#A#/3);                S3=S#/3;
T=(2#A#A#A#/27)-(A#B#/3)+C;  T2=T#/2;
U=(T2#T2)+(S3#S3#S3);
ONE3=1#/3;

IF U GT 0 THEN DO;
PIN=-T2+SQRT(U);
```

```
IF PIN LT 0 THEN P=-((ABS(PIN))##ONE3);
IF PIN GE 0 THEN P=(PIN##ONE3);
QIN=T2+SQRT(U);
IF QIN LE 0 THEN Q=(ABS(QIN))##ONE3;
IF QIN GT 0 THEN Q=-(QIN##ONE3);
II=-(A#/3)+P+Q;
END;

IF U LE 0 THEN DO;
CC=J(1,4,0);
CC(1,1)=1;
CC(1,2)=A;
CC(1,3)=B;
CC(1,4)=C;
ROOTS=MYPOLY(CC);
REAL=ROOTS(,1);
II=REAL(<>,1);
END;

COUNT=COUNT+1;
DATAI(1,COUNT+1)=RP;
DATAW(1,COUNT+1)=RP;
DATAJ(1,COUNT+1)=RP;
DATAB(1,COUNT+1)=RP;
DATAV(1,COUNT+1)=RP;
DATAI(2,COUNT+1)=RM;
DATAW(2,COUNT+1)=RM;
DATAJ(2,COUNT+1)=RM;
DATAB(2,COUNT+1)=RM;
DATAV(2,COUNT+1)=RM;
DATAI(VBNO+2,COUNT+1)=II;
DATAW(VBNO+2,COUNT+1)=IIB;

VMSE=V+((W#K#(1-RP))#/II);
THETA=(V+W#K#(1-RP)#/II)#/(VB#((PHI#((Z#II-W)##2))+(2#Y#W)));
BMSE=THETA#((PHI#((Z#II-W)##2))+(2#Y#W));
JMSE=VMSE+BMSE;
DATAJ(VBNO+2,COUNT+1)=JMSE;

BBIAS=THETA#((PHI#((Z#IIB-W)##2))+(2#Y#W));
VBIAS=V+((W#K#(1-RP))#/IIB);
JBIAS=VBIAS+BBIAS;
DATAB(VBNO+2,COUNT+1)=JBIAS;

VVAR=V+((W#K#(1-RP))#/IIV);
BVAR=THETA#((PHI#((Z#IIV-W)##2))+(2#Y#W));
JVAR=VVAR+BVAR;
DATAV(VBNO+2,COUNT+1)=JVAR;

END;
END;
DATAI(VBNO+2,1)=VB;
DATAJ(VBNO+2,1)=VB;
DATAB(VBNO+2,1)=VB;
DATAV(VBNO+2,1)=VB;
END;

IF STR=1 THEN DATAI=DATAI(,1:2);
IF STR=1 THEN DATAJ=DATAJ(,1:2);
IF STR=1 THEN DATAB=DATAB(,1:2);
IF STR=1 THEN DATAV=DATAV(,1:2);
IF STR=2 THEN DATAI=DATAI(,1:NRP+1);
IF STR=2 THEN DATAW=DATAW(,1:NRP+1);
IF STR=2 THEN DATAJ=DATAJ(,1:NRP+1);
IF STR=2 THEN DATAB=DATAB(,1:NRP+1);
IF STR=2 THEN DATAV=DATAV(,1:NRP+1);
DATAW=DATAW(1:3,);
```

```
C1='K' 'NC' 'PHI' 'Region' 'LSTech' 'CIStrat' 'N2C' 'N2A';

IF STR=1 THEN C2='V/B' 'II';
IF STR=2 THEN C2='V/B' 'IRP0' 'IRP3' 'IRP6' 'IRP9';
IF STR=3 THEN C2='V/B' 'IRP00' 'IRP30' 'IRP33' 'IRP60' 'IRP63'
                       'IRP66' 'IRP90' 'IRP93' 'IRP96' 'IRP99';
IF STR=1 THEN C3='V/B' 'JMSE';
IF STR=2 THEN C3='V/B' 'JRP0' 'JRP3' 'JRP6' 'JRP9';
IF STR=3 THEN C3='V/B' 'JRP00' 'JRP30' 'JRP33' 'JRP60' 'JRP63'
                       'JRP66' 'JRP90' 'JRP93' 'JRP96' 'JRP99';
IF STR=1 THEN C4='V/B' 'BMSE';
IF STR=2 THEN C4='V/B' 'BRP0' 'BRP3' 'BRP6' 'BRP9';
IF STR=3 THEN C4='V/B' 'BRP00' 'BRP30' 'BRP33' 'BRP60' 'BRP63'
                       'BRP66' 'BRP90' 'BRP93' 'BRP96' 'BRP99';
IF STR=1 THEN C5='V/B' 'VMSE';
IF STR=2 THEN C5='V/B' 'VRP0' 'VRP3' 'VRP6' 'VRP9';
IF STR=3 THEN C5='V/B' 'VRP00' 'VRP30' 'VRP33' 'VRP60' 'VRP63'
                       'VRP66' 'VRP90' 'VRP93' 'VRP96' 'VRP99';
R1='RP' 'RM'
   '1' '2' '3' '4' '5' '6' '7' '8' '9' '10' '11' '12' '13' '14'
   '15' '16' '17' '18' '19' '20' '21' '22' '23' '24' '25' '26' '27';

DATAK=K||NC||PHI||REGION||LST||STR||N2C||N2A;
PRINT DATAK COLNAME=C1 ROWNAME=R1;

IF OUTPUT NE 3 AND OUTPUT NE 4 THEN DO;
PRINT DATAI COLNAME=C2 ROWNAME=R1;
END;
IF OUTPUT NE 3 AND OUTPUT NE 4 AND LST=2 AND NC>0 THEN DO;
PRINT DATAW COLNAME=C2 ROWNAME=R1;
END;
IF OUTPUT NE 1 THEN DO;
PRINT DATAJ COLNAME=C3 ROWNAME=R1;
END;
IF OUTPUT=4 OR OUTPUT=5 THEN DO;
PRINT DATAB COLNAME=C4 ROWNAME=R1;
PRINT DATAV COLNAME=C5 ROWNAME=R1;
END;

END;
/*
//
```

```
//A0392JMD JOB 33F02,SAS,CLASS=Q
/*PRIORITY IDLE

*  This program computes the J*-optimal values of J*
*               as a function of  Theta  OR  V*/B* ;

OPTIONS NOCENTER NODATE NONUMBER LS=120;
TITLE 'Optimal Values of J-Star';
PROC MATRIX FUZZ FW=5;
II=1;

*Input Desired Output:  1=J* vs. theta     2=J* vs. V*/B* ;
OUTPUT=2;
*Input Region Shape: 1=Cube 2=Sphere;
REGION=2;
*Input Number of factors, k;
KDATA=(2 6);
*Input Values of rho plus;
RPDATA=(0 .3 .6 .9);
*Input Values of Theta;
TDATA=(.01 .05 .1 .5 .75 1
          1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 8 8.5 9 9.5 10);
*Input Values of Variance to Bias Ratio;
VBDATA=(.01 .05 .1 .5 .75 1
          1.5 2 2.5 3 3.5 4 4.5 5 5.5 6 6.5 7 7.5 8 8.5 9 9.5 10);

NK=NCOL(KDATA);
DO KNO=1 TO NK BY 1;
K=KDATA(1,KNO);

*Cuboidal Region;
IF REGION=1 THEN W=1#/3;
*Spherical Region;
IF REGION=2 THEN W=1#/(K+2);

NTHE=NCOL(TDATA);
NVB=NCOL(VBDATA);
NRP=NCOL(RPDATA);
DATAJVB=J(NVB+1,NRP+1,0);
DATAJT=J(NTHE+1,NRP+1,0);

*If Inputting Values of THETA;
IF OUTPUT=1 THEN DO;

DO THENO=1 TO NTHE BY 1;
THETA=TDATA(1,THENO);
COUNTT=0;

DO RPNO=1 TO NRP BY 1;
RP=RPDATA(1,RPNO);

COUNTT=COUNTT+1;
DATAJT(1,COUNTT+1)=RP;

VSTAR=(K#(1-RP))#/II;
BSTAR=4#W#THETA;
JSTAR=VSTAR+BSTAR;
DATAJT(THENO+1,COUNTT+1)=JSTAR;

END;  *RP;
DATAJT(THENO+1,1)=THETA;
END;  *THE;
END;  *IF THE;

*If Inputting VB Ratios;
IF OUTPUT=2 THEN DO;
```

```
DO VBNO=1 TO NVB BY 1;
VB=VBDATA(1,VBNO);
COUNT=0;

DO RPNO=1 TO NRP BY 1;
RP=RPDATA(1,RPNO);

COUNT=COUNT+1;
DATAJVB(1,COUNT+1)=RP;

VSTAR=(K#(1-RP))#/II;
BSTAR=VSTAR#/VB;
JSTAR=VSTAR+BSTAR;
DATAJVB(VBNO+1,COUNT+1)=JSTAR;

END;  *RP;
DATAJVB(VBNO+1,1)=VB;
END;  *VB;
END;  *IF VB;

C1='K' 'Region';

IF OUTPUT=1 THEN C2='Theta' 'RP0' 'RP3' 'RP6' 'RP9';
IF OUTPUT=2 THEN C2='V/B' 'RP0' 'RP3' 'RP6' 'RP9';
R1='*';
R2='RP' '1' '2' '3' '4' '5' '6' '7' '8' '9' '10' '11' '12' '13' '14'
   '15' '16' '17' '18' '19' '20' '21' '22' '23' '24' '25' '26' '27';

DATAK=K||REGION;
PRINT DATAK COLNAME=C1 ROWNAME=R1;

IF OUTPUT=1 THEN DO;
PRINT DATAJT COLNAME=C2 ROWNAME=R2;
END;

IF OUTPUT=2 THEN DO;
PRINT DATAJVB COLNAME=C2 ROWNAME=R2;
END;

END; *K;
```

Appendix J

Second Order Response Model

For the situation in which an experimenter fits a second order polynomial model to experimental data, the equation for the *fitted* response model can be written as

$$\hat{y} = b_0 + \sum_{i=1}^{k} b_i x_i + \sum_{i=1}^{k} b_{ii} x_i^2 + \sum_{i<j}\sum b_{ij} x_i x_j \quad .$$

If there is uncertainty as to whether or not the fitted second order model can adequately describe the response surface curvature, then an appropriate experimental plan would protect against biases due to missing third order terms. The equation for the third order *protection* model becomes

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k} \beta_{iii} x_i^3 + \sum_{i<j}\sum \beta_{ij} x_i x_j + \sum_{i<j}\sum \beta_{iij} x_i^2 x_j$$

$$+ \sum_{i<j}\sum \beta_{ijj} x_i x_j^2 + \sum_{h<i<j}\sum\sum \beta_{hij} x_h x_i x_j + \varepsilon \quad .$$

The general linear model form of the protection model, partitioned into a fitted part and an unfitted part, can be written as

$$\underline{y} = X_1\underline{\beta}_1 + X_2\underline{\beta}_2 + \underline{\varepsilon} \,.$$

For the fit-protection situation of $d_1 = 2$ and $d_2 = 3$, the $\underline{\beta}_1$ and $\underline{\beta}_2$ vectors are

$$\underline{\beta}_1 = [\beta_0, \ldots, \beta_k \quad \beta_{11}, \ldots, \beta_{kk} \quad \beta_{12}, \ldots, \beta_{k-1,k}]'$$

$$\underline{\beta}_2 = [\ \beta_{111}, \ldots, \beta_{kkk} \quad \beta_{112}, \ldots, \beta_{k-1,k-1,k} \quad \beta_{122}, \ldots, \beta_{k-1,k,k} \quad \beta_{123}, \ldots, \beta_{k-2,k-1,k}\ ]'$$

where $\underline{\beta}_1$ has dimensions $(p_1 \times 1)$, with $p_1 = 1 + 2k + \binom{k}{2} = \frac{1}{2}(k+1)(k+2)$ coefficients, and $\underline{\beta}_2$ has dimensions $(p_2 \times 1)$, with $p_2 = k + 2\binom{k}{2} + \binom{k}{3} = k^2 + \binom{k}{3} = \frac{1}{6}k(k+1)(k+2)$ coefficients. The corresponding X_1 and X_2 matrices can be written as

$$X_1 = \begin{bmatrix} 1 & x_{11} & \cdots & x_{k1} & x_{11}^2 & \cdots & x_{k1}^2 & x_{11}x_{21} & \cdots & x_{k-1,1}x_{k1} \\ 1 & x_{12} & \cdots & x_{k2} & x_{12}^2 & \cdots & x_{k1}^2 & x_{12}x_{22} & \cdots & x_{k-1,2}x_{k2} \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ 1 & x_{1N} & \cdots & x_{kN} & x_{1N}^2 & \cdots & x_{kN}^2 & x_{1N}x_{2N} & \cdots & x_{k-1,N}x_{kN} \end{bmatrix} (N \times p_1)$$

$$X_2 = \begin{bmatrix} x_{11}^3 \ldots x_{k1}^3 & x_{11}^2 x_{21} \ldots x_{k-1,1}^2 x_{k1} & x_{11}x_{21}^2 \ldots x_{k-1,1}x_{k1}^2 & x_{11}x_{21}x_{31} \ldots x_{k-2,1}x_{k-1,1}x_{k1} \\ x_{12}^3 \ldots x_{k2}^3 & x_{12}^2 x_{22} \ldots x_{k-1,2}^2 x_{k2} & x_{12}x_{22}^2 \ldots x_{k-1,2}x_{k2}^2 & x_{12}x_{22}x_{32} \ldots x_{k-2,2}x_{k-1,2}x_{k2} \\ . \quad . \quad . & . \quad . \quad . & . \quad . \quad . & . \quad . \quad . \\ . \quad . \quad . & . \quad . \quad . & . \quad . \quad . & . \quad . \quad . \\ . \quad . \quad . & . \quad . \quad . & . \quad . \quad . & . \quad . \quad . \\ x_{1N}^3 \ldots x_{kN}^3 & x_{1N}^2 x_{2N} \ldots x_{k-1,N}^2 x_{kN} & x_{1N}x_{2N}^2 \ldots x_{k-1,N}x_{kN}^2 & x_{1N}x_{2N}x_{3N} \ldots x_{k-2,N}x_{k-1,N}x_{kN} \end{bmatrix}$$

$$(N \times p_2)$$

where k is the number of design variables and N is the number of design points.

The design moment matrices are useful in describing the average variance and bias properties of specific design classes. For the case of fitting a second order model ($d_1 = 2$), with protection against a true third order model ($d_2 = 3$), the M_{11} design moment matrix contains moments through order $2d_1 = 4$, the M_{21} matrix contains moments through order $d_1 + d_2 = 5$, and the M_{22} matrix contains moments through order $2d_2 = 6$. For convenience, the (k-1)th variable is denoted by j and the (k-2)th variable is denoted by i in the matrices which follow.

$$M_{11} = N^{-1} X'_1 X_1$$

$$\left[\begin{array}{c|cccc|cccc|cccc}
1 & [1] & [2] & \cdots & [k] & [11] & [22] & \cdots & [kk] & [12] & [13] & \cdots & [jk] \\
\hline
[1] & [11] & [12] & \cdots & [1k] & [111] & [122] & \cdots & [1kk] & [112] & [113] & \cdots & [1jk] \\
[2] & [12] & [22] & \cdots & [2k] & [112] & [222] & \cdots & [2kk] & [122] & [123] & \cdots & [2jk] \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
[k] & [1k] & [2k] & \cdots & [kk] & [11k] & [22k] & \cdots & [kkk] & [12k] & [13k] & \cdots & [jkk] \\
\hline
[11] & [111] & [112] & \cdots & [11k] & [1111] & [1122] & \cdots & [11kk] & [1112] & [1113] & \cdots & [11jk] \\
[22] & [122] & [222] & \cdots & [22k] & [1122] & [2222] & \cdots & [22kk] & [1222] & [1223] & \cdots & [22jk] \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
[kk] & [1kk] & [2kk] & \cdots & [kkk] & [11kk] & [22kk] & \cdots & [kkkk] & [12kk] & [13kk] & \cdots & [jkkk] \\
\hline
[12] & [112] & [122] & \cdots & [12k] & [1112] & [1222] & \cdots & [12kk] & [1122] & [1123] & \cdots & [12jk] \\
[13] & [113] & [123] & \cdots & [13k] & [1113] & [1223] & \cdots & [13kk] & [1123] & [1133] & \cdots & [13jk] \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
[jk] & [1jk] & [2jk] & \cdots & [jkk] & [11jk] & [22jk] & \cdots & [jkkk] & [12jk] & [13jk] & \cdots & [jjkk]
\end{array}\right]$$

$$M_{21} = N^{-1} X'_2 X_1$$

$$
\left[\begin{array}{c|cccc|cccc|cccc}
[111] & [1111] & [1112] & \cdots & [111k] & [11111] & [11122] & \cdots & [111kk] & [11112] & [11113] & \cdots & [111jk] \\
\vdots & \vdots & [2222] & \cdots & [222k] & \vdots & [22222] & \cdots & [222kk] & \vdots & [12223] & \cdots & [222jk] \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
[kkk] & [1kkk] & \cdots & \cdots & [kkkk] & [11kkk] & \cdots & \cdots & [kkkkk] & [12kkk] & \cdots & \cdots & [jkkkk] \\
\hline
[112] & [1112] & [1122] & \cdots & [112k] & [11112] & [11222] & \cdots & [112kk] & [11122] & [11123] & \cdots & [112jk] \\
\vdots & \vdots & [1123] & \cdots & [113k] & \vdots & [11223] & \cdots & [113kk] & \vdots & [11133] & \cdots & [113jk] \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
[ijk] & [1ijk] & \cdots & \cdots & [ijkk] & [11jjk] & \cdots & \cdots & [jjkkk] & [12jjk] & \cdots & \cdots & [jjjkk] \\
\hline
[122] & [1122] & [1222] & \cdots & [122k] & [11122] & [12222] & \cdots & [122kk] & [11222] & [11223] & \cdots & [122jk] \\
\vdots & \vdots & [1233] & \cdots & [133k] & \vdots & [12233] & \cdots & [133kk] & \vdots & [11333] & \cdots & [133jk] \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
[jkk] & [1jkk] & \cdots & \cdots & [jkkk] & [11jkk] & \cdots & \cdots & [ijkkk] & [12jkk] & \cdots & \cdots & [ijkkk] \\
\hline
[123] & [1123] & [1223] & \cdots & [123k] & [11123] & [12223] & \cdots & [123kk] & [11223] & [11233] & \cdots & [123jk] \\
\vdots & \vdots & [1224] & \cdots & [124k] & \vdots & [12224] & \cdots & [124kk] & \vdots & [11234] & \cdots & [124jk] \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
[ijk] & [1ijk] & \cdots & \cdots & [ijkk] & [11ijk] & \cdots & \cdots & [ijkkk] & [12ijk] & \cdots & \cdots & [ijjkk]
\end{array}\right]
$$

$$M_{22} = N^{-1} X'_2 X_2$$

$$
\left[\begin{array}{cccc|cccc|cccc|cccc}
[111111] & [111222] & \cdots & [111kkk] & [111112] & [111113] & \cdots & [111jjk] & [111122] & [111133] & \cdots & [111jkk] & [111123] & [111124] & \cdots & [111ijk] \\
\vdots & [222222] & \cdots & [222kkk] & \vdots & [112223] & \cdots & [222jjk] & \vdots & [122233] & \cdots & [222jkk] & \vdots & [122224] & \cdots & [222ijk] \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
[111kkk] & \cdots & \cdots & [kkkkkk] & [112kkk] & \cdots & \cdots & [jjkkkk] & [122kkk] & \cdots & \cdots & [jkkkkk] & [123kkk] & \cdots & \cdots & [ijkkkk] \\
\hline
[111112] & [112222] & \cdots & [112kkk] & [111122] & [111123] & \cdots & [112jjk] & [111222] & [111233] & \cdots & [112jkk] & [111223] & [111224] & \cdots & [112ijk] \\
\vdots & [112223] & \cdots & [113kkk] & \vdots & [111133] & \cdots & [113jjk] & \vdots & [111333] & \cdots & [113jkk] & \vdots & [111234] & \cdots & [113ijk] \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
[111jjk] & \cdots & \cdots & [jjkkkk] & [112jjk] & \cdots & \cdots & [jjjjkk] & [122jjk] & \cdots & \cdots & [jjjkkk] & [123jjk] & \cdots & \cdots & [ijjjkk] \\
\hline
[111122] & [122222] & \cdots & [122kkk] & [111222] & [111223] & \cdots & [122jjk] & [112222] & [112233] & \cdots & [122jkk] & [112223] & [112224] & \cdots & [122ijk] \\
\vdots & [122233] & \cdots & [133kkk] & \vdots & [111333] & \cdots & [133jjk] & \vdots & [113333] & \cdots & [133jkk] & \vdots & [112334] & \cdots & [133ijk] \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
[111jkk] & \cdots & \cdots & [jkkkkk] & [112jkk] & \cdots & \cdots & [jjjkkk] & [122jkk] & \cdots & \cdots & [jjkkkk] & [123jkk] & \cdots & \cdots & [ijjkkk] \\
\hline
[111123] & [122223] & \cdots & [123kkk] & [111223] & [111233] & \cdots & [123jjk] & [112223] & [112333] & \cdots & [123jkk] & [112233] & [112234] & \cdots & [123ijk] \\
\vdots & [122224] & \cdots & [124kkk] & \vdots & [111234] & \cdots & [124jjk] & \vdots & [112334] & \cdots & [124jkk] & \vdots & [112244] & \cdots & [124ijk] \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
[111ijk] & \cdots & \cdots & [ijkkkk] & [112ijk] & \cdots & \cdots & [ijjjkk] & [122ijk] & \cdots & \cdots & [ijjkkk] & [123ijk] & \cdots & \cdots & [iijjkk]
\end{array}\right]
$$

where the elements of the moment matrices, $[\, 1^{\delta_1}\ 2^{\delta_2} \ldots\ldots k^{\delta_k} \,]$, are defined in Appendix B (pages 278-282).

The corresponding region moment matrices, μ_{11}, μ_{21}, and μ_{22}, are identical in form to the design moment matrices, but the design moments are replaced with region moments (for example, [ii] would be replaced with w_{ii}). Due to the similarity of the design and region moment matrices, the region moment matrices are not illustrated here.

Appendix K

Second Order Model For Estimation of Slopes

For the situation in which a second order polynomial model is being used to fit experimental data, with the purpose of predicting the rate of change of the response variable with respect to the input variables, the equation used to estimate the i^{th} partial derivative of the response function can be written as

$$\hat{\gamma}_{(x_i)} = \frac{\partial \hat{y}}{\partial x_i} = b_i + 2\, b_{ii}\, x_i + \sum_{j \neq i} b_{ij}\, x_j \ .$$

If the experimenter suspects that a third order polynomial model may be needed to describe the true response surface and desires protection against biases due to unfitted third order terms, then the equation for the i^{th} partial derivative of the protection model becomes

$$\gamma_{(x_i)} = \frac{\partial y}{\partial x_i} = \beta_i + 2\, \beta_{ii}\, x_i + 3\, \beta_{iii}\, x_i^2 + \sum_{j \neq i} \beta_{ij}\, x_j + 2 \sum_{j \neq i} \beta_{iij}\, x_i\, x_j$$

$$+ \sum_{j \neq i} \beta_{ijj}\, x_j^2 + \sum_{h \neq i} \ \sum_{< j \neq i} \beta_{hij}\, x_h\, x_j \ .$$

The matrix form of the partial derivative of the response function can be partitioned into two parts, a fitted part and an unfitted part, when model uncertainty exists, written as

$$\underline{\gamma}_{(\underline{x})} = \Lambda'_{1\,(\underline{x})}\underline{\beta}_1 + \Lambda'_{2\,(\underline{x})}\underline{\beta}_2 .$$

For the fit-protection situation of $d_1 = 2$ and $d_2 = 3$, the $\underline{\beta}$ vectors of the partitioned model are

$$\underline{\beta}_1 = [\beta_0, \ldots, \beta_k \quad \beta_{11}, \ldots, \beta_{kk} \quad \beta_{12}, \ldots, \beta_{k-1,k}]'$$

$$\underline{\beta}_2 = [\beta_{111}, \ldots, \beta_{kkk} \quad \beta_{112}, \ldots, \beta_{k-1,k-1,k} \quad \beta_{122}, \ldots, \beta_{k-1,k,k} \quad \beta_{123}, \ldots, \beta_{k-2,k-1,k}]'$$

where $\underline{\beta}_1$ has dimensions $(p_1 \times 1)$, with $p_1 = 1 + 2k + \binom{k}{2} = \frac{1}{2}(k+1)(k+2)$ coefficients, and $\underline{\beta}_2$ has dimensions $(p_2 \times 1)$, with $p_2 = k + 2\binom{k}{2} + \binom{k}{3} = \frac{1}{6}k(k+1)(k+2)$ coefficients. The corresponding $\Lambda'_{1(\underline{x})}$ matrix can be written as

$$\Lambda'_{1\,(\underline{x})} = \left[\begin{array}{c|ccccc|ccccc|cccccccccccc}
0 & 1 & 0 & . & . & 0 & 2x_1 & 0 & . & . & 0 & x_2 & x_3 & . & . & x_k & 0 & 0 & . & . & 0 & . & 0 \\
0 & 0 & 1 & . & . & 0 & 0 & 2x_2 & . & . & 0 & x_1 & 0 & . & . & 0 & x_3 & x_4 & . & . & x_k & . & 0 \\
0 & 0 & 0 & . & . & 0 & 0 & 0 & . & . & 0 & 0 & x_1 & . & . & 0 & x_2 & 0 & . & . & 0 & . & 0 \\
0 & 0 & 0 & . & . & 0 & 0 & 0 & . & . & 0 & 0 & 0 & . & . & 0 & 0 & x_2 & . & . & 0 & . & 0 \\
. & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
0 & 0 & 0 & . & . & 1 & 0 & 0 & . & . & 2x_k & 0 & 0 & . & . & x_1 & 0 & 0 & . & . & x_2 & . & x_{k-1}
\end{array}\right]$$

where $\Lambda'_{1(\underline{x})}$ is a $k \times p_1$ matrix. The correspnding $\Lambda_{2(\underline{x})}$ matrix can be written as

$$
\Lambda_{2(\underline{x})} = \left[\begin{array}{cccccc}
3x_1^2 & 0 & 0 & . & 0 & 0 \\
0 & 3x_2^2 & 0 & . & 0 & 0 \\
. & . & . & . & . & . \\
0 & 0 & 0 & . & 0 & 3x_k^2 \\
\hline
2x_1x_2 & x_1^2 & 0 & . & 0 & 0 \\
2x_1x_3 & 0 & x_1^2 & . & 0 & 0 \\
. & . & . & . & . & . \\
2x_1x_k & 0 & 0 & . & 0 & x_1^2 \\
0 & 2x_2x_3 & x_2^2 & . & 0 & 0 \\
0 & 2x_2x_4 & 0 & . & 0 & 0 \\
. & . & . & . & . & . \\
0 & 2x_2x_k & 0 & . & 0 & x_2^2 \\
0 & 0 & 0 & . & 2x_{k-1}x_k & x_{k-1}^2 \\
\hline
x_2^2 & 2x_1x_2 & 0 & . & 0 & 0 \\
x_3^2 & 0 & 2x_1x_3 & . & 0 & 0 \\
0 & . & . & . & . & . \\
x_k^2 & 0 & 0 & . & 0 & 2x_1x_k \\
0 & x_3^2 & 2x_2x_3 & . & 0 & 0 \\
0 & x_4^2 & 0 & . & 0 & 0 \\
. & . & . & . & . & . \\
0 & x_k^2 & 0 & . & 0 & 2x_2x_k \\
0 & 0 & 0 & . & x_k^2 & 2x_{k-1}x_k \\
\hline
x_2x_3 & x_1x_3 & x_1x_2 & . & 0 & 0 \\
x_2x_4 & x_1x_4 & 0 & . & 0 & 0 \\
. & . & . & . & . & . \\
0 & 0 & 0 & . & x_{k-2}x_k & x_{k-2}x_{k-1}
\end{array}\right]
$$

where $\Lambda_{2(\underline{x})}$ is a $p_2 \times k$ matrix.

The matrices of the partial derivatives of the response function, integrated over the region of interest, are useful in describing the average variance and bias properties of designs used in the estimation of the slope coefficients. The region matrices for the partial derivatives of the response function are denoted by μ_{11}^*, μ_{21}^*, and μ_{22}^* . For convenience, the $(k-1)^{th}$ variable is denoted by j and the

$(k\text{-}2)^{th}$ variable is denoted by i in the matrices which follow. (The elements of the region matrices are defined in Appendix D on pages 287-291.)

$$\mu^*_{11} = \Omega_r \int_R \Lambda_{1\,(\underline{x})}\, \Lambda'_{1\,(\underline{x})}\, d\underline{x}$$

$$\left[\begin{array}{c|cccc|cccc|ccccc}
0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \;\cdot\cdot & 0 \\
\hline
0 & 1 & 0 & \cdots & 0 & 2w_1 & 0 & \cdots & 0 & w_2 & w_3 & \cdots & w_k \;\cdot\cdot & 0 \\
: & 0 & 1 & \cdots & 0 & 0 & 2w_2 & \cdots & 0 & w_1 & 0 & \cdots & 0 \;\cdot\cdot & : \\
: & : & : & \cdots & : & : & : & \cdots & : & : & : & \cdots & : \;\cdot\cdot & : \\
0 & 0 & \cdots & \cdots & 1 & 0 & \cdots & \cdots & 2w_k & 0 & \cdots & \cdots & w_1 \;\cdot\cdot & w_j \\
\hline
0 & 2w_1 & 0 & \cdots & 0 & 4w_{11} & 0 & \cdots & 0 & 2w_{12} & 2w_{13} & \cdots & 2w_{1k} \;\cdot\cdot & 0 \\
: & 0 & 2w_2 & \cdots & 0 & 0 & 4w_{22} & \cdots & 0 & 2w_{12} & 0 & \cdots & 0 \;\cdot\cdot & : \\
: & : & : & \cdots & : & : & : & \cdots & : & : & : & \cdots & : \;\cdot\cdot & : \\
0 & 0 & \cdots & \cdots & 2w_k & 0 & \cdots & \cdots & 4w_{kk} & 0 & \cdots & \cdots & 2w_{1k} \;\cdot\cdot & 2w_{jk} \\
\hline
0 & w_2 & w_1 & \cdots & 0 & 2w_{12} & 2w_{12} & \cdots & 0 & w_{11}+w_{22} & w_{23} & \cdots & w_{2k} \;\cdot\cdot & 0 \\
: & w_3 & 0 & \cdots & 0 & 2w_{13} & 0 & \cdots & 0 & w_{23} & w_{11}+w_{33} & \cdots & w_{3k} \;\cdot\cdot & : \\
: & : & : & \cdots & : & : & : & \cdots & : & : & : & \cdots & : \;\cdot\cdot & : \\
0 & w_k & \cdots & \cdots & w_1 & 2w_{1k} & \cdots & \cdots & 2w_{1k} & w_{2k} & \cdots & \cdots & w_{11}+w_{kk} \;\cdot\cdot & w_{1j} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \;\cdot\cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \;\cdot\cdot & \cdot \\
0 & 0 & \cdots & \cdots & w_j & 0 & \cdots & \cdots & 2w_{jk} & 0 & \cdots & \cdots & w_{1j} \;\cdot\cdot & w_{jj}+w_{kk}
\end{array}\right]$$

$$\mu^{*}_{21} = \Omega_r \int_R \Lambda_{2\,(\underline{x})}\, \Lambda'_{1\,(\underline{x})}\, d\underline{x}$$

$$\left[\begin{array}{c|cccc|cccc|cccccc}
0 & 3w_{11} & 0 & \cdots & 0 & 6w_{111} & 0 & \cdots & 0 & 3w_{112} & 3w_{113} & \cdots & 3w_{11k} & \cdot\;\cdot & 0 \\
\vdots & 0 & 3w_{22} & \cdots & 0 & 0 & 6w_{222} & \cdots & 0 & 3w_{122} & 0 & \cdots & 0 & \cdot\;\cdot & 0 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \cdot\;\cdot & \vdots \\
0 & 0 & \cdots & \cdots & 3w_{kk} & 0 & \cdots & \cdots & 6w_{kkk} & 0 & \cdots & \cdots & 3w_{1kk} & \cdot\;\cdot & 3w_{jkk} \\
\hline
0 & 2w_{12} & w_{11} & \cdots & 0 & 4w_{112} & 2w_{112} & \cdots & 0 & 2w_{112}+w_{111} & 2w_{123} & \cdots & 2w_{12k} & \cdot\;\cdot & 0 \\
\vdots & 2w_{13} & 0 & \cdots & 0 & 4w_{113} & 0 & \cdots & 0 & 2w_{123} & 2w_{133}+w_{111} & \cdots & 2w_{13k} & \cdot\;\cdot & 0 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \cdot\;\cdot & \vdots \\
0 & 0 & \cdots & \cdots & w_{jj} & 0 & \cdots & \cdots & 2w_{jjk} & 0 & \cdots & \cdots & 2w_{1jj} & \cdot\;\cdot & w_{jjj} \\
\hline
0 & w_{22} & 2w_{12} & \cdots & 0 & 2w_{122} & 4w_{122} & \cdots & 0 & w_{222}+2w_{112} & w_{223} & \cdots & w_{22k} & \cdot\;\cdot & 0 \\
\vdots & w_{33} & 0 & \cdots & 0 & 2w_{133} & 0 & \cdots & 0 & w_{233} & w_{333}+2w_{113} & \cdots & w_{33k} & \cdot\;\cdot & 0 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \cdot\;\cdot & \vdots \\
0 & 0 & \cdots & \cdots & 2w_{jk} & 0 & \cdots & \cdots & 4w_{jkk} & 0 & \cdots & \cdots & 2w_{1jk} & \cdot\;\cdot & w_{kkk}+2w_{jjk} \\
\hline
0 & w_{23} & w_{13} & \cdots & 0 & 2w_{123} & 2w_{113} & \cdots & 0 & w_{223}+w_{113} & w_{233}+w_{112} & \cdots & w_{23k} & \cdot\;\cdot & 0 \\
\vdots & w_{24} & w_{14} & \cdots & 0 & 2w_{124} & 2w_{124} & \cdots & 0 & w_{224}+w_{114} & w_{234} & \cdots & w_{24k} & \cdot\;\cdot & 0 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \cdot\;\cdot & \vdots \\
0 & 0 & \cdots & \cdots & w_{ij} & 0 & \cdots & \cdots & 2w_{ijk} & 0 & \cdots & \cdots & w_{1ij} & \cdot\;\cdot & w_{ijj}+w_{ikk}
\end{array}\right]$$

$$\overset{*}{\mu}_{22} = \Omega_r \int_R \Lambda_{2(x)} \Lambda'_{2(x)} \, d\underline{x}$$

$$\left[\begin{array}{cccc|cccc|cccc|cccc}
9w_{1111} & 0 & \cdots & 0 & 6w_{1112} & 6w_{1113} & \cdots & 0 & 3w_{1122} & 3w_{1133} & \cdots & 0 & 3w_{1123} & 3w_{1124} & \cdots & 0 \\
0 & 9w_{2222} & \cdots & 0 & 3w_{1122} & 0 & \cdots & 0 & 6w_{1222} & 0 & \cdots & 0 & 3w_{1223} & 3w_{1224} & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
0 & \cdots & \cdots & 9w_{kkkk} & 0 & \cdots & \cdots & 3w_{jjkk} & 0 & \cdots & \cdots & 2w_{jkkk} & 0 & \cdots & \cdots & 3w_{ijkk} \\
\hline
6w_{1112} & 3w_{1122} & \cdots & 0 & 4w_{1122}+w_{1111} & 4w_{1123} & \cdots & 0 & 2w_{1222}+2w_{1112} & 2w_{1233} & \cdots & 0 & 2w_{1223}+w_{1113} & 2w_{1224}+w_{1114} & \cdots & 0 \\
6w_{1113} & 0 & \cdots & 0 & 4w_{1123} & 4w_{1133}+w_{1111} & \cdots & 0 & 2w_{1223} & 2w_{1333}+2w_{1113} & \cdots & 0 & 2w_{1233}+w_{1112} & 2w_{1234} & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
0 & \cdots & \cdots & 3w_{ijkk} & 0 & \cdots & \cdots & 4w_{ijkk}+w_{iiii} & 0 & \cdots & \cdots & 2w_{jkkk}+2w_{iijk} & 0 & \cdots & \cdots & 2w_{ijkk}+w_{iiij} \\
\hline
3w_{1122} & 6w_{1222} & \cdots & 0 & 2w_{1222}+2w_{1112} & 2w_{1223} & \cdots & 0 & w_{2222}+4w_{1122} & w_{2233} & \cdots & 0 & w_{2223}+2w_{1123} & w_{2224}+2w_{1124} & \cdots & 0 \\
3w_{1133} & 0 & \cdots & 0 & 2w_{1233} & 2w_{1333}+2w_{1113} & \cdots & 0 & w_{2233} & w_{3333}+4w_{1133} & \cdots & 0 & w_{2333}+2w_{1123} & w_{2334} & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
0 & \cdots & \cdots & 2w_{jkkk} & 0 & \cdots & \cdots & 2w_{iijk} & 0 & \cdots & \cdots & w_{kkkk}+4w_{jjkk} & 0 & \cdots & \cdots & w_{ikkk}+2w_{iijk} \\
\hline
3w_{1123} & 3w_{1223} & \cdots & 0 & 2w_{1223}+w_{1113} & 2w_{1233}+w_{1112} & \cdots & 0 & w_{2223}+2w_{1123} & w_{2333}+2w_{1123} & \cdots & 0 & w_{2233}+w_{1133}+w_{1122} & w_{2234}+w_{1134} & \cdots & 0 \\
3w_{1124} & 3w_{1224} & \cdots & 0 & 2w_{1224}+w_{1114} & 2w_{1234} & \cdots & 0 & w_{2224}+2w_{1124} & w_{2234} & \cdots & 0 & w_{2234}+w_{1134} & . & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
0 & \cdots & \cdots & 3w_{ijkk} & 0 & \cdots & \cdots & 2w_{ijkk}+w_{iiij} & 0 & \cdots & \cdots & w_{ikkk}+2w_{iijk} & 0 & \cdots & \cdots & w_{jjkk}+w_{iikk}+w_{iiij}
\end{array}\right]$$

Appendix L

Second Order Design Plans

Second order response surface designs are experimental plans useful in the estimation of second order polynomial models. Such designs require that each factor be present at three or more levels and that there be at least one design point for each of the $p_1 = \frac{1}{2}(k+1)(k+2)$ estimated coefficients, of which $k+1$ are linear, k are quadratic, and $\binom{k}{2}$ are two-way interaction coefficients.

If there is doubt as to whether or not a second order polynomial model can adequately describe the response surface curvature in the region of interest, then it may be desirable to use a second order design which affords some protection against cubic curvature. The $p_2 = \frac{1}{6}k(k+1)(k+2)$ third order coefficients would not be included in the fitted model, but the experimental plan would be designed to protect against bias due to missing cubic terms.

For second order polynomial models with $k = 2, 3, 4, 5, 6, 7$ factors, Table L-1 summarizes the number of fitted model terms (p_1) and the number of unfitted third order model terms (p_2).

Table L-1. Fitting a Second Order Response Surface Model with Protection Against Third Order Bias. Number of factors, fitted model terms (p_1), and unfitted third order terms (p_2).

Values of :			Number of Fitted Coefficients $p_1 = \frac{1}{2}(k+1)(k+2)$				Number of Unfitted Coefficients $p_2 = \frac{1}{6}k(k+1)(k+2)$			
k	p_1	p_2	β_0 1	β_i k	β_{ii} k	β_{ij} $\binom{k}{2}$	β_{iii} k	β_{iij} $\binom{k}{2}$	β_{ijj} $\binom{k}{2}$	β_{ijk} $\binom{k}{3}$
2	6	4	1	2	2	1	2	1	1	0
3	10	10	1	3	3	3	3	3	3	1
4	15	20	1	4	4	6	4	6	6	4
5	21	35	1	5	5	10	5	10	10	10
6	28	57	1	6	6	15	6	15	15	20
7	36	84	1	7	7	21	7	21	21	35

The following three design properties are particularly important in the evaluation of the correlation induction strategies examined in this research:

1. *Odd* design moments through order five are equal to zero.
2. *Even*-order design moments are constant.
3. The design can be partitioned into orthogonal blocks.

The orthogonal blocking property is necessary for implementation of the assignment rule blocking strategy and the design moment conditions enable the equations for **B** and **B*** to be written in terms of three (instead of p_2) unknown third order parameters, regardless of the value of *k*. These conditions also lead to a common form of Var $[\underline{b}_1]$, thereby simplifying the equations for **V** and **V***. Because of the importance of the orthogonal blocking property and the design moment conditions, all of the designs used in this research (except the small composite designs) have been selected for achievement of these properties. The SCDs have the last two properties, but do not have the first, and therefore the optimal SCDs cannot be determined analytically.

The four second order design classes which are used in this research for evaluation of the correlation induction strategies are as follows:

1. Central Composite Designs (CCDs)

   a) Full factorial replications $k = 2, 3, 4, 5$ factors.
   b) ½-fractional factorials $k = 6, 7$ factors.

2. Box-Behnken Designs (BBDs) $k = 4, 5, 7$ factors.

3. Three-level Factorial Designs (FACs)

   a) Full factorial replications $k = 3, 4, 5$ factors.
   b) ⅓-fractional factorials $k = 6, 7$ factors.

4. Small Composite Designs (SCDs)

   a) Hartley Designs (SCD-H) $k = 3, 4, 6$ factors.
   b) Draper Designs (SCD-D) $k = 5, 7$ factors.

For each of the specific second order designs used in this research, Table L-2 summarizes the number of factors (k), the number of blocks (b), the number of non-center design points, $(N - N_c)$, the ratio of [iiii] to [iijj] (r), and whether the design is orthogonally blockable. For the CCDs and SCDs, the orthogonal blocking property is indicated by the axial distance, α, needed to achieve the property.

The specific CCDs, BBDs, FACs (for $k = 3, 4$ only), and SCDs used in this research are illustrated on the following pages of this appendix:

Design Class	# Factors (k)	Pages
CCDs	2, 3, 4, 5, 6, 7	363-369
BBDs	4, 5, 7	371-372
FACs	3, 4	373-375
SCDs	3, 4, 5, 6, 7	376-380

Table L-2. Second Order Response Surface Designs.
Design classes, number of factors, blocks, and design points.

DESIGN CLASS	Number of Factors (k)	Number of Blocks (b)	Number of Non-Center Points ($N-N_c$)	Ratio of [iiii] to [iijj] (r)	Orthogonal Blocking
CCD	2	2	8	3	$\alpha = 1.414$
CCD	3	2	14	3.42	$\alpha = 1.764$
CCD	4	2	24	3.24	$\alpha = 2.058$
CCD	5	2	42	2.78	$\alpha = 2.309$
CCD	6 (½ fr.)	2	44	3.48	$\alpha = 2.511$
CCD	7 (½ fr.)	2	78	2.70	$\alpha = 2.717$
BBD	4	3	24	3	Yes
BBD	5	2	40	4	Yes
BBD	7	2	56	3	Yes
FAC	3	3	27	1.5	Yes
FAC	4	3	81	1.5	Yes
FAC	5	3	243	1.5	Yes
FAC	6 (⅓ fr.)	3	243	1.5	Yes
FAC	7 (⅓ fr.)	3	729	1.5	Yes
SCD-H	3	2	10	*	*
SCD-H	4	2	16	*	*
SCD-D	5	2	22	*	*
SCD-H	6	2	28	*	*
SCD-D	7	2	42	4.75	$\alpha = 2.691$

Note: The * indicates that the value of α needed for orthogonal blocking in the $k=3, 4, 5, 6$ SCDs, and the resulting value of r, depend on the number of center runs required for the **J** and **J*** optimal designs. (See pages 376-379 for the values of α and r.)

k=2 CCD

$$
D = \begin{array}{c} \\ \end{array}
\begin{array}{cc} x_1 & x_2 \end{array}
$$

$$
D = \left[\begin{array}{cc}
-1 & -1 \\
1 & -1 \\
-1 & 1 \\
1 & 1 \\
0 & 0 \\
\hdashline
-\alpha & 0 \\
\alpha & 0 \\
0 & -\alpha \\
0 & \alpha \\
0 & 0 \\
\hdashline
0 & 0
\end{array}\right] \times g
$$

(1) for Min-V^*

$F = 4$
$n_a = 4$

$\alpha = 1.4142136$ for orthogonal blocking

$r = [iiii] / [iijj] = 3$

Spherical Region :

g = .5773503 for Min-**B**

g = .7071068 for Min-$\mathbf{B}^*$

$N_c \cong 2$ for Min-**V**|Min-**B**

$N_c \cong 3$ for Min $\mathbf{V}^*$|Min-$\mathbf{B}^*$

Cuboidal Region :

g = .6831301 for Min-**B**

g = .8164966 for Min-$\mathbf{B}^*$

$N_c \cong 2$ for Min **V**|Min-**B**

$N_c \cong 3$ for Min $\mathbf{V}^*$|Min-$\mathbf{B}^*$

k = 3 CCD

$$
\begin{array}{cc}
 & \begin{array}{ccc} X_1 & X_2 & X_3 \end{array} \\
D = & \left[\begin{array}{ccc}
-1 & -1 & -1 \\
-1 & 1 & 1 \\
1 & -1 & 1 \\
1 & 1 & -1 \\
-1 & -1 & -1 \\
-1 & 1 & 1 \\
1 & -1 & 1 \\
1 & 1 & -1 \\
0 & 0 & 0 \\
\hdashline
\alpha & 0 & 0 \\
-\alpha & 0 & 0 \\
0 & \alpha & 0 \\
0 & -\alpha & 0 \\
0 & 0 & \alpha \\
0 & 0 & -\alpha \\
0 & 0 & 0 \\
\hdashline
0 & 0 & 0
\end{array}\right] \times g
\end{array}
$$

(1) for Min-V^*

$F = 8$
$n_a = 6$

$\alpha = 1.7638342$ for orthogonal blocking

$r = [iiii] / [iijj] = 3.4197531$

Spherical Region :

$g = .4840441$ for Min-$\mathbf{B}$
$g = .5727288$ for Min-$\mathbf{B}^*$

$N_c \cong 2$ for Min-$\mathbf{V}$|Min-$\mathbf{B}$
$N_c \cong 3$ for Min $\mathbf{V}^*$|Min-$\mathbf{B}^*$

Cuboidal Region :

$g = .6445849$ for Min-$\mathbf{B}$
$g = .7393896$ for Min-$\mathbf{B}^*$

$N_c \cong 2$ for Min $\mathbf{V}$|Min-$\mathbf{B}$
$N_c \cong 3$ for Min $\mathbf{V}^*$|Min-$\mathbf{B}^*$

k = 4 CCD

$$
D = \begin{array}{c}
\begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \end{array} \\
\left[\begin{array}{cccc}
-1 & -1 & -1 & -1 \\
-1 & -1 & 1 & 1 \\
-1 & 1 & -1 & 1 \\
-1 & 1 & 1 & -1 \\
1 & -1 & -1 & 1 \\
1 & -1 & 1 & -1 \\
1 & 1 & -1 & -1 \\
1 & 1 & 1 & 1 \\
-1 & 1 & 1 & 1 \\
-1 & -1 & -1 & 1 \\
-1 & 1 & -1 & -1 \\
-1 & -1 & 1 & -1 \\
1 & 1 & -1 & 1 \\
1 & 1 & 1 & -1 \\
1 & -1 & -1 & -1 \\
1 & -1 & 1 & -1 \\
0 & 0 & 0 & 0 \\
\hdashline
-\alpha & 0 & 0 & 0 \\
\alpha & 0 & 0 & 0 \\
0 & -\alpha & 0 & 0 \\
0 & \alpha & 0 & 0 \\
0 & 0 & -\alpha & 0 \\
0 & 0 & \alpha & 0 \\
0 & 0 & 0 & -\alpha \\
0 & 0 & 0 & \alpha \\
0 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & 0
\end{array}\right]
\end{array}
\begin{array}{l} \times g \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ (N_c - 2) \end{array}
$$

$F = 16$

$n_a = 8$

$\alpha = 2.0579830$ for orthogonal blocking

$r = [iiii] / [iijj] = 3.2422145$

Spherical Region :

$g = .4286704$ for Min-B

$g = .4949859$ for Min-B*

$N_c \cong 2$ for Min-V|Min-B

$N_c \cong 4$ for Min V*|Min-B*

Cuboidal Region :

$g = .6261132$ for Min-B

$g = .7000158$ for Min-B*

$N_c \cong 3$ for Min V|Min-B

$N_c \cong 4$ for Min V*|Min-B*

k=5 CCD

$$
D = \begin{array}{c}
\begin{array}{ccccc} X_1 & X_2 & X_3 & X_4 & X_5 \end{array} \\
\left[\begin{array}{ccccc}
-1 & -1 & -1 & -1 & -1 \\
-1 & 1 & -1 & 1 & 1 \\
-1 & 1 & 1 & -1 & 1 \\
-1 & -1 & 1 & 1 & -1 \\
1 & -1 & 1 & -1 & 1 \\
1 & -1 & 1 & 1 & -1 \\
1 & 1 & -1 & -1 & -1 \\
1 & 1 & -1 & 1 & 1 \\
-1 & -1 & -1 & -1 & 1 \\
-1 & -1 & -1 & 1 & -1 \\
-1 & 1 & 1 & -1 & -1 \\
-1 & 1 & 1 & 1 & 1 \\
1 & -1 & 1 & -1 & -1 \\
1 & -1 & 1 & 1 & 1 \\
1 & 1 & -1 & -1 & 1 \\
1 & 1 & -1 & 1 & -1 \\
-1 & -1 & 1 & -1 & 1 \\
-1 & -1 & 1 & 1 & -1 \\
-1 & 1 & -1 & -1 & -1 \\
-1 & 1 & -1 & 1 & 1 \\
1 & -1 & -1 & -1 & -1 \\
1 & -1 & -1 & 1 & 1 \\
1 & 1 & 1 & -1 & 1 \\
1 & 1 & 1 & 1 & -1 \\
-1 & -1 & 1 & -1 & -1 \\
-1 & -1 & 1 & 1 & 1 \\
-1 & 1 & -1 & -1 & 1 \\
-1 & 1 & -1 & 1 & -1 \\
1 & -1 & -1 & -1 & 1 \\
1 & -1 & -1 & 1 & -1 \\
1 & 1 & 1 & -1 & -1 \\
1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 \\
\hline
\pm\alpha & 0 & 0 & 0 & 0 \\
0 & \pm\alpha & 0 & 0 & 0 \\
0 & 0 & \pm\alpha & 0 & 0 \\
0 & 0 & 0 & \pm\alpha & 0 \\
0 & 0 & 0 & 0 & \pm\alpha \\
0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0
\end{array}\right]
\end{array}
\begin{array}{l}
\\ \\ \times g \\ \\ (2) \\ (2) \\ (2) \\ (2) \\ (2) \\ \\ (N_c - 2)
\end{array}
$$

$F = 32$ $\qquad$ $\alpha = 2.3094011$ for orthogonal blocking

$n_a = 10$ $\qquad$ $r = [iiii] / [iijj] = 2.7777778$

Spherical Region :

$g = .3911591$ for Min-B

$g = .4435328$ for Min-B*

$N_c \cong 3$ for Min-V|Min-B

$N_c \cong 5$ for Min V*|Min-B*

Cuboidal Region :

$g = .6167073$ for Min-B

$g = .6775075$ for Min-B*

$N_c \cong 4$ for Min V|Min-B

$N_c \cong 5$ for Min V*|Min-B*

k = 6 (½ fraction) CCD

$$
\mathbf{D} = \begin{array}{c}
\begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{array} \\
\left[\begin{array}{cccccc}
-1 & -1 & -1 & -1 & -1 & 1 \\
-1 & -1 & -1 & -1 & 1 & -1 \\
-1 & -1 & -1 & 1 & -1 & -1 \\
-1 & -1 & -1 & 1 & 1 & 1 \\
-1 & 1 & 1 & -1 & -1 & 1 \\
-1 & 1 & 1 & -1 & 1 & -1 \\
-1 & 1 & 1 & 1 & -1 & -1 \\
-1 & 1 & 1 & 1 & 1 & 1 \\
1 & -1 & 1 & -1 & -1 & 1 \\
1 & -1 & 1 & -1 & 1 & -1 \\
1 & -1 & 1 & 1 & -1 & -1 \\
1 & -1 & 1 & 1 & 1 & 1 \\
1 & 1 & -1 & -1 & -1 & 1 \\
1 & 1 & -1 & -1 & 1 & -1 \\
1 & 1 & -1 & 1 & -1 & -1 \\
1 & 1 & -1 & 1 & 1 & 1 \\
-1 & -1 & 1 & -1 & -1 & -1 \\
-1 & -1 & 1 & -1 & 1 & 1 \\
-1 & -1 & 1 & 1 & -1 & 1 \\
-1 & -1 & 1 & 1 & 1 & -1 \\
-1 & 1 & -1 & -1 & -1 & -1 \\
-1 & 1 & -1 & -1 & 1 & 1 \\
-1 & 1 & -1 & 1 & -1 & 1 \\
-1 & 1 & -1 & 1 & 1 & -1 \\
1 & -1 & -1 & -1 & -1 & -1 \\
1 & -1 & -1 & -1 & 1 & 1 \\
1 & -1 & -1 & 1 & -1 & 1 \\
1 & -1 & -1 & 1 & 1 & -1 \\
1 & 1 & 1 & -1 & -1 & -1 \\
1 & 1 & 1 & -1 & 1 & 1 \\
1 & 1 & 1 & 1 & -1 & 1 \\
1 & 1 & 1 & 1 & 1 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
\pm\alpha & 0 & 0 & 0 & 0 & 0 \\
0 & \pm\alpha & 0 & 0 & 0 & 0 \\
0 & 0 & \pm\alpha & 0 & 0 & 0 \\
0 & 0 & 0 & \pm\alpha & 0 & 0 \\
0 & 0 & 0 & 0 & \pm\alpha & 0 \\
0 & 0 & 0 & 0 & 0 & \pm\alpha \\
0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]
\end{array}
\begin{array}{l}
\\ \\
\times g \\
(2) \\ (2) \\ (2) \\ (2) \\ (2) \\ (2) \\ \\
(N_c - 2)
\end{array}
$$

Note: Defining Contrast = $x_1x_2x_3x_4x_5x_6$

$F = 32$ $\qquad \alpha = 2.5105837$ for orthogonal blocking

$n_a = 12$ $\qquad r = [iiii] / [iijj] = 3.4830119$

Spherical Region :

g = .3625700 for Min-**B**

g = .4053656 for Min-**B***

$N_c \cong 2$ for Min-V|Min-**B**

$N_c \cong 5$ for Min **V***|Min-**B***

Cuboidal Region :

g = .6102963 for Min-**B**

g = .6619593 for Min-**B***

$N_c \cong 3$ for Min **V**|Min-**B**

$N_c \cong 5$ for Min **V***|Min-**B***

k = 7 (½ Fraction) CCD

$$
D = \begin{array}{c}
\begin{matrix} X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 \end{matrix} \\
\left[\begin{matrix}
-1 & -1 & -1 & -1 & 1 & 1 & -1 \\
-1 & -1 & -1 & 1 & -1 & -1 & 1 \\
-1 & 1 & 1 & -1 & 1 & -1 & 1 \\
-1 & 1 & 1 & 1 & -1 & 1 & -1 \\
1 & -1 & 1 & -1 & -1 & 1 & 1 \\
1 & -1 & 1 & 1 & 1 & -1 & -1 \\
1 & 1 & -1 & -1 & -1 & -1 & -1 \\
1 & 1 & -1 & 1 & 1 & 1 & 1 \\
-1 & -1 & -1 & -1 & -1 & -1 & -1 \\
-1 & -1 & -1 & 1 & 1 & 1 & 1 \\
-1 & 1 & 1 & -1 & -1 & 1 & 1 \\
-1 & 1 & 1 & 1 & 1 & -1 & -1 \\
1 & -1 & 1 & -1 & 1 & -1 & 1 \\
1 & -1 & 1 & 1 & -1 & 1 & -1 \\
1 & 1 & -1 & -1 & 1 & 1 & -1 \\
1 & 1 & -1 & 1 & -1 & -1 & 1 \\
-1 & -1 & -1 & -1 & 1 & -1 & 1 \\
-1 & -1 & -1 & 1 & -1 & 1 & -1 \\
-1 & 1 & 1 & -1 & 1 & 1 & -1 \\
-1 & 1 & 1 & 1 & -1 & -1 & 1 \\
1 & -1 & 1 & -1 & -1 & -1 & -1 \\
1 & -1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & -1 & -1 & -1 & 1 & 1 \\
1 & 1 & -1 & 1 & 1 & -1 & -1 \\
-1 & -1 & -1 & -1 & -1 & 1 & 1 \\
-1 & -1 & -1 & 1 & 1 & -1 & -1 \\
-1 & 1 & 1 & -1 & -1 & -1 & -1 \\
-1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & -1 & 1 & -1 & 1 & 1 & -1 \\
1 & -1 & 1 & 1 & -1 & -1 & 1 \\
1 & 1 & -1 & -1 & 1 & -1 & 1 \\
1 & 1 & -1 & 1 & -1 & 1 & -1 \\
-1 & -1 & 1 & -1 & -1 & -1 & 1 \\
-1 & -1 & 1 & 1 & 1 & 1 & -1 \\
-1 & 1 & -1 & -1 & -1 & 1 & -1 \\
-1 & 1 & -1 & 1 & 1 & -1 & 1 \\
1 & -1 & -1 & -1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1 & -1 & 1 & 1 \\
1 & 1 & 1 & -1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & -1 & -1 & -1 \\
-1 & -1 & 1 & -1 & 1 & 1 & 1 \\
-1 & -1 & 1 & 1 & -1 & -1 & -1 \\
-1 & 1 & -1 & -1 & 1 & -1 & -1 \\
-1 & 1 & -1 & 1 & -1 & 1 & 1 \\
1 & -1 & -1 & -1 & -1 & 1 & -1 \\
1 & -1 & -1 & 1 & 1 & -1 & 1 \\
1 & 1 & 1 & -1 & -1 & -1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & -1
\end{matrix}\right]
\end{array} \times g
$$

(continued on next page)

k = 7 (½ Fraction) CCD

(continuation)

x_1	x_2	x_3	x_4	x_5	x_6	x_7	
-1	-1	1	-1	-1	1	-1	
-1	-1	1	1	1	-1	1	
-1	1	-1	-1	-1	-1	1	
-1	1	-1	1	1	1	-1	
1	-1	-1	-1	1	1	1	
1	-1	-1	1	-1	-1	-1	
1	1	1	-1	1	-1	-1	
1	1	1	1	-1	1	1	
-1	-1	1	-1	1	-1	-1	
-1	-1	1	1	-1	1	1	
-1	1	-1	-1	1	1	1	
-1	1	-1	1	-1	-1	-1	
1	-1	-1	-1	-1	-1	1	
1	-1	-1	1	1	1	-1	
1	1	1	-1	-1	1	-1	
1	1	1	1	1	-1	1	
0	0	0	0	0	0	0	
$\pm\alpha$	0	0	0	0	0	0	(2)
0	$\pm\alpha$	0	0	0	0	0	(2)
0	0	$\pm\alpha$	0	0	0	0	(2)
0	0	0	$\pm\alpha$	0	0	0	(2)
0	0	0	0	$\pm\alpha$	0	0	(2)
0	0	0	0	0	$\pm\alpha$	0	(2)
0	0	0	0	0	0	$\pm\alpha$	(2)
0	0	0	0	0	0	0	
0	0	0	0	0	0	0	$(N_c - 2)$

Note: Defining Contrast = $x_1x_2x_3x_4x_5x_6x_7$

$F = 64$
$n_a = 14$

$\alpha = 2.7174649$ for orthogonal blocking

$r = [iiii] / [iijj] = 2.7041420$

Spherical Region :

$g = .3401342$ for Min-B

$g = .3760325$ for Min-B*

$N_c \cong 3$ for Min-V|Min-B

$N_c \cong 7$ for Min V*|Min-B*

Cuboidal Region :

$g = .6063341$ for Min-B

$g = .6513073$ for Min-B*

$N_c \cong 4$ for Min V|Min-B

$N_c \cong 7$ for Min V*|Min-B*

k = 4 BBD

$$D = \begin{array}{c} \\ \end{array}\begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \end{array}$$

$$D = \left[\begin{array}{cccc}
-1 & -1 & 0 & 0 \\
-1 & 1 & 0 & 0 \\
1 & -1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
0 & 0 & -1 & -1 \\
0 & 0 & -1 & 1 \\
0 & 0 & 1 & -1 \\
0 & 0 & 1 & 1 \\
\hdashline
-1 & 0 & 0 & -1 \\
-1 & 0 & 0 & 1 \\
1 & 0 & 0 & -1 \\
1 & 0 & 0 & 1 \\
0 & -1 & -1 & 0 \\
0 & -1 & 1 & 0 \\
0 & 1 & -1 & 0 \\
0 & 1 & 1 & 0 \\
\hdashline
-1 & 0 & -1 & 0 \\
-1 & 0 & 1 & 0 \\
1 & 0 & -1 & 0 \\
1 & 0 & 1 & 0 \\
0 & -1 & 0 & -1 \\
0 & -1 & 0 & 1 \\
0 & 1 & 0 & -1 \\
0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 \; (N_c)
\end{array}\right] \times g$$

F = 24

r = [iiii] / [iijj] = 3

Spherical Region :

g = .6123724 for Min-**B**

g = .7071068 for Min-**B***

$N_c \cong 2$ for Min-**V**|Min-**B**

$N_c \cong 4$ for Min **V***|Min-**B***

Cuboidal Region :

g = .8944272 for Min-**B**

g = 1.0 for Min-**B***

$N_c \cong 3$ for Min **V**|Min-**B**

$N_c \cong 4$ for Min **V***|Min-**B***

k = 5 BBD

$$
D = \begin{array}{c}
\begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\
\left[\begin{array}{ccccc}
\pm 1 & \pm 1 & 0 & 0 & 0 \\
0 & 0 & \pm 1 & \pm 1 & 0 \\
0 & \pm 1 & 0 & 0 & \pm 1 \\
\pm 1 & 0 & \pm 1 & 0 & 0 \\
0 & 0 & 0 & \pm 1 & \pm 1 \\
0 & 0 & 0 & 0 & 0 \\
\hdashline
0 & \pm 1 & \pm 1 & 0 & 0 \\
\pm 1 & 0 & 0 & \pm 1 & 0 \\
0 & 0 & \pm 1 & 0 & \pm 1 \\
\pm 1 & 0 & 0 & 0 & \pm 1 \\
0 & \pm 1 & 0 & \pm 1 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & 0 & 0
\end{array}\right]
\end{array}
\begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ (N_c - 2) \end{array}
$$

Note: The ± 1's indicate the following combinations of the two variables:

$$
\begin{bmatrix} -1 & -1 \\ -1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}
$$

F = 40

r = [iiii] / [iijj] = 4

Spherical Region :

g = .6236096 for Min-B

g = .7071068 for Min-B*

$N_c \cong 3$ for Min-V|Min-B

$N_c \cong 5$ for Min V*|Min-B*

Cuboidal Region :

g = .9831921 for Min-B

g = 1.0801234 for Min-B*

$N_c \cong 3$ for Min V|Min-B

$N_c \cong 5$ for Min V*|Min-B*

k = 7 BBD

$$
D = \left[\begin{array}{ccccccc}
x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \\
0 & 0 & 0 & \pm 1 & \pm 1 & \pm 1 & 0 \\
\pm 1 & 0 & 0 & 0 & 0 & \pm 1 & \pm 1 \\
0 & \pm 1 & 0 & 0 & \pm 1 & 0 & \pm 1 \\
\pm 1 & \pm 1 & 0 & \pm 1 & 0 & 0 & 0 \\
0 & 0 & \pm 1 & \pm 1 & 0 & 0 & \pm 1 \\
\pm 1 & 0 & \pm 1 & 0 & \pm 1 & 0 & 0 \\
0 & \pm 1 & \pm 1 & 0 & 0 & \pm 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & \pm 1 & \pm 1 & \pm 1 & 0 \\
\pm 1 & 0 & 0 & 0 & 0 & \pm 1 & \pm 1 \\
0 & \pm 1 & 0 & 0 & \pm 1 & 0 & \pm 1 \\
\pm 1 & \pm 1 & 0 & \pm 1 & 0 & 0 & 0 \\
0 & 0 & \pm 1 & \pm 1 & 0 & 0 & \pm 1 \\
\pm 1 & 0 & \pm 1 & 0 & \pm 1 & 0 & 0 \\
0 & \pm 1 & \pm 1 & 0 & 0 & \pm 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]
\begin{array}{l} \\ \\ \\ \\ \\ \\ \\ \\ \times g \\ \\ \\ \\ \\ \\ \\ \\ \\ (N_c - 2) \end{array}
$$

Note: The ± 1's indicate the following combinations of the three variables:

First Block

$$\begin{bmatrix} -1 & -1 & -1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

Second Block

$$\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

F = 56

r = [iiii] / [iijj] = 3

Spherical Region :

g = .5222330 for Min-B

g = .5773503 for Min-B*

$N_c \cong 2$ for Min-V|Min-B

$N_c \cong 5$ for Min V*|Min-B*

Cuboidal Region :

g = .9309493 for Min-B

g = 1.0 for Min-B*

$N_c \cong 3$ for Min V|Min-B

$N_c \cong 5$ for Min V*|Min-B*

k = 3 Factorial Design

$$D = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \left[\begin{array}{ccc} -1 & -1 & 1 \\ -1 & 0 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & -1 \\ \hdashline -1 & -1 & 0 \\ -1 & 0 & -1 \\ -1 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & -1 & -1 \\ 0 & 0 & 0 \\ \hdashline -1 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \end{array}\right] \end{array} \times g$$

F = 27

$N_c \cong 0$ for Min-V|Min-B and Min V*|Min-B*

r = [iiii] / [iijj] = 1.5

Spherical Region :

g = .5532833 for Min-B

g = .6546537 for Min-B*

Cuboidal Region :

g = .7367884 for Min-B

g = .8451543 for Min-B*

k = 4 Factorial Design

$$
D = \begin{array}{c}
\begin{array}{cccc} X_1 & X_2 & X_3 & X_4 \end{array} \\
\left[\begin{array}{cccc}
-1 & -1 & -1 & 1 \\
-1 & -1 & 0 & 0 \\
-1 & -1 & 1 & -1 \\
-1 & 0 & -1 & 0 \\
-1 & 0 & 0 & -1 \\
-1 & 0 & 1 & 1 \\
-1 & 1 & -1 & -1 \\
-1 & 1 & 0 & 1 \\
-1 & 1 & 1 & 0 \\
0 & -1 & -1 & 0 \\
0 & -1 & 0 & -1 \\
0 & -1 & 1 & 1 \\
0 & 0 & -1 & -1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & -1 & 1 \\
0 & 1 & 0 & 0 \\
0 & 1 & 1 & -1 \\
1 & -1 & -1 & -1 \\
1 & -1 & 0 & 1 \\
1 & -1 & 1 & 0 \\
1 & 0 & -1 & 1 \\
1 & 0 & 0 & 0 \\
1 & 0 & 1 & -1 \\
1 & 1 & -1 & 0 \\
1 & 1 & 0 & -1 \\
1 & 1 & 1 & 1 \\
\hdashline
-1 & -1 & -1 & -1 \\
-1 & -1 & 0 & 1 \\
-1 & -1 & 1 & 0 \\
-1 & 0 & -1 & 1 \\
-1 & 0 & 0 & 0 \\
-1 & 0 & 1 & -1 \\
-1 & 1 & -1 & 0 \\
-1 & 1 & 0 & -1 \\
-1 & 1 & 1 & 1 \\
0 & -1 & -1 & 1 \\
0 & -1 & 0 & 0 \\
0 & -1 & 1 & -1 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1 \\
0 & 0 & 1 & 1 \\
0 & 1 & -1 & -1 \\
0 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 \\
1 & -1 & -1 & 0 \\
1 & -1 & 0 & -1 \\
1 & -1 & 1 & 1 \\
1 & 0 & -1 & -1 \\
1 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 \\
1 & 1 & -1 & 1 \\
1 & 1 & 0 & 0 \\
1 & 1 & 1 & -1 \\
\hdashline
\end{array}\right]
\end{array} \times g
$$

(continued on next page)

k = 4 Factorial Design

(continuation)

x_1	x_2	x_3	x_4
-1	-1	-1	0
-1	-1	0	-1
-1	-1	1	1
-1	0	-1	-1
-1	0	0	1
-1	0	1	0
-1	1	-1	1
-1	1	0	0
-1	1	1	-1
0	-1	-1	-1
0	-1	0	1
0	-1	1	0
0	0	-1	1
0	0	0	0
0	0	1	-1
0	1	-1	0
0	1	0	-1
0	1	1	1
1	-1	-1	1
1	-1	0	0
1	-1	1	-1
1	0	-1	0
1	0	0	-1
1	0	1	1
1	1	-1	-1
1	1	0	1
1	1	1	0

Note: Blocking Contrast = $x_1x_2x_3x_4$

F = 81

$N_c \cong 0$ for Min-V|Min-B and Min V*|Min-B*

r = [iiii] / [iijj] = 1.5

Spherical Region :

g = .5 for Min-B

g = .5773503 for Min-B*

Cuboidal Region :

g = .7302967 for Min-B

g = .8164966 for Min-B*

k = 3 SCD-Hartley

$$D = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \left[\begin{array}{ccc} 1 & 1 & 1 \\ -1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ 0 & 0 & 0 \\ \hdashline \alpha & 0 & 0 \\ -\alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & -\alpha & 0 \\ 0 & 0 & \alpha \\ 0 & 0 & -\alpha \\ 0 & 0 & 0 \end{array}\right] \end{array} \times g$$

(0) for Min-B in a Spherical Region

Note: Defining Contrast = $x_1x_2x_3$

Alias Pairs: $x_1 \equiv x_2x_3$
$x_2 \equiv x_1x_3$
$x_3 \equiv x_1x_2$

(Two-way interaction terms are NOT aliased with each other.)

$F = 4$
$n_a = 6$

$\alpha = 1.5491933$ for orthogonal blocking when $N_c = 1$
$\alpha = 1.6733201$ for orthogonal blocking when $N_c = 2$

$r = [iiii] / [iijj] = 3.88$ when $N_c = 1$
$r = [iiii] / [iijj] = 4.92$ when $N_c = 2$

Spherical Region :

$g = .5169620$ for Min-B
$g = .5889149$ for Min-B*

$N_c \cong 1$ for Min-V|Min-B
$N_c \cong 2$ for Min V*|Min-B*

Cuboidal Region :

$g = .6628019$ for Min-B
$g = .7602859$ for Min-B*

$N_c \cong 2$ for Min V|Min-B
$N_c \cong 2$ for Min V*|Min-B*

k = 4 SCD-Hartley

$$
D = \begin{array}{c}
\begin{array}{cccc} x_1 & \quad x_2 & \quad x_3 & \quad x_4 \end{array} \\
\left[\begin{array}{cccc}
-1 & -1 & -1 & -1 \\
-1 & -1 & 1 & 1 \\
-1 & 1 & -1 & 1 \\
-1 & 1 & 1 & -1 \\
1 & -1 & -1 & 1 \\
1 & -1 & 1 & -1 \\
1 & 1 & -1 & -1 \\
1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 \\
\hdashline
-\alpha & 0 & 0 & 0 \\
\alpha & 0 & 0 & 0 \\
0 & -\alpha & 0 & 0 \\
0 & \alpha & 0 & 0 \\
0 & 0 & -\alpha & 0 \\
0 & 0 & \alpha & 0 \\
0 & 0 & 0 & -\alpha \\
0 & 0 & 0 & \alpha \\
0 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & 0
\end{array}\right]
\end{array}
$$

(0) for Min-B in Sph. Region

$\times g$

(1) for Min-V*

Note: Defining Contrast = $x_1x_2x_3$

Alias Pairs:

$x_1 \equiv x_2x_3$ $x_1x_4 \equiv x_2x_3x_4$

$x_2 \equiv x_1x_3$ $x_2x_4 \equiv x_1x_3x_4$

$x_3 \equiv x_1x_2$ $x_3x_4 \equiv x_1x_2x_4$

$x_4 \equiv x_1x_2x_3x_4$

(Two-way interaction terms are NOT aliased with each other.)

$F = 8$

$n_a = 8$

$\alpha = 2.1213203$ for orthogonal blocking when $N_c = 1$

$\alpha = 2.0$ for orthogonal blocking when $N_c = 2$

$r = [iiii] / [iijj] = 6.0625$ when $N_c = 1$

$r = [iiii] / [iijj] = 5$ when $N_c = 2$

Spherical Region :

$g = .4193591$ for Min-B

$g = .5$ for Min-B*

$N_c \cong 1$ for Min-V|Min-B

$N_c \cong 3$ for Min V*|Min-B*

Cuboidal Region :

$g = .6324555$ for Min-B

$g = .7071068$ for Min-B*

$N_c \cong 2$ for Min V|Min-B

$N_c \cong 3$ for Min V*|Min-B*

k = 5 SCD-Draper

$$
D = \begin{array}{c}
\begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\
\left[\begin{array}{rrrrr}
-1 & -1 & -1 & -1 & -1 \\
1 & 1 & 1 & 1 & 1 \\
-1 & -1 & 1 & 1 & 1 \\
-1 & 1 & 1 & -1 & 1 \\
1 & -1 & -1 & 1 & 1 \\
1 & 1 & -1 & 1 & -1 \\
1 & 1 & 1 & -1 & -1 \\
1 & -1 & 1 & -1 & -1 \\
1 & -1 & -1 & -1 & 1 \\
-1 & 1 & -1 & 1 & -1 \\
-1 & 1 & -1 & -1 & 1 \\
-1 & -1 & 1 & 1 & -1 \\
0 & 0 & 0 & 0 & 0 \\
\hline
\pm\alpha & 0 & 0 & 0 & 0 \\
0 & \pm\alpha & 0 & 0 & 0 \\
0 & 0 & \pm\alpha & 0 & 0 \\
0 & 0 & 0 & \pm\alpha & 0 \\
0 & 0 & 0 & 0 & \pm\alpha \\
0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0
\end{array}\right]
\end{array}
\begin{array}{l}
\\ \\ \\ \\ \\ \\ \\ \\
\times g \\
\\ \\ \\ \\
\text{(0) for Min-}\mathbf{B} \\
\text{in Sph. Region} \\
(2) \\
(2) \\
(2) \\
(2) \\
(2) \\
\\
\\
\text{(1) for Min-}\mathbf{V}^*
\end{array}
$$

Note: Alias structure is that of a 12-run Plackett-Burman design.

$F = 12$

$n_a = 10$

$\alpha = 2.3452079$ for orthogonal blocking when $N_c = 1$

$\alpha = 2.2532028$ for orthogonal blocking when $N_c = 2$

$r = [iiii] / [iijj] = 6.0416667$ when $N_c = 1$

$r = [iiii] / [iijj] = 5.2958580$ when $N_c = 2$

Spherical Region :

$g = .3852992$ for Min-**B**

$g = .4456452$ for Min-**B***

$N_c \cong 1$ for Min-**V**|Min-**B**

$N_c \cong 3$ for Min **V***|Min-**B***

Cuboidal Region :

$g = .6196445$ for Min-**B**

$g = .6807342$ for Min-**B***

$N_c \cong 2$ for Min **V**|Min-**B**

$N_c \cong 3$ for Min **V***|Min-**B***

k = 6 SCD-Hartley

$$D = \begin{array}{c} \begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{array} \\ \left[\begin{array}{cccccc}
1 & -1 & -1 & 1 & -1 & -1 \\
-1 & 1 & -1 & 1 & -1 & -1 \\
-1 & -1 & 1 & 1 & -1 & -1 \\
1 & 1 & 1 & 1 & -1 & -1 \\
1 & -1 & -1 & -1 & 1 & -1 \\
-1 & 1 & -1 & -1 & 1 & -1 \\
-1 & -1 & 1 & -1 & 1 & -1 \\
1 & 1 & 1 & -1 & 1 & -1 \\
1 & -1 & -1 & -1 & -1 & 1 \\
-1 & 1 & -1 & -1 & -1 & 1 \\
-1 & -1 & 1 & -1 & -1 & 1 \\
1 & 1 & 1 & -1 & -1 & 1 \\
1 & -1 & -1 & 1 & 1 & 1 \\
-1 & 1 & -1 & 1 & 1 & 1 \\
-1 & -1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\hline
\pm\alpha & 0 & 0 & 0 & 0 & 0 \\
0 & \pm\alpha & 0 & 0 & 0 & 0 \\
0 & 0 & \pm\alpha & 0 & 0 & 0 \\
0 & 0 & 0 & \pm\alpha & 0 & 0 \\
0 & 0 & 0 & 0 & \pm\alpha & 0 \\
0 & 0 & 0 & 0 & 0 & \pm\alpha \\
0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0
\end{array}\right] \end{array}
\begin{array}{l}
\\ \\ \\ \\ \\ \\ \\ \\ \\ \times g \\ \\ \\ \\ \\ \\ \\ \\ \text{(0) for Min-}\mathbf{B} \\
(2) \\ (2) \\ (2) \\ (2) \\ (2) \\ (2) \\ \\
\text{(1) for Min-}\mathbf{V}^*
\end{array}$$

Note: Defining Contrasts = $x_1x_2x_3$ and $x_4x_5x_6$

Alias Pairs: $x_1 \equiv x_2x_3$ $\quad$ $x_4 \equiv x_5x_6$

$x_2 \equiv x_1x_3$ $\quad$ $x_5 \equiv x_6x_7$

$x_3 \equiv x_1x_2$ $\quad$ $x_6 \equiv x_4x_5$

and other higher order alias pairs

$F = 16 \qquad n_a = 12$

$\alpha = 2.5495098$ when $N_c = 1$

$r = 6.2812500$ when $N_c = 1$

Spherical Region :

$g = .3585134$ for Min-$\mathbf{B}$

$g = .4065249$ for Min-$\mathbf{B}^*$

$N_c \cong 1$ for Min-$\mathbf{V}$|Min-$\mathbf{B}$

$N_c \cong 3$ for Min $\mathbf{V}^*$|Min-$\mathbf{B}^*$

$\alpha = 2.4733878$ when $N_c = 2$

$r = 5.6782007$ when $N_c = 2$

Cuboidal Region :

$g = .6120416$ for Min-$\mathbf{B}$

$g = .6638523$ for Min-$\mathbf{B}^*$

$N_c \cong 2$ for Min $\mathbf{V}$|Min-$\mathbf{B}$

$N_c \cong 3$ for Min $\mathbf{V}^*$|Min-$\mathbf{B}^*$

k = 7 SCD-Draper

$$
D = \begin{array}{c}
\begin{array}{ccccccc} X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 \end{array} \\
\left[\begin{array}{ccccccc}
1 & 1 & -1 & 1 & 1 & -1 & 1 \\
-1 & 1 & 1 & -1 & 1 & 1 & -1 \\
1 & -1 & 1 & 1 & -1 & 1 & 1 \\
1 & 1 & -1 & 1 & 1 & -1 & 1 \\
1 & 1 & 1 & -1 & 1 & 1 & -1 \\
1 & 1 & 1 & 1 & -1 & 1 & 1 \\
-1 & 1 & 1 & 1 & 1 & -1 & 1 \\
-1 & -1 & 1 & 1 & 1 & 1 & -1 \\
-1 & -1 & -1 & 1 & 1 & 1 & 1 \\
-1 & -1 & -1 & -1 & 1 & 1 & 1 \\
1 & -1 & -1 & -1 & -1 & 1 & 1 \\
-1 & 1 & -1 & -1 & -1 & -1 & 1 \\
-1 & -1 & 1 & -1 & -1 & -1 & -1 \\
-1 & -1 & -1 & 1 & -1 & -1 & -1 \\
1 & -1 & -1 & -1 & 1 & -1 & -1 \\
-1 & 1 & -1 & -1 & -1 & 1 & -1 \\
-1 & -1 & 1 & -1 & -1 & -1 & 1 \\
1 & -1 & -1 & 1 & -1 & -1 & -1 \\
1 & 1 & -1 & -1 & 1 & -1 & -1 \\
1 & 1 & 1 & -1 & -1 & 1 & -1 \\
-1 & 1 & 1 & 1 & -1 & -1 & 1 \\
1 & -1 & 1 & 1 & 1 & -1 & -1 \\
-1 & 1 & -1 & 1 & 1 & 1 & -1 \\
1 & -1 & 1 & -1 & 1 & 1 & 1 \\
1 & 1 & -1 & 1 & -1 & 1 & 1 \\
-1 & 1 & 1 & -1 & 1 & -1 & 1 \\
1 & -1 & 1 & 1 & -1 & 1 & -1 \\
-1 & -1 & -1 & -1 & -1 & -1 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
\pm\alpha & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \pm\alpha & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \pm\alpha & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \pm\alpha & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \pm\alpha & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \pm\alpha & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \pm\alpha \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right] \times g
\end{array}
\begin{array}{l}
\\ \\
(2) \\ (2) \\ (2) \\ (2) \\ (2) \\ (2) \\ (2) \\ \\
\text{(1) for Min-V*}
\end{array}
$$

Note: Alias structure is that of a 28-run Plackett-Burman design.

$F = 28$ $\qquad \alpha = 2.6909811$ for orthogonal blocking

$n_a = 14$ $\qquad r = [iiii] / [iijj] = 4.7455410$

Spherical Region :

g = .3398898 for Min-B

g = .3757623 for Min-B*

$N_c \cong 2$ for Min-V|Min-B

$N_c \cong 3$ for Min V*|Min-B*

Cuboidal Region :

g = .6058985 for Min-B

g = .6508394 for Min-B*

$N_c \cong 2$ for Min V|Min-B

$N_c \cong 3$ for Min V*|Min-B*

Appendix M

Variance-Covariance Matrix of Second Order Model Coefficients

The variance components of the MSE of response and MSE of slope criteria involve the variance-covariance matrix of the fitted model coefficients. For the situation in which ordinary least squares is used for estimation of the model parameters, the $(p_1 \times p_1)$ $\mathrm{Var}[\underline{b}_1]$ matrix can be written as

$$\begin{aligned}\mathrm{Var}[\underline{b}_1] &= (X'_1 X_1)^{-1} X'_1 \; V X_1 (X'_1 X_1)^{-1} \sigma^2 \\ &= Z' V Z \sigma^2\end{aligned}$$

where $\mathbf{Z}$ is an $(N \times p_1)$ matrix defined for convenience reasons as

$$Z = X_1 (X'_1 X_1)^{-1}$$

and V is the $(N \times N)$ variance-covariance matrix of the response, illustrated in Appendix A (pages 273-277) for the **IR**, **CR**, and **AR** correlation induction strategies. Substituting the equations for V_{IR}, V_{CR}, and V_{AR} into the equation for $\mathrm{Var}[\underline{b}_1]$ yields the following equations for the variance-covariance matrix under the **IR**, **CR**, and **AR** correlation induction strategies:

$$\mathrm{Var}[\underline{b}_1]_{IR} = Z' \, [\, I_N \,] \, Z \, \sigma^2$$

$$= (X'_1 X_1)^{-1} \sigma^2$$

$$\mathrm{Var}[\underline{b}_1]_{CR} = Z' \left[\, I_N + \rho_+ \underline{u}\,\underline{u}' - \rho_+ U \,\right] Z \, \sigma^2$$

$$= \left[\, (X'_1 X_1)^{-1} + \rho_+ Z'\underline{u}\,\underline{u}'Z - \rho_+ Z'UZ \,\right] \sigma^2$$

$$\mathrm{Var}[\underline{b}_1]_{AR} = Z' \left[\, I_N + \tfrac{1}{2}(\rho_+ - \rho_-)\underline{u}\,\underline{u}' + \tfrac{1}{2}(\rho_+ + \rho_-)\underline{v}\,\underline{v}' - \rho_+ U \,\right] Z \, \sigma^2$$

$$= \left[\, (X'_1 X_1)^{-1} + \tfrac{1}{2}(\rho_+ - \rho_-) Z'\underline{u}\,\underline{u}'Z + \tfrac{1}{2}(\rho_+ + \rho_-) Z'\underline{v}\,\underline{v}'Z - \rho_+ Z'UZ \,\right] \sigma^2$$

where the elements of $\underline{u}$, $\underline{v}$, and U are defined in Appendix A (pages 273-277). Inspection of the three forms of the variance-covariance matrix indicates that the following four matrices are needed for determination of $\mathrm{Var}[\underline{b}_1]$ for the three strategies:

1. $(X'_1 X_1)^{-1}$
2. $Z' \underline{u}\,\underline{u}' Z$
3. $Z' \underline{v}\,\underline{v}' Z$
4. $Z' U Z$.

In the case of a first order two-level factorial design, inversion of $(X'_1 X_1)$ is simply an inversion of the diagonal elements. However, for the second order designs used in this research, $(X'_1 X_1)$ is not a diagonal matrix and the form of $(X'_1 X_1)^{-1}$ is not necessarily the same for each second order design class. Appendix J (pages 348-352) illustrates the $M_{11} = N^{-1}(X'_1 X_1)$ moment matrix in the second order case. Inspection of the M_{11} matrix indicates that when all of the *odd* design moments through order four are equal to zero, the matrix is block diagonal (with the exception of the [ii] terms in the first row and column). The CCDs, BBDs, and FACs considered in this research have this property, resulting in a common form of the inverted M_{11} matrix. However, the SCDs do not necessarily have odd design moments equal to zero, and inversion of the M_{11} matrix is more complicated.

The remainder of this appendix is restricted to the CCDs, BBDs, FACs (fractional CCDs and FACs with $k \geq 6$ factors) which have the following two design moment conditions (for all *i, j, k, l, m*):

1. *Odd* design moments through order five are equal to zero.
   $[i] = [ij] = [iij] = [ijk] = [iiij] = [iijk] = [ijkl] = [iiiij] = [iiijj] = [iiijk] = [iijjk] = [iijkl] = [ijklm] = 0$
2. *Even* order design moments through order four are constant.
   $[ii] = [jj] \qquad [iiii] = [jjjj] \qquad [iijj] = [kkll]$

In the case of the AR strategy, the forms of the $Z' \underline{u}\,\underline{u}' Z$ and $Z' U Z$ matrices depend on the number of orthogonal blocks in the design. For the designs which partition into two orthogonal blocks (CCDs and $k = 5, 7$ BBDs), a common form for each of these matrices exists. For the designs which partition into three orthogonal blocks ($k = 4$ BBD and the FACs), the form of $Z' \underline{u}\,\underline{u}' Z$ can be altered through the use of a term which accounts for the number of blocks in the design. The $Z' U Z$ matrix, however, is different for the designs which partition into three blocks and separate forms of this matrix are presented here.

For the CCDs, BBDs, and FACs, the $(p_1 \times p_1)$ M_{11} moment matrix shown in Appendix J (page 350) simplifies to

$$
M_{11} = \left[\begin{array}{c|c|c|c}
1 & \mathbf{0} & [ii]\,\mathbf{1}_k' & \mathbf{0} \\
\hline
\mathbf{0} & [ii]\,\mathbf{I}_k & \mathbf{0} & \mathbf{0} \\
\hline
[ii]\,\mathbf{1}_k & \mathbf{0} & \begin{array}{c}[iiii]\,\mathbf{I}_k \\ +\,[iijj]\,\mathbf{1}_k\mathbf{1}_k' \\ -\,[iijj]\,\mathbf{I}_k\end{array} & \mathbf{0} \\
\hline
\mathbf{0} & \mathbf{0} & \mathbf{0} & [iijj]\,\mathbf{I}_{\binom{k}{2}}
\end{array}\right].
$$

Upon inversion of the M_{11} matrix and division by N, the $(X'_1 X_1)^{-1}$ matrix becomes

$$
(X'_1 X_1)^{-1} = \left[\begin{array}{c|c|c|c}
\dfrac{a}{N\,D} & \mathbf{0} & \dfrac{-[ii]}{N\,D}\,\mathbf{1}_k' & \mathbf{0} \\
\hline
\mathbf{0} & \dfrac{1}{N\,[ii]}\,\mathbf{I}_k & \mathbf{0} & \mathbf{0} \\
\hline
\dfrac{-[ii]}{N\,D}\,\mathbf{1}_k & \mathbf{0} & \begin{array}{c}\dfrac{(b-c)}{N\,D_2}\,\mathbf{I}_k \\ +\,\dfrac{c}{N\,D_2}\,\mathbf{1}_k\mathbf{1}_k'\end{array} & \mathbf{0} \\
\hline
\mathbf{0} & \mathbf{0} & \mathbf{0} & \dfrac{1}{N\,[iijj]}\,\mathbf{I}_{\binom{k}{2}}
\end{array}\right]
$$

where $a = [iiii] + (k-1)\,[iijj]$

$b = [iiii] + (k-2)\,[iijj] - (k-1)\,[ii]^2$

$c = [ii]^2 - [iijj]$

$D = [iiii] + (k-1)\,[iijj] - k\,[ii]^2$

$D_2 = D\,(\,[iiii] - [iijj]\,)$.

Recall from Appendix A that $\underline{u}$ is an $(N\times1)$ vector whose i^{th} element is $u_i = 1$ if a common or antithetic random number stream is used to generate the i^{th} design point and $u_i = 0$ if an independent random number stream is used. The CR strategy utilizes independent streams for replications of center runs and the AR strategy utilizes independent streams for replications of center runs within blocks and for all design points in the third block (if one exists). For the CR and AR strategies, the $(p_1 \times p_1)$ $Z'\underline{u}\,\underline{u}'Z$ matrix becomes

$$Z'\underline{u}\,\underline{u}'Z = \left[\begin{array}{c|c|c|c} s^2 & 0 & s\,t\,\underline{1}_k' & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline s\,t\,\underline{1}_k & 0 & t^2\,\underline{1}_k\underline{1}_k' & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right]$$

where $s = \dfrac{(N-N_2)\,a \;-\; h\,k\,N\,[ii]^2}{N\,D}$

$t = \dfrac{N_2\,[ii] \;+\; (h-1)\,N\,[ii]}{N\,D}$

N_2 = number of design points generated with independent streams

$$h = \left\{\begin{array}{cc} 2\,(\text{number of blocks})^{-1} & \text{for the AR strategy} \\ 1 & \text{for the CR strategy} \end{array}\right\}.$$

Recall from Appendix A that $\underline{v}$ is an $(N \times 1)$ vector whose i^{th} element is $v_i = +1$ if a common stream is used to generate the i^{th} design point, $v_i = 0$ if an independent stream is used, and $v_i = -1$ if an antithetic stream is used. For the AR strategy, the $(p_1 \times p_1)$ $Z' \underline{v}\underline{v}' Z$ matrix becomes

$$Z' \underline{v}\underline{v}' Z = \left[\begin{array}{c|c|c|c} p^2 & 0 & m\,p\,\underline{1}_k' & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline m\,p\,\underline{1}_k & 0 & m^2\,\underline{1}_k\underline{1}_k' & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right]$$

where $p = \dfrac{(N_{1a}-N_{1b})\,a \; - \; k\,[ii]\,\Delta}{N\,D}$

$$m = \frac{(N_{1b}-N_{1a})\,[ii] \; + \; \Delta}{N\,D}$$

N_{1a} = number of design points generated with common streams (block 1)

N_{1b} = number of design points generated with antithetic streams (block 2)

$\Delta = N_{1a}\,[ii]_{1a} \; - \; N_{1b}\,[ii]_{1b}$ = Block 1 sum of squares — Block 2 sum of squares .

For the BBDs and FACs, $Z' \underline{v}\underline{v}' Z$ is a null matrix because the sizes of the first two blocks and the sum of squares within these blocks are equal.

Recall from Appendix A that U is an $(N \times N)$ diagonal matrix of the vector $\underline{\mathbf{u}}$. Rather than present the $Z'\,U\,Z$ matrix, it is convenient to take advantage of the sparsity of the matrix obtained by subtracting U from the identity matrix, denoted by W, and defined as

$$W = I_N - U$$

$$Z'\,W\,Z = (X'_1\,X_1)^{-1} - Z'\,U\,Z\,.$$

The $Z'\,W\,Z$ matrix is more sparse than $Z'\,U\,Z$ because the CR and AR strategies generally use fewer independent streams than common and/or antithetic streams. For the CR strategy, the number of blocks into which the design partitions is irrelevant and the $Z'\,W\,Z$ matrix shown below applies to the CCDs, BBDs, and FACs under the CR strategy. For the AR strategy, the $Z'\,W\,Z$ matrix shown below only applies to the designs which partition into two orthogonal blocks (CCDs and $k = 5, 7$ BBDs). Thus, for the AR strategy in the case of two blocks, and for the CR strategy in general, the $(p_1 \times p_1)$ $Z'\,W\,Z$ matrix becomes

$$Z'\,W\,Z = \left[\begin{array}{c|c|c|c} q^2 & 0 & r\,q\,\mathbf{1}_k' & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline r\,q\,\mathbf{1}_k & 0 & r^2\,\mathbf{1}_k\mathbf{1}_k' & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right]$$

where $q = \sqrt{N_2}\,\dfrac{a}{N\,D}$

$$r = -\sqrt{N_2}\,\frac{[ii]}{N\,D}\,.$$

Under the AR strategy, the form of $Z' W Z$ is slightly different for the designs which partition into three orthogonal blocks. For the 3^k designs, which have no center runs except the one inherent in the full factorial designs ($N_c = 0$), the $Z' W Z$ matrix for the FACs under the AR strategy only becomes

$$Z' W Z = (1-h)\,(X'_1 X_1)^{-1} \, .$$

In addition to the FACs, the $k=4$ BBD also partitions into three orthogonal blocks. Under the AR strategy, all of the center runs are placed in the 3rd block, and therefore utilize independent random number streams. The $Z' W Z$ matrix for the $k=4$ BBD under the AR strategy only becomes

$$Z' W Z = \left[\begin{array}{c|c|c|c}
\dfrac{a}{N D} & 0 & -\dfrac{[ii]}{N D}\,\mathbf{1}_k' & 0 \\
\hline
0 & \dfrac{(1-h)}{N\,[ii]}\,\mathbf{I}_k & 0 & 0 \\
\hline
-\dfrac{[ii]}{N D}\,\mathbf{1}_k & 0 & \begin{array}{cc|cc} e & f & e & f \\ f & e & f & e \\ \hdashline e & f & e & f \\ f & e & f & e \end{array} & 0 \\
\hline
0 & 0 & 0 & \begin{array}{c|c} \dfrac{1}{N\,[iijj]}\,\mathbf{I}_2 & 0 \\ \hdashline 0 & 0 \end{array}
\end{array}\right]$$

where
$$e = \frac{(1-h)\,b}{N D_2} + \frac{h\,N_c\,[ii]^2}{(N D)^2}$$
$$f = \frac{N_c\,[ii]^2}{(N D)^2} - \frac{4}{(N\,[ii])^2} \, .$$

Appendix N

V and V* Equations in the Second Order Case

The variance components of **J** and **J*** involve the variance-covariance matrix of the fitted model coefficients, $\mathrm{Var}[\underline{b}_1]$, and the region matrices, μ_{11} and μ_{11}^*. For the MSE of response and MSE of slope criteria, respectively, the variance components can be written as

$$\mathbf{V} = \frac{N}{\sigma_2} \text{ Trace}\left\{\mathrm{Var}[\underline{b}_1]\ \mu_{11}\right\}$$

$$\mathbf{V}^* = \frac{N}{\sigma_2} \text{ Trace}\left\{\mathrm{Var}[\underline{b}_1]\ \mu_{11}^*\right\}$$

where the component parts of the $\mathrm{Var}[\underline{b}_1]$ matrix are shown in Appendix M (pages 381-388) for the second order designs used in this research. The μ_{11} region moment matrix has the same form as the M_{11} moment matrix shown in Appendix M, except that the design moments are replaced with region moments. (For example, [iijj] is replaced with w_{iijj}.) For the second order designs considered in this research, the $(p_1 \times p_1)$ μ_{11} matrix becomes

$$
\mu_{11} = \left[\begin{array}{c|c|c|c}
1 & \mathbf{0} & w_{ii}\,\mathbf{1}_k' & \mathbf{0} \\
\hline
\mathbf{0} & w_{ii}\,\mathbf{I}_k & \mathbf{0} & \mathbf{0} \\
\hline
w_{ii}\,\mathbf{1}_k & \mathbf{0} & \begin{array}{c} w_{iiii}\,\mathbf{I}_k \\ + w_{iijj}\,\mathbf{1}_k\mathbf{1}_k' \\ - w_{iijj}\,\mathbf{I}_k \end{array} & \mathbf{0} \\
\hline
\mathbf{0} & \mathbf{0} & \mathbf{0} & w_{iijj}\,\mathbf{I}_{\binom{k}{2}}
\end{array}\right]
$$

where the w_{ii}, w_{iiii}, and w_{iijj} terms are the region moments of the design, defined in Appendix D (pages 287-291) for spherical and cuboidal regions of interest.

The μ_{11}^* region matrix is shown in Appendix K (page 356) for the general case of fitting a second order model. For the second order designs used in this research, which have *odd* design moments through order four equal to zero, the $(p_1 \times p_1)$ μ_{11}^* matrix simplifies to

$$
\mu_{11}^* = \left[\begin{array}{c|c|c|c}
0 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\hline
\mathbf{0} & \mathbf{I}_k & \mathbf{0} & \mathbf{0} \\
\hline
\mathbf{0} & \mathbf{0} & 4\,w_{ii}\,\mathbf{I}_k & \mathbf{0} \\
\hline
\mathbf{0} & \mathbf{0} & \mathbf{0} & 2\,w_{ii}\,\mathbf{I}_{\binom{k}{2}}
\end{array}\right].
$$

The **V** component of the MSE of response criteria is obtained by taking the *trace* of the product of the μ_{11} and Var [$\underline{b}_1$] matrices. For the CCDs, BBDs, and FACs considered here, the **V** component of **J** for the **IR** strategy becomes

$$\mathbf{V}_{IR} = \frac{a - 2k\,w_{ii}\,[ii]}{D} + \frac{k\,w_{ii}}{[ii]} + \frac{b\,k\,w_{iiii} + c\,k\,(k-1)\,w_{iijj}}{D_2} + \frac{k\,(k-1)\,w_{iijj}}{2\,[iijj]}$$

where a, b, c, D, and D_2 are components of the elements of the $(X'_1\,X_1)^{-1}$ matrix (defined on page 384).

For the **CR** strategy, two forms of the **V** component are relevant. For the *pure* **CR** strategy, in which there are no replicated center runs, the **V** component of **J** becomes

$$\mathbf{V}_{pure\ CR} = (1-\rho_+)\,\mathbf{V}_{IR} + N\,\rho_+$$

and for the *modified* **CR** strategy, in which there are replicated center runs, **V** becomes

$$\mathbf{V}_{mod.\ CR} = (1-\rho_+)\,\mathbf{V}_{IR} + N\,\rho_+\,(s^2 + 2k\,w_{ii}\,s\,t + k\,t^2\,g) + N\,\rho_+\,(q^2 + 2k\,w_{ii}\,r\,q + k\,r^2\,g)$$

where $g = w_{iiii} + (k-1)\,w_{iijj}$ is defined to simplify notation, s and t are elements of the $Z'\,\underline{u}\,\underline{u}'\,Z$ matrix (defined on page 385) and q and r are elements of the $Z'\,U\,Z$ matrix (defined on page 387).

For the **AR** strategy, four forms of the **V** component are relevant. For the *pure* **AR** strategy, in which the design partitions into two orthogonal blocks and there are no replicated center runs within blocks, the **V** component of **J** for the CCDs and $k = 5,\,7$ BBDs becomes

$$\mathbf{V}_{pure\ AR} = (1-\rho_+)\,\mathbf{V}_{IR} + \tfrac{1}{2}\,N\,(\rho_+ - \rho_-) + \tfrac{1}{2}\,(\rho_+ + \rho_-)\,(N_{1a} - N_{1b})^2\,/\,N$$

where N_{1a} and N_{1b} are the number of design points in blocks 1 and 2 of the partitioned design (defined on page 386).

For the designs which partition into two orthogonal blocks, but have replicated center runs within blocks, the **V** component for the *modified* **AR** strategy becomes

$$\mathbf{V}_{mod.\,AR} = (1-\rho_+)\,\mathbf{V}_{IR} + \tfrac{1}{2}\,N\,(\rho_+ - \rho_-)\,(s^2 + 2\,k\,w_{ii}\,s\,t + k\,t^2\,g) \\ + N\,\rho_+\,(q^2 + 2\,k\,w_{ii}\,r\,q + k\,r^2\,g) + \tfrac{1}{2}\,N\,(\rho_+ + \rho_-)\,(p^2 + 2\,k\,w_{ii}\,m\,p + k\,m^2\,g)$$

where p and m are elements of the $\mathbf{Z}'\,\underline{y}\,\underline{y}'\,\mathbf{Z}$ matrix (defined on page 386).

The third form of the **V** component under the **AR** strategy applies to the FACs, which partition into three orthogonal blocks and have no replicated center runs. The **V** component of **J** for the 3^k designs becomes

$$\mathbf{V}_{AR,\,3fac} = (1 - h\,\rho_+)\,\mathbf{V}_{IR} + \tfrac{1}{2}\,N\,s^2\,(\rho_+ - \rho_-)$$

where $h = 2\,/\,$(number of blocks) is a term which takes the additional design block into account. The t, m, and p terms are equal to zero because the block sizes are equal, and the q and r terms are not relevant here.

The fourth form of the **V** component under the **AR** strategy applies to the $k=4$ BBD, which partitions into three orthogonal blocks, with all of the center runs in the 3rd block. The **V** component of **J** for the $k=4$ BBD becomes

$$\mathbf{V}_{AR,\,bbd4} = (1 - \rho_+)\,\mathbf{V}_{IR} + \tfrac{1}{2}\,N\,k\,t^2\,g\,(\rho_+ - \rho_-) \\ + \rho_+ \left[\frac{a - 8\,w_{ii}\,[ii]}{D} + \frac{4\,w_{ii}}{3\,[ii]} + \frac{2\,w_{iijj}}{[iijj]} + 4\,N\,(e\,w_{iiii} + (2f+e)\,w_{iijj}) \right]$$

where e and f are components of the $Z' W Z$ matrix (defined on page 388). The s , m , and p terms are equal to zero and the q and r terms are not relevant to the $k=4$ BBD because it partitions into three orthogonal blocks.

The V^* component of the MSE of slope criteria is obtained by taking the *trace* of the product of the $\mu_{11}{}^*$ and Var [$\underline{b}_1$] matrices. For the CCDs, BBDs, and FACs considered in this research, the V^* component of J^* for the IR strategy becomes

$$V^*_{IR} = \frac{k}{[ii]} + \frac{4\,b\,k\,w_{ii}}{D_2} + \frac{k\,(k-1)\,w_{ii}}{[iijj]} .$$

For the CR strategy, two forms of the V^* component are relevant. For the *pure* CR strategy, in which there are no replicated center runs, the V^* component of J^* becomes

$$V^*_{pure\ CR} = (1-\rho_+)\, V^*_{IR}$$

and for the *modified* CR strategy, in which there are replicated center runs, V^* becomes

$$V^*_{mod.\ CR} = (1-\rho_+)\, V^*_{IR} + 4\,k\,w_{ii}\,N\,\rho_+\,[\,t^2 + r^2\,] .$$

For the AR strategy, four forms of the V^* component are relevant. For the *pure* AR strategy, in which the design partitions into two orthogonal blocks and there are no replicated center runs within blocks, the V^* component of J^* for the CCDs and $k=5, 7$ BBDs becomes

$$V^*_{pure\ AR} = (1-\rho_+)\, V^*_{IR}$$

which is the same result as previously given for the *pure* CR strategy.

For the designs which partition into two orthogonal blocks (CCDs and $k=5,\ 7$ BBDs), but have replicated center runs within blocks, the **V*** component for the *modified* AR strategy becomes

$$\mathbf{V}^*_{mod.\,AR} = (1-\rho_+)\,\mathbf{V}^*_{IR} + 4\,k\,w_{ii}\left[\ \tfrac{1}{2}\,N\,t^2\,(\rho_+-\rho_-) + N\,r^2\,\rho_+ + \tfrac{1}{2}\,N\,m^2\,(\rho_+ + \rho_-)\ \right].$$

The third form of the **V*** component under the AR strategy applies to the FACs, which have no replicated center runs and partition into three orthogonal blocks of equal size. The **V*** component of **J*** for the 3^k designs becomes

$$\mathbf{V}^*_{AR,\,3fac} = (\,1 - h\,\rho_+)\,\mathbf{V}^*_{IR}\ .$$

The fourth form of the **V*** component under the AR strategy applies to the $k=4$ BBD, which partitions into three orthogonal blocks with all of the center runs in the 3rd block. The **V*** component of **J*** for the $k=4$ BBD becomes

$$\mathbf{V}^*_{AR,\,bbd4} = (\,1-\rho_+)\,\mathbf{V}^*_{IR} + 8\,w_{ii}\,N\,t^2\,(\rho_+-\rho_-) + \rho_+\left[\ \frac{4}{3\,[ii]} + \frac{4\,w_{ii}}{[iijj]} + 16\,N\,w_{ii}\,e\ \right].$$

In order to compute **V** and **V*** for the correlation induction strategies, the a, b, c, D, D_2, g, h, s, t, q, r, m, p, Δ, N, N_{1a} , N_{1b}, and N_2 constants (defined in Appendix M on pages 381-388) must be specified. The constants which involve [iiii] and/or [iijj] can be redefined using the following notation:

$$\theta = \frac{[iiii]}{[ii]} \qquad\qquad r = \frac{[iiii]}{[iijj]}$$

resulting in the following equations for [iiii] and [iijj]:

$$[iiii] = \theta\,[ii] \qquad\qquad [iijj] = \frac{\theta\,[ii]}{r}\ .$$

Through the use of the θ and r design moment ratios, **V** and **V*** can be computed without specification of the values of [iiii] and [iijj]. The constants involving [iiii] and/or [iijj] can now be redefined (originally defined in Appendix M) using the θ and r ratios, as follows:

$$a = \frac{\theta\,[ii]\,(r+k-1)}{r}$$

$$b = \frac{\theta\,[ii]\,(r+k-2)}{r} - (k-1)\,[ii]^2$$

$$c = [ii]\left[\frac{r\,[ii]-\theta}{r}\right]$$

$$D = a - k\,[ii]^2$$

$$D_2 = \frac{D\,\theta\,[ii]\,(r-1)}{r} .$$

The second order designs result in *singular* $(X'_1 X_1)$ matrices if value of *D* is equal to zero, or equivalently, if the value of [iijj] is equal to

$$[iijj] = \frac{[ii]}{k-1}\,(k\,[ii] - \theta) .$$

Each of the second order designs are checked for singularity of the $(X'_1 X_1)$ matrix by comparing the actual values of [iijj] to the "singular" values.

The formulas given in this appendix enable computation of $\mathbf{V}_{IR}$, $\mathbf{V}_{CR}$, $\mathbf{V}_{AR}$, $\mathbf{V}^*_{IR}$, $\mathbf{V}^*_{CR}$, and $\mathbf{V}^*_{AR}$ for the CCDs, BBDs, FACs, and fractional CCDs and 3^k designs with $k \geq 6$ factors (enabling the use of a 6-way interaction term as the defining contrast). Due to the complications involved in inverting the $(X'_1 X_1)^{-1}$ matrix, no computational formulas have been given for the SCDs. However, the **V** and **V*** components for the SCDs can be approximated using the formulas for the designs which partition into two orthogonal blocks, but for exact computations, direct matrix algebra on the specific SCD design plan must be used.

Appendix O

B and B* Equations in the Second Order Case

The bias components of **J** and **J*** involve the alias matrix, A, and the region matrices, μ_{11}, μ_{11}^*, μ_{21}, μ_{21}^*, μ_{22}, and μ_{22}^* . For the MSE of response and MSE of slope criteria, respectively, the bias components can be written as

$$\mathbf{B} = \frac{N}{\sigma_2}\ \underline{\beta}'_2\ \{A'\mu_{11}A - 2\mu_{21}A + \mu_{22}\}\ \underline{\beta}_2$$

$$\mathbf{B}^* = \frac{N}{\sigma^2}\ \underline{\beta}'_2\ \{A'\overset{*}{\mu}_{11}A - 2\overset{*}{\mu}_{21}A + \overset{*}{\mu}_{22}\}\ \underline{\beta}_2\ .$$

Because the CCDs, BBDs, and FACs considered here have odd design moments through order five equal to zero, the equations for **B** and **B*** can be simplified to scalar equations in the design and region moments. For the second order designs examined in this research, the $(p_2 \times p_1)$ alias matrix under ordinary least squares estimation can be written as

$$
A = \left[\begin{array}{c|c}
\mathbf{0}' & \\
\hline
G_A & \mathbf{0} \\
\hline
\mathbf{0} &
\end{array}\right]
$$

where G_A is a $(k \times k^2)$ block diagonal matrix with k identical blocks of the $(1 \times k)$ g_A vector

$$g_A = \tfrac{1}{[\text{illegible}]} \left([iiii] , [iijj] , \ldots\ldots , [iijj] \right) .$$

Under weighted least squares estimation, the alias matrix is the same for the CR strategy, but for the AR strategy, the alias matrix is the same only for the designs which partition into two orthogonal blocks (CCDs and [illegible] BBDs). For the $k=4$ BBD and the FACs, the weighted least squares alias matrix for the AR strategy is considerably more complicated.

In order to derive equations for B and B*, the following component matrices are needed:

1. $A' \mu_{11} A$, $A' \mu_{11}^* A$
2. $\mu_{21} A$, $\mu_{21}^* A$
3. μ_{22} , and μ_{22}^*

The remainder of this appendix develops these component matrices for the second order designs examined in this research. The components of **B** are presented first, followed by the components of **B***.

The μ_{21} and μ_{22} matrices needed for the **B** component of **J** have the same form as the M_{21} and M_{22} matrices shown in Appendix J (pages 348-352) for the general case of fitting a second order model, except that the design moments are replaced with region moments. For the second order designs considered, the $(p_2 \times p_1)$ μ_{21} matrix becomes

$$\mu_{21} = \left[\begin{array}{c|c|c} \underline{0}_{k^2} & G_{21} & 0 \\ \hline \multicolumn{3}{c}{0} \end{array}\right]$$

where G_{21} is a $(k^2 \times k)$ block diagonal matrix, with k identical blocks of the $(k \times 1)$ $\underline{g}_{21}$ vector

$$\underline{g}_{21}' = [\, w_{iiii} \,,\, w_{iijj} \,,\, \ldots\ldots \,,\, w_{iijj} \,]$$

and the $(p_2 \times p_2)$ μ_{22} matrix becomes

$$\mu_{22} = \left[\begin{array}{c|c} G_{22} & 0 \\ \hline 0 & w_{iijjkk}\, \mathbf{I}_{\binom{k}{3}} \end{array}\right]$$

where G_{22} is a $(k^2 \times k^2)$ block diagonal matrix, with k identical blocks of the $(k \times k)$ H_{22} matrix

$$H_{22} = (\, w_{iiiiii} - w_{iiiijj} \,)\, \mathbf{I}_k + w_{iiiijj}\, \underline{1}_k \underline{1}_k' \,.$$

In order to obtain the first component matrix of **B**, the μ_{11} matrix shown in Appendix N (page 390) is pre- and post-multiplied by the alias matrix, yielding the $(p_2 \times p_2)$ $A' \mu_{11} A$ matrix

$$
A' \mu_{11} A = \left[\begin{array}{c|c} G_{A11A} & 0 \\ \hline 0 & 0 \end{array}\right]
$$

where G_{A11A} is a $(k^2 \times k^2)$ block diagonal matrix, with k identical blocks of the $(k \times k)$ H_{A11A} matrix

$$
H_{A11A} = \frac{w_{ii}}{[ii]^2} \left[\begin{array}{c|c} [iiii]^2 & [iiii]\,[iijj]\,\mathbf{1}'_{k-1} \\ \hline [iiii]\,[iijj]\,\mathbf{1}_{k-1} & [iijj]^2\,\mathbf{1}_{k-1}\mathbf{1}'_{k-1} \end{array}\right].
$$

In order to obtain the second component matrix of **B**, the μ_{21} matrix is post-multiplied by the alias matrix, yielding the $(p_2 \times p_2)$ $\mu_{21} A$ matrix

$$
\mu_{21} A = \left[\begin{array}{c|c} G_{21A} & 0 \\ \hline 0 & 0 \end{array}\right]
$$

where G_{21A} is a $(k^2 \times k^2)$ block diagonal matrix, with k identical blocks of the $(k \times k)$ H_{21A} matrix

$$
H_{21A} = \frac{1}{[ii]} \left[\begin{array}{c|c} w_{iiii}\,[iiii] & w_{iiii}\,[iijj]\,\mathbf{1}'_{k-1} \\ \hline w_{iijj}\,[iiii]\,\mathbf{1}_{k-1} & w_{iijj}\,[iijj]\,\mathbf{1}_{k-1}\mathbf{1}_{k-1}' \end{array} \right].
$$

The $A'\mu_{11}A$, $\mu_{21}A$, and μ_{22} matrices are summed to form a matrix, which upon pre- and post-multiplication by $\underline{\beta}_2 \sqrt{N/\sigma^2}$, yields the $\mathbf{B}$ component of $\mathbf{J}$

$$
\mathbf{B} = \frac{N}{\sigma^2}\, \underline{\beta}_2' \left[\begin{array}{c|c} G_B & \mathbf{0} \\ \hline \mathbf{0} & w_{iijjkk}\, \mathbf{I}_{\binom{k}{3}} \end{array} \right] \underline{\beta}_2
$$

where G_B is a $(k^2 \times k^2)$ block diagonal matrix, with k identical blocks of the $(k \times k)$ H_B matrix

$$
H_B = \left[\begin{array}{c|c} h_1 & h_2\, \mathbf{1}'_{k-1} \\ \hline h_3\, \mathbf{1}_{k-1} & (h_4 - h_5)\, \mathbf{I}_{k-1} + h_5\, \mathbf{1}_{k-1}\mathbf{1}_{k-1}' \end{array} \right]
$$

where
$$
h_1 = \frac{w_{ii}\,[iiii]^2}{[ii]^2} - \frac{2\, w_{iiii}\,[iiii]}{[ii]} + w_{iiiiii}
$$

$$
h_2 = \frac{w_{ii}\,[iiii]\,[iijj]}{[ii]^2} - \frac{2\, w_{iiii}\,[iijj]}{[ii]} + w_{iiiijj}
$$

$$
h_3 = \frac{w_{ii}\,[iiii]\,[iijj]}{[ii]^2} - \frac{2\, w_{iijj}\,[iiii]}{[ii]} + w_{iiiijj}
$$

$$
h_4 = \frac{w_{ii}\,[iijj]^2}{[ii]^2} - \frac{2\, w_{iijj}\,[iijj]}{[ii]} + w_{iiiiii}
$$

$$h_5 = \frac{w_{ii}\,[iijj]^2}{[ii]^2} - \frac{2\,w_{iijj}\,[iijj]}{[ii]} + w_{iiiijj}\ .$$

The **B** component of **J** can now be written in equation form as

$$\mathbf{B} = \frac{N}{\sigma^2}\left[\{h_1 + (k-1)\,h_3\}\sum_{i=1}^{k}\beta_{iii}^{\,2} + \{h_2 + h_4 + (k-2)\,h_5\}\sum_{i \neq j}\sum\beta_{ijj}^{\,2} + w_{iijjkk}\sum_{i \neq j \neq k}\sum\sum\beta_{ijk}^{\,2}\right]$$

$$= \phi_1\,\Theta_1 + \phi_2\,\Theta_2 + \phi_3\,\Theta_3$$

where $\phi_1 = h_1 + (k-1)\,h_3$ $\qquad \Theta_1 = \frac{N}{\sigma^2}\sum_{i=1}^{k}\beta_{iii}^{\,2}$

$\phi_2 = h_2 + h_4 + (k-2)\,h_5$ $\qquad \Theta_2 = \frac{N}{\sigma^2}\sum_{i \neq j}\sum\beta_{ijj}^{\,2}$

$\phi_3 = w_{iijjkk}$ $\qquad \Theta_3 = \frac{N}{\sigma^2}\sum_{i \neq j \neq k}\sum\sum\beta_{ijk}^{\,2}$.

For second order designs with odd design moments through order five equal to zero, and even order moments through order six equal to known constants, the **B** component of **J** can be computed for specified values of the unknown Θ_1, Θ_2, and Θ_3 parameters.

To simplify notation, θ is defined as the ratio of the pure 4th to pure 2nd order design moment, and r is defined as the ratio of the pure 4th to mixed 4th order design moment, yielding θ/r as the ratio of the mixed 4th to pure 2nd order design moment

$$\theta = \frac{[iiii]}{[ii]}\ , \qquad r = \frac{[iiii]}{[iijj]}\ , \text{ and } \qquad \frac{\theta}{r} = \frac{[iijj]}{[ii]}\ ,$$

which leads to the following equations for ϕ_1 and ϕ_2

$$\phi_1 = \frac{w_{ii}\,\theta^2\,(r+k-1)}{r} - 2\theta\,(w_{iiii} + (k-1)\,w_{iijj}) + (w_{iiiiii} + (k-1)\,w_{iiiijj})$$

$$\phi_2 = \frac{w_{ii}\,\theta^2\,(r+k-1)}{r^2} - \frac{2\theta\,(w_{iiii} + (k-1)\,w_{iijj})}{r} + (w_{iiiiii} + (k-1)\,w_{iiiijj})\ .$$

For a spherical region of interest, the ϕ_1, ϕ_2, and ϕ_3 coefficients become

$$\phi_1 = \frac{\theta^2\,(r+k-1)}{r\,(k+2)} - \frac{2\theta}{k+4} + \frac{3}{(k+2)\,(k+6)}$$

$$\phi_2 = \frac{\theta^2\,(r+k-1)}{r^2\,(k+2)} - \frac{2\theta}{r\,(k+4)} + \frac{3}{(k+2)\,(k+6)}$$

$$\phi_3 = \frac{1}{(k+2)\,(k+4)\,(k+6)}$$

and for *spherically-rotatable* designs, which have $r = 3$, the above equations can be further simplified.

For a cuboidal region of interest, the ϕ_1, ϕ_2, and ϕ_3 coefficients become

$$\phi_1 = \frac{\theta^2\,(r+k-1)}{3\,r} - \frac{2\theta\,(5k+4)}{45} + \frac{7k+8}{105}$$

$$\phi_2 = \frac{\theta^2\,(r+k-1)}{3\,r^2} - \frac{2\theta\,(5k+4)}{45\,r} + \frac{7k+8}{105}$$

$$\phi_3 = \frac{1}{27}$$

and for *cuboidally-rotatable* designs, which have $r = {}^9/_5$, the above equations can be further simplified.

The $\mu_{21}{}^*$ and $\mu_{22}{}^*$ matrices needed for the $\mathbf{B}^*$ component of $\mathbf{J}^*$ are shown in Appendix K (pages 353-358) for the general case of fitting a second order model. For the second order designs used in this research, the $(p_2 \times p_1)$ $\mu_{21}{}^*$ matrix becomes

$$\mu_{21}{}^* = \left[\begin{array}{c|c|c} \underset{\sim}{0}_{k^2} & G_{21\bullet} & 0 \\ \hline \multicolumn{3}{c}{0} \end{array}\right]$$

where $G_{21\bullet}$ is a $(k^2 \times k)$ block diagonal matrix, with k identical blocks of the $(k \times 1)$ $\underset{\sim}{g}_{21\bullet}$ vector

$$\underset{\sim}{g}_{21\bullet}{}' = [\, 3\,w_{ii} \,,\, w_{ii} \,,\, \ldots\ldots \,,\, w_{ii} \,]$$

and the $(p_2 \times p_2)$ $\mu_{22}{}^*$ matrix becomes

$$\mu_{22}{}^* = \left[\begin{array}{c|c} G_{22\bullet} & 0 \\ \hline 0 & 3\,w_{iijj}\,\mathbf{I}_{\binom{k}{3}} \end{array}\right]$$

where $G_{22\bullet}$ is a $(k^2 \times k^2)$ block diagonal matrix, with k identical blocks of the $(k \times k)$ $H_{22\bullet}$ matrix

$$H_{22\bullet} = \left[\begin{array}{c|c} 9\,w_{iiii} & 3\,w_{iiii}\,\underset{\sim}{1}'_{k-1} \\ \hline 3w_{iiii}\,\underset{\sim}{1}_{k-1} & (\,w_{iiii} + 3\,w_{iijj}\,)\,\mathbf{I}_{k-1} + w_{iijj}\,\underset{\sim}{1}_{k-1}\underset{\sim}{1}_{k-1}{}' \end{array}\right].$$

In order to obtain the first component matrix of $\mathbf{B}^*$, the μ_{11}^* matrix shown in Appendix N (page 390) is pre- and post-multiplied by the alias matrix, yielding the $(p_2 \times p_2)$ $A' \mu_{11}^* A$ matrix

$$
A' \mu_{11}^* A = \left[\begin{array}{c|c} G_{A11 \cdot A} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right]
$$

where $G_{A11 \cdot A}$ is a $(k^2 \times k^2)$ block diagonal matrix, with k identical blocks of the $(k \times k)$ $H_{A11 \cdot A}$ matrix

$$
H_{A11 \cdot A} = \frac{1}{[ii]^2} \left[\begin{array}{c|c} [iiii]^2 & [iiii]\,[iijj]\,\mathbf{1}'_{k-1} \\ \hline [iiii]\,[iijj]\,\mathbf{1}_{k-1} & [iijj]^2\,\mathbf{1}_{k-1}\mathbf{1}'_{k-1} \end{array} \right].
$$

In order to obtain the second component matrix of $\mathbf{B}^*$, the μ_{21}^* matrix is post-multiplied by the alias matrix, yielding the $(p_2 \times p_2)$ $\mu_{21}^* A$ matrix

$$
\mu_{21}^* A = \left[\begin{array}{c|c} G_{21 \cdot A} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right]
$$

where $G_{21 \cdot A}$ is a $(k^2 \times k^2)$ block diagonal matrix, with k identical blocks of the $(k \times k)$ $H_{21 \cdot A}$ matrix

$$
H_{21 \cdot A} = \frac{1}{[ii]} \left[\begin{array}{c|c} 3\,w_{ii}\,[iiii] & 3\,w_{ii}\,[iijj]\,\mathbf{1}'_{k-1} \\ \hline w_{ii}\,[iiii]\,\mathbf{1}_{k-1} & w_{ii}\,[iijj]\,\mathbf{1}_{k-1}\mathbf{1}_{k-1}' \\ & \end{array} \right].
$$

The $A'\mu_{11}{}^{*}A$, $\mu_{21}{}^{*}A$, and $\mu_{22}{}^{*}$ matrices are summed to form a matrix, which upon pre- and post-multiplication by $\underline{\beta}_2 \sqrt{N/\sigma^2}$, yields the $\mathbf{B}^{*}$ component of $\mathbf{J}^{*}$

$$
\mathbf{B}^{*} = \frac{N}{\sigma^2}\,\underline{\beta}_2{}' \left[\begin{array}{c|c} G_{B^{*}} & 0 \\ \hline 0 & 3\,w_{iijj}\,\mathbf{I}_{\binom{k}{3}} \end{array} \right] \underline{\beta}_2
$$

where $G_{B^{*}}$ is a $(k^2{\times}k^2)$ block diagonal matrix, with k identical blocks of the $(k{\times}k)$ $H_{B^{*}}$ matrix

$$
H_{B^{*}} = \left[\begin{array}{c|c} h_1^{*} & h_2^{*}\;\mathbf{1}'_{k-1} \\ \hline h_3^{*}\;\mathbf{1}_{k-1} & (h_4^{*} - h_5^{*})\,\mathbf{I}_{k-1} + h_5^{*}\;\mathbf{1}_{k-1}\mathbf{1}_{k-1}' \end{array} \right]
$$

$$
\text{where}\quad h_1^{*} = \frac{[iiii]^2}{[ii]^2} - \frac{6\,w_{ii}\,[iiii]}{[ii]} + 9\,w_{iiii}
$$

$$
h_2^{*} = \frac{[iiii]\,[iijj]}{[ii]^2} - \frac{6\,w_{ii}\,[iijj]}{[ii]} + 3\,w_{iiii}
$$

$$
h_3^{*} = \frac{[iiii]\,[iijj]}{[ii]^2} - \frac{2\,w_{ii}\,[iiii]}{[ii]} + 3\,w_{iiii}
$$

$$
h_4^{*} = \frac{[iijj]^2}{[ii]^2} - \frac{2\,w_{ii}\,[iijj]}{[ii]} + w_{iiii} + 4\,w_{iijj}
$$

$$h_5^* = \frac{[iijj]^2}{[ii]^2} - \frac{2\,w_{ii}\,[iijj]}{[ii]} + w_{iijj}\,.$$

The B^* component of J^* can now be written in equation form as

$$B^* = \frac{N}{\sigma^2}\left[\{h_1^* + (k-1)\,h_3^*\}\sum_{i=1}^{k}\beta_{iii}{}^2 + \{h_2^* + h_4^* + (k-2)\,h_5^*\}\sum_{i \neq j}\sum \beta_{iij}{}^2 + 3\,w_{iijj}\sum_{i \neq j \neq k}\sum\sum \beta_{ijk}{}^2\right]$$

$$= \phi_1^*\,\Theta_1 + \phi_2^*\,\Theta_2 + \phi_3^*\,\Theta_3$$

where $\phi_1^* = h_1^* + (k-1)\,h_3^*$ $\qquad \Theta_1 = \frac{N}{\sigma^2}\sum_{i=1}^{k}\beta_{iii}{}^2$

$\phi_2^* = h_2^* + h_4^* + (k-2)\,h_5^*$ $\qquad \Theta_2 = \frac{N}{\sigma^2}\sum_{i \neq j}\sum \beta_{iij}{}^2$

$\phi_3^* = 3\,w_{iijj}$ $\qquad \Theta_3 = \frac{N}{\sigma^2}\sum_{i \neq j \neq k}\sum\sum \beta_{ijk}{}^2$.

For second order designs with odd design moments through order five equal to zero, and even order moments through order six equal to known constants, the B^* component of J^* can be computed for specified values of the unknown Θ_1, Θ_2, and Θ_3 parameters.

The equations for ϕ_1^* and ϕ_2^* can be simplified through the use of the $\theta = [iiii]/[ii]$ and $r = [iiii]/[iijj]$ notation, yielding

$$\phi_1^* = \frac{\theta^2\,(r+k-1)}{r} - 2\theta\,w_{ii}\,(k+2) + 3\,w_{iiii}\,(k+2)$$

$$\phi_2^* = \frac{\theta^2\,(r+k-1)}{r^2} - \frac{2\theta\,w_{ii}\,(k+2)}{r} + (4\,w_{iiii} + (k+2)\,w_{iijj})\,.$$

For a spherical region of interest, the ϕ_1^*, ϕ_2^*, and ϕ_3^* coefficients simplify to

$$\phi_1^* = \frac{\theta^2\,(r+k-1)}{r} - 2\theta + \frac{9}{k+4}$$

$$\phi_2^* = \frac{\theta^2 (r+k-1)}{r^2} - \frac{2\theta}{r} + \frac{k+14}{(k+2)(k+4)}$$

$$\phi_3^* = \frac{3}{(k+2)(k+4)}$$

and for *spherically-rotatable* designs, which have $r=3$, the above equations can be further simplified.

For a cuboidal region of interest, the ϕ_1^*, ϕ_2^*, and ϕ_3^* coefficients simplify to

$$\phi_1^* = \frac{\theta^2 (r+k-1)}{r} - \frac{2\theta (k+2)}{3} + \frac{3(k+2)}{5}$$

$$\phi_2^* = \frac{\theta^2 (r+k-1)}{r^2} - \frac{2\theta (k+2)}{3r} + \frac{5k+46}{45}$$

$$\phi_3^* = \frac{1}{3}$$

and for *cuboidally-rotatable* designs, which have $r = 9/5$, the above equations can be further simplified.

The formulas presented in this appendix enable computation of **B** and **B*** for the CCDs, BBDs, and FACs. Due to the complicatedness of the alias matrix, no computational formulas have been given for the SCDs. For these designs, the **B** and **B*** components cannot be minimized analytically.

Appendix P

Min-V | Min-B and Min-V* | Min-B* Second Order Designs

In the first order case, the **J**- and **J***-optimal designs were determined by taking the partial derivatives of **J** and **J*** with respect to the pure second order design moment, [ii]. In the second order case, **J** and **J*** are functions of the pure and mixed fourth order design moments, [iiii] and [iijj], as well as [ii]. Therefore, in the second order case it is not possible to minimize **J** and **J*** with respect to [ii] alone. For the second order designs examined in this research, the bias components are first minimized by taking the partial derivatives of **B** and **B*** with respect to $\theta = [iiii]/[ii]$, yielding the **B**- and **B***-optimal values of θ. Utilizing the Min-**B** and Min-**B*** values of θ, the partial derivatives of **V** and **V*** are then taken with respect to [ii], resulting in Min-**V** | Min-**B** and Min-**V*** | Min-**B*** designs. The **B**- and **B***-optimal values of θ are obtained through appropriate choices of the scaling factor, g, and the **V**- and **V***-optimal values of [ii] are obtained (or *nearly* obtained) with appropriate choices for the number of center runs, N_c.

For the second order designs examined in this research, the bias and variance components of **J** and **J*** are functions of the following parameters and constants:

$$\mathbf{B} = f\,(\theta, r, k, w_{ii}, w_{iiii}, w_{iijj}, w_{iiiiii}, w_{iiiijj}, w_{iijjkk}, \Theta_1, \Theta_2, \Theta_3)$$

$$\mathbf{B}^* = f\,(\theta, r, k, w_{ii}, w_{iiii}, \Theta_1, \Theta_2, \Theta_3)$$

$$\mathbf{V} = f\,(\theta, [ii], r, k, w_{ii}, w_{iiii}, w_{iijj}, \rho_+, \rho_-, N, N_{1a}, N_{1b})$$

$$\mathbf{V}^* = f\,(\theta, [ii], r, k, w_{ii}, \rho_+, \rho_-, N, N_{1a}, N_{1b})\ .$$

The **B** and **B*** components are functions of constant terms, with the exception of θ, Θ_1, Θ_2, and Θ_3. The Θ parameters disappear when the partial derivatives of **B** and **B*** are taken with respect to θ, yielding the optimal values of θ as functions of constant terms only. The **V** and **V*** components are functions of constant terms, with the exception of θ and [ii]. Inserting the **B**- and **B***-optimal values of θ into **V** and **V***, and taking the partial derivatives with respect to [ii], yields the optimal values of [ii] as functions of constant terms only.

In order to minimize the **B** and **B*** components of **J** and **J***, the partial derivatives with respect to θ are set equal to zero, for both spherical and cuboidal regions of interest. Appendix O (pages 396-407) illustrates the **B** component of **J**, and when set equal to zero, the partial derivative of **B** with respect to θ becomes

$$\frac{\partial \mathbf{B}}{\partial \theta} = \frac{\partial}{\partial \theta}\left[\frac{w_{ii}\,\theta^2\,(r+k-1)\,\Theta_1}{r} - 2\theta\,(w_{iiii} + (k-1)\,w_{iijj})\,\Theta_1 + (w_{iiiiii} + (k-1)\,w_{iiiijj})\,\Theta_1\right]$$

$$+ \frac{\partial}{\partial \theta}\left[\frac{w_{ii}\,\theta^2\,(r+k-1)\,\Theta_2}{r^2} - \frac{2\theta\,(w_{iiii} + (k-1)\,w_{iijj})\,\Theta_2}{r} + (w_{iiiiii} + (k-1)\,w_{iiiijj})\,\Theta_2\right]$$

$$+ \frac{\partial}{\partial \theta}\left[w_{iijjkk}\,\Theta_3\right] = 0$$

$$0 = \frac{2\,w_{ii}\,\theta\,(r+k-1)}{r}\left[\Theta_1 + \frac{\Theta_2}{r}\right] - 2\,(w_{iiii} + (k-1)\,w_{iijj})\left[\Theta_1 + \frac{\Theta_2}{r}\right].$$

Solving the partial derivative equation for θ, the **B**-optimal value of θ becomes

$$\theta = \frac{r\,(w_{iiii} + (k-1)\,w_{iijj})}{w_{ii}\,(r+k-1)}$$

and for a spherical region of interest, the equation for **B**-optimal value of θ can be written as

$$\theta = \frac{r\,(k+2)}{(r+k-1)\,(k+4)}$$

and for a cuboidal region of interest, the Min-**B** value of θ is computed as

$$\theta = \frac{r\,(5k+4)}{15\,(r+k-1)}\ .$$

The **B*** component of **J*** is also shown in Appendix O. When set equal to zero, the partial derivative of **B*** with respect to θ becomes

$$\frac{\partial \mathbf{B}^*}{\partial\theta} = \frac{\partial}{\partial\theta}\left[\frac{\theta^2\,(r+k-1)\,\Theta_1}{r} - 2\,\theta\,w_{ii}\,(k+2)\,\Theta_1 + 3\,w_{iiii}\,(k+2)\,\Theta_1\right]$$

$$+ \frac{\partial}{\partial\theta}\left[\frac{\theta^2\,(r+k-1)\,\Theta_2}{r^2} - \frac{2\,\theta\,w_{ii}\,(k+2)\,\Theta_2}{r} + 4\,w_{iiii}\,\Theta_2 + w_{iijj}\,(k+2)\,\Theta_2\right]$$

$$+ \frac{\partial}{\partial\theta}[\,3\,w_{iijj}\,\Theta_3\,] = 0$$

$$0 = \frac{2\,\theta\,(r+k-1)}{r}\left[\Theta_1 + \frac{\Theta_2}{r}\right] - 2\,w_{ii}\,(k+2)\left[\Theta_1 + \frac{\Theta_2}{r}\right]$$

Solving the partial derivative equation for θ, the **B***-optimal value of θ becomes

$$\theta = \frac{r\, w_{ii}\,(k+2)}{(r+k-1)}$$

and for a spherical region of interest, the **B***-optimal value of θ is computed as

$$\theta = \frac{r}{(r+k-1)}$$

and for a cuboidal region of interest, the Min-**B*** value of θ becomes

$$\theta = \frac{r\,(k+2)}{3\,(r+k-1)}\ .$$

The second partial derivatives of **B** and **B*** with respect to θ are both positive quantities and, therefore, the optimal values of θ minimize **B** and **B***.

For the CCDs examined in this research, θ can be written as

$$\theta = \frac{(F\,g^4 + 2\,\alpha^4\,g^4)\,/\,N}{(F\,g^2 + 2\,\alpha^2\,g^2)\,/\,N} = \frac{g^2\,(F + 2\,\alpha^4)}{(F + 2\,\alpha^2)}\ .$$

Solving for the scaling factor needed to achieve the optimal value of θ, the optimal value of g for the CCDs becomes

$$g = \sqrt{\frac{\theta\,(F + 2\,\alpha^2)}{(F + 2\,\alpha^4)}}\ .$$

For the BBDs and FACs, θ can be written as

$$\theta = \frac{f\,F\,g^4\,/\,N}{f\,F\,g^2\,/\,N}$$

where f is the fraction of factorial design points in which a particular design variable equal to zero ($f = 2/3$ for the FACs, $f = 1/2$ for the $k=4$ BBD, $f = 2/5$ for the $k=5$ BBD, and $f = 3/7$ for the $k=7$

BBD). Solving for the scaling factor needed to achieve the optimal value of θ, the optimal value of g for the BBDs and FACs becomes

$$g = \sqrt{\theta} \ .$$

Table P-1 on page 418 shows the Min-**B** and Min-**B*** values of θ for each of the second order designs and Table P-2 on page 419 shows the values of the scaling factor, g, needed to achieve the optimal values of θ.

In order to minimize the **V** and **V*** components of **J** and **J***, given the Min-**B** and Min-**B*** values of θ, the partial derivatives with respect to [ii] are set equal to zero. Appendix M (pages 381-388) illustrates the **V** and **V*** components for the IR, CR, and AR strategies. For the *pure* CR and AR strategies, the optimal values of [ii] are the same as those for the IR strategy because the additional terms in the variance components do not involve [ii]. The optimal values of [ii] are obtained through appropriate choices for the number of center runs (restricted to be integer-valued), and therefore the optimal values of [ii] cannot be exactly achieved. For the *modified* CR and AR strategies, the $\mathbf{V}_{IR}$ and $\mathbf{V}^*_{IR}$ optimal values of [ii] are used for two reasons:

1. The optimal values of [ii] under the *modified* CR and AR strategies are a function of N_c (an unknown quantity which is specified in order to achieve the optimal values of [ii]) and, therefore, solving for the optimal values of [ii] and N_c would require an iterative solution.
2. Empirical findings of this research indicate that the optimal values of [ii] under the IR strategy result in values of [ii] which are *near*-optimal under the *modified* CR and AR strategies, and these *near*-optimal values of [ii] generally result in the same Min-**V** values of N_c due to the integer restriction on N_c.

The **V** components of **J** for the IR strategy, the pure and modified CR strategies, and the pure, modified, and 3-block AR strategies are shown in Appendix M. Because the optimal values of N_c are essentially the same for each strategy, only the optimal value of [ii] obtained under the IR strategy (which is used to compute the optimal value of N_c) is derived in this research. When set equal to zero, the partial derivative of V_{IR} with respect to [ii] becomes

$$\frac{\partial V_{IR}}{\partial [ii]} = \frac{\partial}{\partial \theta}\left[\frac{k\, w_{ii}}{[ii]} + \frac{r\,k\,(k-1)\,w_{iijj}}{2\theta\,[ii]} + \frac{\theta\,(r+k-1)}{\theta\,(r+k-1) - r\,k\,[ii]} \right]$$

$$+ \frac{\partial}{\partial \theta}\left[\frac{-2\,k\,r\,w_{ii}}{\theta\,(r+k-1) - r\,k\,[ii]} \right]$$

$$+ \frac{\partial}{\partial \theta}\left[(k\,r\,w_{iiii})\,\frac{\theta\,(r+k-2) - r\,(k-1)\,[ii]}{\theta^2\,(r-1)\,(r+k-1)\,[ii] - r\,k\,\theta\,(r-1)\,[ii]^2} \right]$$

$$+ \frac{\partial}{\partial \theta}\left[(k\,(k-1)\,r\,w_{iijj})\,\frac{r\,[ii] - \theta}{\theta^2\,(r-1)\,(r+k-1)\,[ii] - r\,k\,\theta\,(r-1)\,[ii]^2} \right] = 0$$

$$0 = 2\theta\,k\,w_{ii} + r\,k\,(k-1)\,w_{iijj}$$

$$- \frac{2\theta^2\,r\,k\,\ell_1\,(r+k-1)\,[ii]^2 - 4\theta\,r^2\,k^2\,\ell\,w_{ii}\,[ii]^2}{d_1^{\,2} - 2\,d_1\,e_1\,[ii] + e_1^{\,2}\,[ii]^2}$$

$$- \frac{m_1\,(2\,b_1\,e_1\,[ii] - b_1\,d_1 - c_1\,e_1\,[ii]^2) + n_1\,(r\,e_1\,[ii]^2 + \theta\,d_1 - 2\theta\,e_1\,[ii])}{d_1^{\,2} - 2\,d_1\,e_1\,[ii] + e_1^{\,2}\,[ii]^2}$$

where $b_1 = \theta\,r\,(r+k-2)$

$c_1 = r^2\,(k-1)$

$d_1 = \theta^2\,(r-1)\,(r+k-1)$

$e_1 = r\,k\,\theta\,(r-1)$

$\ell_1 = \theta^2\,(r-1)^2$

$m_1 = 2\theta\,k\,w_{iiii}$

$n_1 = 2\theta\,r\,k\,(k-1)\,w_{iijj}$.

Collecting the $[ii]^2$, $[ii]^1$, and $[ii]^0$ terms, the quadratic equation for the partial derivative of V_{IR} set equal to zero becomes

$$0 = (-2rk\ell_1\theta^2(r+k-1) + 4r^2k^2\ell_1\theta w_{ii} + c_1 e_1 m_1 - n_1 r e_1 + e_1^2 j_1)\,[ii]^2$$

$$+\ (-2j_1 d_1 e_1 - 2m_1 b_1 e_1 + 2n_1\theta e_1)\,[ii]^1 \ +\ (b_1 d_1 m_1 - n_1\theta d_1 + d_1^2 j_1)\,[ii]^0$$

$$= \ddot{o}_1\,[ii]^2 \ +\ p_1\,[ii]^1 \ +\ q_1\,[ii]^0$$

where $j_1 = 2\theta k w_{ii} + rk(k-1)w_{iijj}$.

Applying the quadratic formula, the Min-V_{IR} | Min-B value of [ii] becomes

$$[ii] = \frac{-p_1 \pm \sqrt{p_1^2 - 4\ddot{o}_1 q_1}}{2\ddot{o}_1}$$

where the "−" root is the V-optimal value of [ii] because the "+" root leads to a Max-V value of [ii] which is infeasible for the designs examined in this research. (See Figure P-1 on page 421.)

The V* component of J* is also shown in Appendix M for the IR, CR, and AR strategies. When set equal to zero, the partial derivative of V^*_{IR} with respect to [ii] becomes

$$\frac{\partial V^*_{IR}}{\partial [ii]} = \frac{\partial}{\partial\theta}\left[\frac{k}{[ii]} + \frac{k(k-1)w_{ii}r}{\theta[ii]} + (4kw_{ii})\frac{\theta(r+k-2) - r(k-1)[ii]}{\theta^2(r-1)(r+k-1)[ii] - rk\theta(r-1)[ii]^2}\right] = 0$$

$$0 = k\theta + a_1 + 4k\theta w_{ii}\left[\frac{b_1 d_1 - 2b_1 e_1[ii] + c_1 e_1[ii]^2}{d_1^2 - 2d_1 e_1[ii] + e_1^2[ii]^2}\right]$$

where $a_1 = k(k-1)w_{ii}r$.

Collecting the $[ii]^2$, $[ii]^1$, and $[ii]^0$ terms, the quadratic equation for the partial derivative of V^*_{IR} set equal to zero becomes

$$0 = (c_1 e_1 + f_1 e_1^2)\,[ii]^2 \ +\ (-2b_1 e_1 - 2d_1 e_1 f_1)\,[ii]^1 \ +\ (b_1 d_1 + f_1 d_1^2)\,[ii]^0$$

$$= g_1\,[ii]^2 \ +\ h_1\,[ii]^1 \ +\ i_1\,[ii]^0 \ .$$

where $f_1 = \dfrac{k\theta + a_1}{4k\theta w_{ii}}$.

Applying the quadratic formula, the Min-$\mathbf{V^*}_{IR}$ | Min-$\mathbf{B^*}$ value of [ii] becomes

$$[ii] = \frac{-h_1 \pm \sqrt{h_1{}^2 - 4g_1 i_1}}{2g_1}$$

where the "−" root is the $\mathbf{V^*}$-optimal value of [ii] because the "+" root leads to a Max-$\mathbf{V^*}$ value of [ii] which is infeasible for the designs examined in this research. (See Figure P-2 on page 422.)

The second partial derivatives of $\mathbf{V}$ and $\mathbf{V^*}$ with respect to [ii] are positive quantities for the feasible values of [ii]; that is, for the values of [ii] requiring a positive number of center runs. The second partial derivatives are negative quantities for the infeasible values of [ii]; that is, for the values of [ii] requiring a negative number of center runs. Therefore, the Min-$\mathbf{V}$ and Min-$\mathbf{V^*}$ values of [ii] are the "−" roots of the partial derivatives of $\mathbf{V}$ and $\mathbf{V^*}$ set equal to zero, both of which result in a positive number of center runs for the designs examined in this research (except for the Min-$\mathbf{V}$ FACs). The $k=2$ CCD is used to illustrate the relationships between [ii], N_c, $\mathbf{V}$, and $\mathbf{V^*}$ in the figures at the end of this appendix. Figure P-1 on page 421 is a plot of $\mathbf{V}$ and N_c versus [ii], and Figure P-2 on page 422 is a plot of $\mathbf{V^*}$ and N_c versus [ii].

For the CCDs examined in this research, [ii] can be written as

$$[ii] = \frac{g^2(F + 2\alpha^2)}{F + n_\alpha + N_c} \ .$$

Solving for the number of center runs needed to achieve the optimal value of [ii], the optimal value of N_c for the CCDs becomes

$$N_c = \frac{g^2(F + 2\alpha^2)}{[ii]} - (F + n_\alpha) \ .$$

For the BBDs and FACs, [ii] can be written as

$$[ii] = \frac{f\,F\,g^2}{F + N_c}$$

and solving for the number of center runs needed to achieve the optimal value of [ii], the optimal value of N_c for the BBDs and FACs becomes

$$N_c = \frac{f\,F\,g^2}{[ii]} - F\ .$$

Table P-1 on page 418 shows the Min-V | Min-B and Min-V* | Min-B* values of [ii] for the second order designs and Table P-2 on page 419 shows the number of center runs, N_c, needed to achieve the optimal values of [ii]. Because the number of center runs must be integer-valued, the actual values of [ii] are not the *exact* Min-V and Min-V* values. The actual designs utilize the closest positive integer to the optimal values of N_c, resulting in *near* optimal values of [ii], and *near* optimal designs.

The optimal values of N_c for the Min-V* | Min-B* designs are independent of the shape of the region of interest; that is, N_c is the same for a spherical and a cuboidal region of interest. However, the scaling factor, *g*, and therefore [ii], are different for spherical and cuboidal regions.

For the CR strategy, common random number streams can be used for a maximum of one center, and any additional center runs must use independent streams. Therefore, the number of design points utilizing independent streams (N_2) in the CR strategy is

$$\begin{array}{lll} N_2 = N_c - 1 & \text{if} & N_c \geq 1\ ,\ \text{and} \\ N_2 = 0 & \text{if} & N_c = 0\ . \end{array}$$

For the AR strategy, a common/antithetic pair of random number streams can be used for a maximum of two center runs, and any additional center runs must use independent streams. Therefore, the number of design points utilizing independent streams for the AR strategy is

$$N_2 = N - N_{1a} - N_{1b}$$

where N_{1a} and N_{1b} are the number of design points in the 1st and 2nd blocks, respectively. The values of N_{1a} and N_{1b} vary with the number of center runs, the design class, and the number of blocks into which the design partitions. For the CCDs and SCDs examined in this research, the values of N_{1a} and N_{1b} become

$$
\begin{array}{llll}
N_{1a} = F + 1 & \text{and} \quad N_{1b} = n_a + 1 & \text{if} & N_c \geq 2 \\
N_{1a} = F & \text{and} \quad N_{1b} = n_a + 1 & \text{if} & N_c = 1 \\
N_{1a} = F & \text{and} \quad N_{1b} = n_a & \text{if} & N_c = 0
\end{array}
$$

with the exception of the $k=3$ SCD with one center run. In order to make the block sizes as close as possible, the single center run of this design is placed in the factorial block.

For the $k=5, 7$ BBDs, which partition into two orthogonal blocks, N_{1a} and N_{1b} become

$$
\begin{array}{lll}
N_{1a} = N_{1b} = F/2 + 1 & \text{if} & N_c \geq 2 \text{ , and} \\
N_{1a} = N_{1b} = F/2 & \text{if} & N_c \leq 1 \text{ .}
\end{array}
$$

For the $k=4$ BBD and the FACs, which partition into three orthogonal blocks, N_2 includes the design points in the 3rd block. For the $k=4$ BBD, N_2 also includes all of the center runs in the design. The values of N_{1a} and N_{1b} for these designs become

$$N_{1a} = N_{1b} = F/3 \text{ .}$$

The optimal values of N_c, and therefore the values N_2, vary with the design criteria (J or J*) and the shape of the region of interest (spherical or cuboidal). Table P-3 on page 420 summarizes the values of N_{1a}, N_{1b}, and N_2 for each of the second order designs.

Table P-1. Optimal Values of θ and [ii] for Second Order Response Surface Designs.
The optimal designs are Min-V | Min-B and Min-V* | Min-B* designs.

DESIGN CLASS	Number of Factors (k)	θ for a Min-B Design		θ for a Min-B* Design		[ii] for a Min-V|Min-B Design		[ii] for a Min-V*|Min-B* Design	
		Spherical Region	Cuboidal Region	Cuboidal Region	Spherical Region	Spherical Region	Cuboidal Region	Cuboidal Region	Spherical Region
CCD	2	.5	.7	1.0	.75	.264929	.363775	.483805	.362854
CCD	3	.450700	.79924	1.05163	.630979	.207920	.360858	.448024	.268814
CCD	4	.389551	.83104	1.03880	.519401	.171326	.357558	.426876	.213438
CCD	5	.318761	.79235	.95628	.409836	.145842	.354767	.412865	.176942
CCD	6 (½ fr.)	.328469	.93066	1.09490	.410587	.127041	.353277	.402830	.151061
CCD	7 (½ fr.)	.254187	.80775	.93202	.310673	.112554	.351236	.395039	.131680
BBD	4	.375	.8	1.0	.5	.171312	.357295	.426804	.213402
BBD	5	.388889	.96667	1.16667	.5	.145940	.355803	.413472	.177202
BBD	7	.272727	.86667	1.0	.333333	.112547	.351351	.394979	.131660
FAC	3	.306122	.54286	.71429	.428571	.211153	.362200	.461377	.276826
FAC	4	.25	.53333	.66667	.333333	.173258	.358747	.436645	.218322
FAC	5	.212121	.52727	.63636	.272727	.147029	.355993	.420065	.180028
FAC	6 (⅓ fr.)	.184615	.52308	.61538	.230769	.127791	.353783	.408188	.153070
FAC	7 (⅓ fr.)	.163636	.52	.6	.2	.113063	.351978	.399257	.133086
SCD-H	3	.471331	.90058	1.18497	.710983	.208090	.363471	.451124	.270674
SCD-H	4	.501724	1.0	1.25	.625	.172208	.359776	.429406	.214703
SCD-D	5	.467958	1.10142	1.32930	.569701	.146380	.357030	.415130	.177913
SCD-H	6	.445429	1.20531	1.41802	.531756	.127408	.354836	.404841	.151816
SCD-D	7	.361333	1.14823	1.32489	.441629	.112656	.352283	.395895	.131965

Note: The optimal values of θ and [ii] for the SCDs are those computed using the formulas for designs with *odd* design moments through order five equal to zero, and therefore these values are not "truly" optimal for the SCDs.

Table P-2. Optimal Values of g and N_c for Second Order Response Surface Designs. The optimal designs are Min-V | Min-B and Min-V* | Min-B* designs.

DESIGN CLASS	Number of Factors (k)	g for a Min-B Design		g for a Min-B* Design		N_c for a Min-V|Min-B Design		N_c for a Min-V*|Min-B* Design
		Spherical Region	Cuboidal Region	Cuboidal Region	Spherical Region	Spherical Region	Cuboidal Region	Both Regions
CCD	2	.577350	.683130	.81650	.707107	$2.07 \cong 2$	$2.26 \cong 2$	$3.02 \cong 3$
CCD	3	.484044	.644585	.73939	.572729	$2.03 \cong 2$	$2.38 \cong 2$	$3.35 \cong 3$
CCD	4	.428670	.626113	.70002	.494986	$2.25 \cong 2$	$2.83 \cong 3$	$4.09 \cong 4$
CCD	5	.391159	.616707	.67751	.443533	$2.76 \cong 3$	$3.74 \cong 4$	$5.44 \cong 5$
CCD	6 (½ fr.)	.362570	.610296	.66196	.405366	$2.16 \cong 2$	$3.03 \cong 3$	$4.52 \cong 5$
CCD	7 (½ fr.)	.340134	.606334	.65131	.376032	$2.96 \cong 3$	$4.45 \cong 4$	$6.58 \cong 7$
BBD	4	.612372	.894427	1.0	.707107	$2.27 \cong 2$	$2.87 \cong 3$	$4.12 \cong 4$
BBD	5	.623610	.983192	1.08012	.707107	$2.64 \cong 3$	$3.47 \cong 3$	$5.15 \cong 5$
BBD	7	.522233	.930949	1.0	.577350	$2.16 \cong 2$	$3.20 \cong 3$	$4.76 \cong 5$
FAC	3	.553283	.736788	.84515	.654654	$-0.90 \cong 0$	$-0.02 \cong 0$	$0.87 \cong 0$
FAC	4	.5	.730297	.81650	.577350	$-3.08 \cong 0$	$-0.72 \cong 0$	$1.45 \cong 0$
FAC	5	.460566	.726135	.79772	.522233	$-9.28 \cong 0$	$-3.06 \cong 0$	$2.42 \cong 0$
FAC	6 (⅓ fr.)	.429669	.723241	.78446	.480384	$-9.0 \cong 0$	$-3.5 \cong 0$	$1.23 \cong 0$
FAC	7 (⅓ fr.)	.404520	.721110	.77460	.447214	$-25.6 \cong 0$	$-11.0 \cong 0$	$1.36 \cong 0$
SCD-H	3	.516962	.662802	.76029	.588915	$1.30 \cong 1$	$1.60 \cong 2$	$2.30 \cong 2$
SCD-H	4	.419359	.632455	.70711	.5	$1.36 \cong 1$	$1.79 \cong 2$	$2.63 \cong 3$
SCD-D	5	.385299	.619644	.68073	.445645	$1.33 \cong 1$	$1.82 \cong 2$	$2.73 \cong 3$
SCD-H	6	.358513	.612042	.66385	.406525	$1.26 \cong 1$	$1.81 \cong 2$	$2.74 \cong 3$
SCD-D	7	.339890	.605898	.65084	.375762	$1.56 \cong 2$	$2.27 \cong 2$	$3.45 \cong 3$

Note: The number of center runs required for the Min-V | Min-B 3^k designs are less than zero and, therefore, N_c has been rounded up to $N_c \cong 0$. For convenience and consistency reasons, the number of center runs has been rounded down to $N_c \cong 0$ for the Min-V* | Min-B* 3^k designs.

Table P-3. Optimal Values of N_1 and N_2 for Second Order Response Surface Designs.
The optimal designs are Min-V | Min-B and Min-V* | Min-B* designs.
N_2 is the number of design points using independent random number streams.
N_{1a} and N_{1b} are the number of design points using common and antithetic streams.

DESIGN CLASS	Number of Factors (k)	Values of N_1 and N_2 for a:											
		Min-V|Min-B Design								Min-V*|Min-B* Design			
		CR Strategy		AR Strategy						CR	AR Strategy		
		Sph. Reg.	Cub. Reg.	Spherical Region			Cuboidal Region			Both Regs.	Both Regions		
		N_2	N_2	N_{1a}	N_{1b}	N_2	N_{1a}	N_{1b}	N_2	N_2	N_{1a}	N_{1b}	N_2
CCD	2	1	1	5	5	0	5	5	0	2	5	5	1
CCD	3	1	1	9	7	0	9	7	0	2	9	7	1
CCD	4	1	2	17	9	0	17	9	1	3	17	9	2
CCD	5	2	3	33	11	1	33	11	2	4	33	11	3
CCD	6 (½ fr.)	1	2	33	13	0	33	13	1	4	33	13	3
CCD	7 (½ fr.)	2	3	65	15	1	65	15	2	6	65	15	5
BBD	4	1	2	8	8	10	8	8	11	3	8	8	12
BBD	5	2	2	21	21	1	21	21	1	4	21	21	3
BBD	7	1	2	29	29	0	29	29	1	4	29	29	3
FAC	3	0	0	9	9	9	9	9	9	0	9	9	9
FAC	4	0	0	27	27	27	27	27	27	0	27	27	27
FAC	5	0	0	81	81	81	81	81	81	1	81	81	81
FAC	6 (1/3 fr.)	0	0	81	81	81	81	81	81	0	81	81	81
FAC	7 (1/3 fr.)	0	0	243	243	243	243	243	243	0	243	243	243
SCD-H	3	0	1	5	6	0	5	7	0	1	5	7	0
SCD-H	4	0	1	8	9	0	9	9	0	2	9	9	1
SCD-D	5	0	1	12	11	0	13	11	0	2	13	11	1
SCD-H	6	0	1	16	13	0	17	13	0	2	17	13	1
SCD-D	7	1	1	29	15	0	29	15	0	2	29	15	1

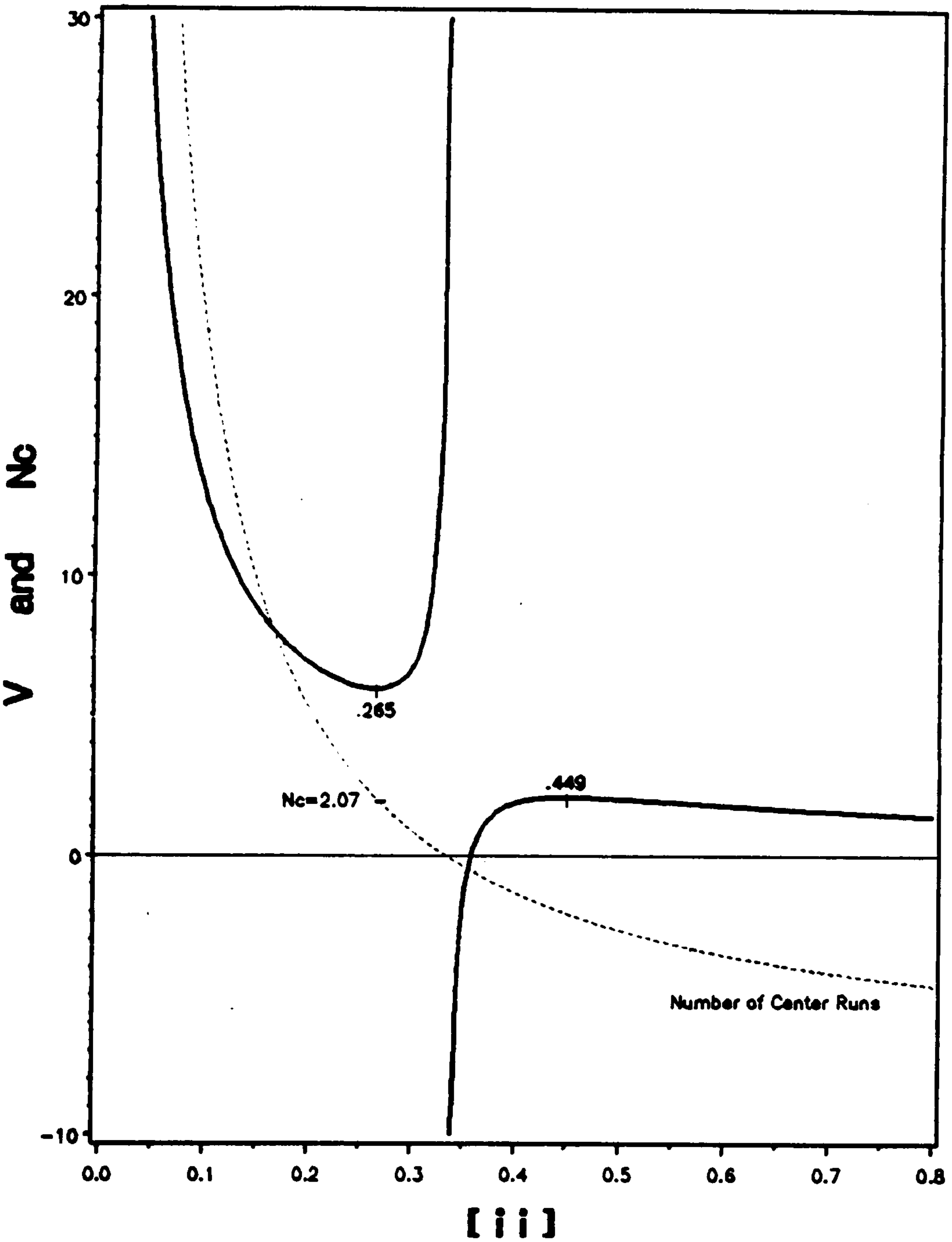


Figure P-1 V and N_c versus [ii] for a Min-B $k=2$ CCD.

Region of interest is Spherical.

The B-optimal value of $\theta =$ [iiii]/[ii] is .50.

The B-optimal value of the scaling factor is $g = .5773503$.

The horizontal line at $N_c = 0$ indicates that a negative number of center runs would be required below this line.

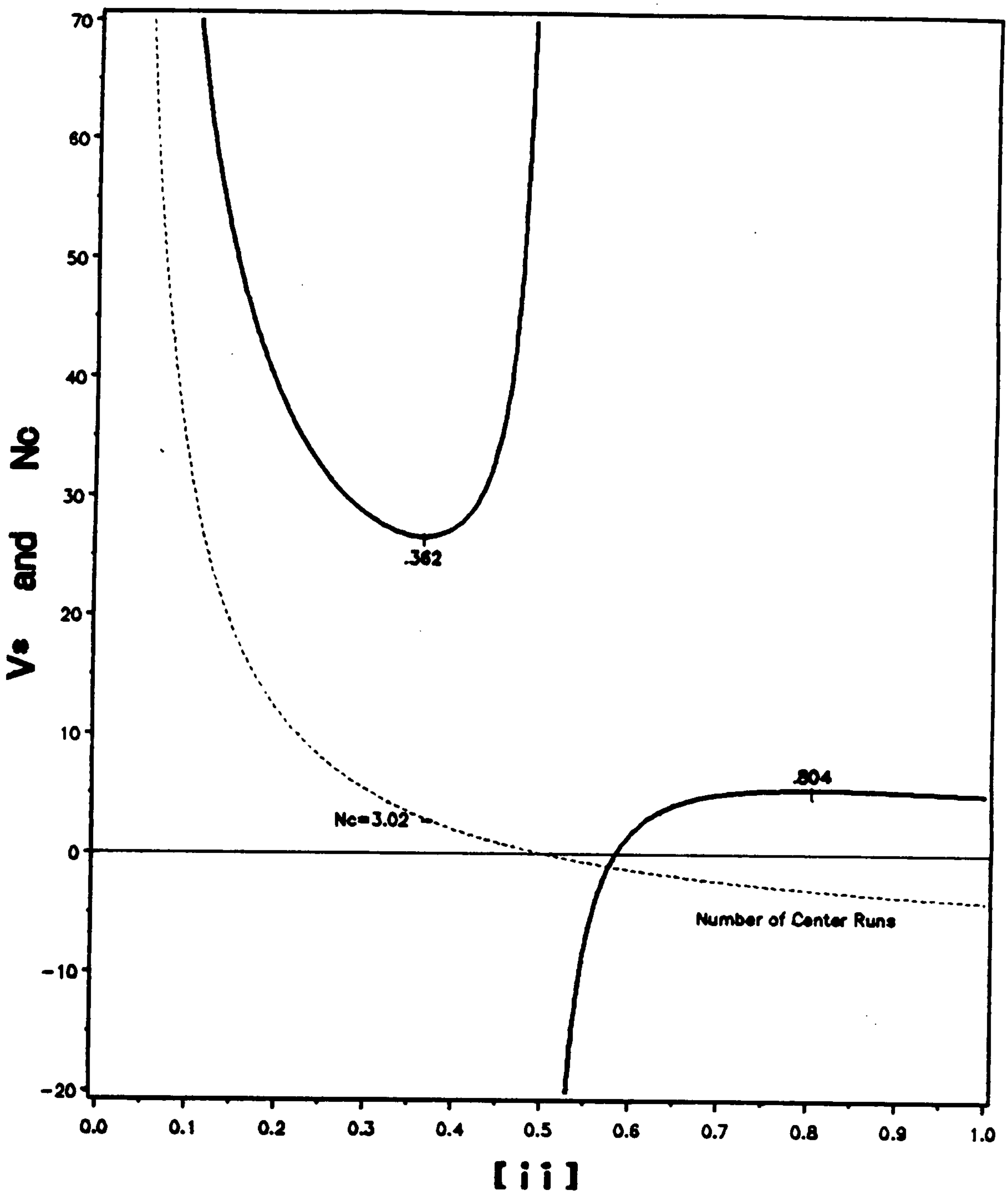


Figure P-2 V* and N_c versus [ii] for a Min-B* $k=2$ CCD.
Region of interest is Spherical.
The B*-optimal value of $\theta=$[iiii]/[ii] is 1.0.
The B*-optimal value of the scaling factor is $g=.7071070$.
The horizontal line at $N_c=0$ indicates that a negative number of center runs would be required below this line.

Appendix Q

Computer Program for Determination of Min-V | Min-B Second Order Designs

```
//A0392JMD JOB 33F02,JMD,MSGLEVEL=(0,0),REGION=2048K
//*
//*
//*   This program computes the Min-V|Min-B values of  V
//*                                 -----------      -
//*              for various amounts of induced correlation
//*              under the IR, CR, and AR strategies.
//*
//*   The SECOND order designs examined are:
//*              k = 2,...,7 CCDs (central composite designs)
//*              k = 4,5,7   BBDs (Box-Behnken designs)
//*              k = 3,...,7 FACs (3-level factorial designs)
//*              k = 3,...,7 SCDs (small composite designs)
//*
//*   This program is not entirely applicable to the SCDs
//*              but the "optimal" values of the scaling
//*              factor (g) and the number of center runs (Nc)
//*              are computed for the SCDs using this program.
//*
//*
/*JOBPARM LINES=4
/*PRIORITY IDLE
//STEP1   EXEC SAS
//SYSIN DD *
OPTIONS CENTER NODATE NONUMBER LS=72;
TITLE1;
TITLE2 'Min-V | Min-B  designs';
DATA Q1;
INPUT DES$ 1-3 DESIGN$ 1-4 DES2$ 2-3 K 4 F 6-8;
*
Define  REGION=1 if Spherical or REGION=2 if Cuboidal;
REGION=1;
TITLE4 'Cuboidal Region';
```

```
TITLE4 'Spherical Region';
*
DES = CCD or BBD or FAC or SCD
DESIGN = CCD2 ... CCD7, BBD4, BBD5, BBD7, FAC3 ... FAC7, SCD3 ... SCD7
K = Number of Factors in Model
F = Number of Factorial Design Points;
*
Define  AX = Number of axial design points;
AX=0; IF DES2='CD' THEN AX=2*K;
*
Define  NB = Number of blocks into which the design partitions;
NB=2; IF DES='FAC' OR DESIGN='BBD4' THEN NB=3;
*
Define  NC1 & NC2 = Number of factorial & axial center points;
IF DES2='CD' THEN NC1=1;
IF DES2='CD' THEN NC2=1;
IF REGION=1 AND DES='SCD' AND K<7 AND K>3 THEN NC1=0;
IF REGION=1 AND DES='SCD' AND K=3 THEN NC2=0;
*
Define Region moments;
IF REGION=1 THEN DO;
WII=1/(K+2);
WIIJJ=WII/(K+4);
WIIII=3*WIIJJ;
WIIJJKK=WIIJJ/(K+6);
WIIIIJJ=3*WIIJJKK;
WIIIIII=15*WIIJJKK;
END;
IF REGION=2 THEN DO;
WII=1/3;
WIIJJ=1/9;
WIIII=1/5;
WIIJJKK=1/27;
WIIIIJJ=1/15;
WIIIIII=1/7;
END;
*
Compute ALPHA for the CCDs and SCDs;
N=F+AX+NC1+NC2;
ALP=0;
IF DES2='CD' THEN ALP=SQRT((F*N-F*F-F*NC1)/(2*F+2*NC1));
ALP22=2*ALP*ALP;
ALP24=ALP22*ALP*ALP;
FPALP22=F+ALP22;
FPALP24=F+ALP24;
*
Define  IIFAC = proportion of non-zero factorial levels of i'th point;
IIFAC=1;
IF DES='FAC' THEN IIFAC=2/3;
IF DESIGN='BBD4' THEN IIFAC=1/2;
IF DESIGN='BBD5' THEN IIFAC=2/5;
IF DESIGN='BBD7' THEN IIFAC=3/7;
*
Compute  R = [iiii] / [iijj] ;
R=(FPALP24/F)/IIFAC;
IF DES='BBD' THEN R=3;
IF DESIGN='BBD5' THEN R=4;
*
Determine the Min-B Optimal Values of  THE=Theta  &  G=Scaling factor;
KM1=K-1;
KP2=K+2;
RM1=R-1;
RK1=R+K-1;
RK2=R+K-2;
GFAC=WIIII+KM1*WIIJJ;
THE=(R*GFAC)/(WII*RK1);
THE2=THE*THE;
```

```
G=SQRT(THE*FPALP22/FPALP24);
*
Determine the Min-V|Min-B Optimal Values of  IIOPT=[ii]  &  NCOPT=Nc;
B1=THE*R*RK2;
C1=R*R*KM1;
D1=THE2*RM1*RK1;
E1=R*K*THE*RM1;
J1=2*THE*K*WII+R*K*KM1*WIIJJ;
L1=THE2*RM1*RM1;
M1=2*THE*K*WIIII;
N1=2*THE*R*K*KM1*WIIJJ;
O1=(-2)*R*K*L1*THE2*RK1+4*R*R*K*K*L1*THE*WII+C1*E1*M1-N1*R*E1+E1*E1*J1;
P1=(-2)*J1*D1*E1-2*M1*B1*E1+2*N1*THE*E1;
Q1=(-1)*N1*THE*D1+B1*D1*M1+D1*D1*J1;
INSQRT=P1*P1-4*O1*Q1;
IIOPT=(P1*(-1)-SQRT(INSQRT))/(2*O1);
NCOPT=(IIFAC*G*G*FPALP22/IIOPT)-F-AX;
*
Determine the actual value of II=[ii] using the integer value of NC=Nc;
NC=ROUND(NCOPT);
IF DES='FAC' THEN NC=0;
N=F+AX+NC;
II=IIFAC*G*G*FPALP22/N;
II2=II*II;
*
Determine IIJJ=[iijj] and check for a singular X'X matrix;
IIJJ=THE*II/R;
IIJJ_SIN=(II/KM1)*(K*II-THE);
IF IIJJ=IIJJ_SIN THEN SINGULAR='Yes';
ELSE SINGULAR='No';
IF DES='SCD' THEN IIJJ_SIN='.';
IF DES='SCD' THEN SINGULAR='.';
*
Determine DELTA = the difference between blocks 1 & 2 sum of squares;
DELTA=0;
IF DES2='CD' THEN DELTA=G*G*(F-ALP22);
*
Determine the components of the inverted X'X matrix;
A=THE*II*RK1/R;
B=(THE*II*RK2/R)-(KM1*II2);
C=(II/R)*(R*II-THE);
D=A-K*II2;
D2=D*THE*II*RM1/R;
*
Determine the value of  V  under the  IR Strategy;
VMIR=((A-2*K*WII*II)/D)+(K*WII/II)+(((B*K*WIIII)+(C*K*KM1*WIIJJ))/D2)
      +((K*KM1*R*WIIJJ)/(2*THE*II));
*
Determine N1 & N2 = Number design points in blocks 1 & 2 (AR Strategy);
N1=F/NB;
N2=N1;
IF DES2='CD' THEN N1=F+1;
IF DES2='CD' THEN N2=AX+1;
IF REGION=1 AND DESIGN='SCD3' THEN N2=AX;
IF REGION=1 AND DES='SCD' AND K>3 AND K<7 THEN N1=F;
*
Determine N3C & N3A = Number of design points using independent
                              streams in the  CR  &  AR Strategies;
N3C=0;
IF NC>1 THEN N3C=NC-1;
IF N3C=0 THEN SQRTN3C=0;
ELSE SQRTN3C=SQRT(N3C);
N3A=0;
IF NC>2 THEN N3A=NC-2;
IF DESIGN='BBD4' OR DES='FAC' THEN N3A=N1+NC;
IF N3A=0 THEN SQRTN3A=0;
ELSE SQRTN3A=SQRT(N3A);
```

```
FORMAT THE G ALP 10.8;
FORMAT IIOPT II 6.4;
FORMAT DELTA IIJJ IIJJ_SIN 5.3;
FORMAT NCOPT 5.2;
FORMAT IIFAC R 4.2;
CARDS;
CCD2    4
CCD3    8
CCD4   16
CCD5   32
CCD6   32
CCD7   64
BBD4   24
BBD5   40
BBD7   56
FAC3   27
FAC4   81
FAC5  243
FAC6  243
FAC7  729
SCD3    4
SCD4    8
SCD5   12
SCD6   16
SCD7   28
;
*
Print Input Data,  Optimal values of Theta, g, [ii], Nc, and
      Actual values of [ii], Nc, N1, N2, N3, N;

DATA Q1DOT;
SET Q1;
IF DES2 NE 'CD' THEN ALP='.';
IF DES2 NE 'CD' THEN AX='.';
IF DES='SCD' THEN VMIR='.';
PROC PRINT NOOBS DATA=Q1DOT;
TITLE6 'Second Order Designs  and values of Design Parameters';
VAR DES K F AX R ALP IIFAC NB DELTA;

PROC PRINT NOOBS DATA=Q1;
TITLE6
'THETA & G  for Min-B   and   [ii] & Nc  for Min-V|Min-B  Designs';
VAR DESIGN THE G IIOPT NCOPT NC;

PROC PRINT NOOBS DATA=Q1;
TITLE6 'Check for a SINGULAR X"X matrix';
VAR DESIGN SINGULAR IIJJ IIJJ_SIN;

PROC PRINT NOOBS DATA=Q1DOT;
TITLE6 'Actual Values of [ii], Nc, N1, N2, N3A, N3C, N, V-IR ';
VAR DESIGN II NC N1 N2 N3A N3C N VMIR;
*
*
Determine the value of  V  under the  CR Strategy;
DATA Q1VCR
            DATA Q1CCD2C (KEEP=RP VCCCD2B)
            DATA Q1CCD3C (KEEP=RP VCCCD3B)
            DATA Q1CCD4C (KEEP=RP VCCCD4B)
            DATA Q1CCD5C (KEEP=RP VCCCD5B)
            DATA Q1CCD6C (KEEP=RP VCCCD6B)
            DATA Q1CCD7C (KEEP=RP VCCCD7B)
            DATA Q1BBD4C (KEEP=RP VCBBD4B)
            DATA Q1BBD5C (KEEP=RP VCBBD5B)
            DATA Q1BBD7C (KEEP=RP VCBBD7B)
            DATA Q1FAC3C (KEEP=RP VCFAC3B)
            DATA Q1FAC4C (KEEP=RP VCFAC4B)
```

```
            DATA Q1FAC5C (KEEP=RP VCFAC5B)
            DATA Q1FAC6C (KEEP=RP VCFAC6B)
            DATA Q1FAC7C (KEEP=RP VCFAC7B);
SET Q1;
IF DES='SCD' THEN DELETE;
*
Define components of  V  for the  CR Strategy;
Q1C=(SQRTN3C*A)/(N*D);
R1C=(-1)*(SQRTN3C*II)/(N*D);
S1C=( ((N-N3C)*A) - (K*N*II*II) )/ (N*D);
T1C=(N3C*II)/(N*D);
*
Compute  V  for   RP=rho+   RP=(0, .3, .6, .9);
DO RP=0 TO .9 BY .3;
IF N3C=0 THEN VMCR=((1-RP)*VMIR) + (N*RP);
ELSE VMCR=((1-RP)*VMIR) +
           (N*RP*(S1C*S1C+2*K*WII*S1C*T1C+K*T1C*T1C*GFAC))+
           (N*RP*(Q1C*Q1C+2*K*WII*Q1C*R1C+K*R1C*R1C*GFAC));
IF DESIGN='CCD2' THEN VCCCD2B=VMCR;
IF DESIGN='CCD3' THEN VCCCD3B=VMCR;
IF DESIGN='CCD4' THEN VCCCD4B=VMCR;
IF DESIGN='CCD5' THEN VCCCD5B=VMCR;
IF DESIGN='CCD6' THEN VCCCD6B=VMCR;
IF DESIGN='CCD7' THEN VCCCD7B=VMCR;
IF DESIGN='BBD4' THEN VCBBD4B=VMCR;
IF DESIGN='BBD5' THEN VCBBD5B=VMCR;
IF DESIGN='BBD7' THEN VCBBD7B=VMCR;
IF DESIGN='FAC3' THEN VCFAC3B=VMCR;
IF DESIGN='FAC4' THEN VCFAC4B=VMCR;
IF DESIGN='FAC5' THEN VCFAC5B=VMCR;
IF DESIGN='FAC6' THEN VCFAC6B=VMCR;
IF DESIGN='FAC7' THEN VCFAC7B=VMCR;
OUTPUT;
END;

DATA QQ1CCD2C; SET Q1CCD2C; IF VCCCD2B='.' THEN DELETE;
DATA QQ1CCD3C; SET Q1CCD3C; IF VCCCD3B='.' THEN DELETE;
DATA QQ1CCD4C; SET Q1CCD4C; IF VCCCD4B='.' THEN DELETE;
DATA QQ1CCD5C; SET Q1CCD5C; IF VCCCD5B='.' THEN DELETE;
DATA QQ1CCD6C; SET Q1CCD6C; IF VCCCD6B='.' THEN DELETE;
DATA QQ1CCD7C; SET Q1CCD7C; IF VCCCD7B='.' THEN DELETE;
DATA QQ1BBD4C; SET Q1BBD4C; IF VCBBD4B='.' THEN DELETE;
DATA QQ1BBD5C; SET Q1BBD5C; IF VCBBD5B='.' THEN DELETE;
DATA QQ1BBD7C; SET Q1BBD7C; IF VCBBD7B='.' THEN DELETE;
DATA QQ1FAC3C; SET Q1FAC3C; IF VCFAC3B='.' THEN DELETE;
DATA QQ1FAC4C; SET Q1FAC4C; IF VCFAC4B='.' THEN DELETE;
DATA QQ1FAC5C; SET Q1FAC5C; IF VCFAC5B='.' THEN DELETE;
DATA QQ1FAC6C; SET Q1FAC6C; IF VCFAC6B='.' THEN DELETE;
DATA QQ1FAC7C; SET Q1FAC7C; IF VCFAC7B='.' THEN DELETE;

DATA Q1CCDC; MERGE QQ1CCD2C QQ1CCD3C QQ1CCD4C QQ1CCD5C QQ1CCD6C QQ1CCD7C;
                                                                  BY RP;
DATA Q1BBDC; MERGE QQ1BBD4C QQ1BBD5C QQ1BBD7C;                    BY RP;
DATA Q1FACC; MERGE QQ1FAC3C QQ1FAC4C QQ1FAC5C QQ1FAC6C QQ1FAC7C; BY RP;

PROC PRINT NOOBS DATA=Q1CCDC;
TITLE6 'V  for the CCDs under the CR Strategy';

PROC PRINT NOOBS DATA=Q1BBDC;
TITLE6 'V  for the BBDs under the CR Strategy';

PROC PRINT NOOBS DATA=Q1FACC;
TITLE6 'V  for the FACs under the CR Strategy';
*
*
Determine the value of  V  under the  AR Strategy;
DATA Q1VAR
```

```
            DATA Q1CCD2A (KEEP=RP RM VACCD2B)
            DATA Q1CCD3A (KEEP=RP RM VACCD3B)
            DATA Q1CCD4A (KEEP=RP RM VACCD4B)
            DATA Q1CCD5A (KEEP=RP RM VACCD5B)
            DATA Q1CCD6A (KEEP=RP RM VACCD6B)
            DATA Q1CCD7A (KEEP=RP RM VACCD7B)
            DATA Q1BBD4A (KEEP=RP RM VABBD4B)
            DATA Q1BBD5A (KEEP=RP RM VABBD5B)
            DATA Q1BBD7A (KEEP=RP RM VABBD7B)
            DATA Q1FAC3A (KEEP=RP RM VAFAC3B)
            DATA Q1FAC4A (KEEP=RP RM VAFAC4B)
            DATA Q1FAC5A (KEEP=RP RM VAFAC5B)
            DATA Q1FAC6A (KEEP=RP RM VAFAC6B)
            DATA Q1FAC7A (KEEP=RP RM VAFAC7B);
SET Q1;
IF DES='SCD' THEN DELETE;
*
Define components of  V  for the  AR Strategy;
HFAC=2/NB;
E1A=((1-HFAC)*B/(N*D2))+((HFAC*NC*II2)/(N*N*D*D));
F1A=((-1)*4/(N*N*II*II)) + (NC*II*II/(N*N*D*D));
M1A=((N2-N1)*II/(N*D)) + (DELTA/(N*D));
P1A=((N1-N2)*A/(N*D)) - (K*II*DELTA/(N*D));
Q1A=SQRTN3A*A/(N*D);
R1A=(-1)*(SQRTN3A*II)/(N*D);
S1A=((N-N3A)*A/(N*D)) - (HFAC*K*N*II*II/(N*D));
T1A=(N3A*II/(N*D)) + ((N*II*HFAC-N*II)/(N*D));
*
Compute  V  for   RM=rho-  < or =  RP=rho+   RP=(0, .3, .6, .9);
DO RP=0 TO .9 BY .3;
DO RM=0 TO .9 BY .3 WHILE(RM LE RP);
RPMRM=RP-RM;
RPPRM=RP+RM;
N1MN2SQ=(N1-N2)*(N1-N2);
IF N3A=0 AND NB=2 THEN
VMAR=((1-RP)*VMIR) + (.5*N*RPMRM) + (.5*RPPRM*N1MN2SQ/N);
IF N3A>0 AND NB=2 THEN
VMAR=((1-RP)*VMIR)
      + ( (.5*N*RPMRM)* (S1A*S1A+2*K*WII*S1A*T1A+K*T1A*T1A*GFAC) )
      + ( (N*RP)* (Q1A*Q1A+2*K*WII*Q1A*R1A+K*R1A*R1A*GFAC) )
      + ( (.5*N*RPPRM)* (P1A*P1A+2*K*WII*M1A*P1A+K*M1A*M1A*GFAC) );
IF DES='FAC' THEN
VMAR=((1-HFAC*RP)*VMIR) + (.5*N*RPMRM*S1A*S1A);
IF DESIGN='BBD4' THEN
VMAR=((1-RP)*VMIR) + (.5*N*K*T1A*T1A*GFAC*RPMRM)
      + (RP*( A/D-8*WII*II/D+(4*WII/(3*II))+(2*WIIJJ*R/(THE*II))
      +(4*N* (E1A*WIIII+2*F1A*WIIJJ+E1A*WIIJJ) )));
IF DESIGN='CCD2' THEN VACCD2B=VMAR;
IF DESIGN='CCD3' THEN VACCD3B=VMAR;
IF DESIGN='CCD4' THEN VACCD4B=VMAR;
IF DESIGN='CCD5' THEN VACCD5B=VMAR;
IF DESIGN='CCD6' THEN VACCD6B=VMAR;
IF DESIGN='CCD7' THEN VACCD7B=VMAR;
IF DESIGN='BBD4' THEN VABBD4B=VMAR;
IF DESIGN='BBD5' THEN VABBD5B=VMAR;
IF DESIGN='BBD7' THEN VABBD7B=VMAR;
IF DESIGN='FAC3' THEN VAFAC3B=VMAR;
IF DESIGN='FAC4' THEN VAFAC4B=VMAR;
IF DESIGN='FAC5' THEN VAFAC5B=VMAR;
IF DESIGN='FAC6' THEN VAFAC6B=VMAR;
IF DESIGN='FAC7' THEN VAFAC7B=VMAR;
OUTPUT;
END;
END;

DATA QQ1CCD2A; SET Q1CCD2A; IF VACCD2B='.' THEN DELETE;
DATA QQ1CCD3A; SET Q1CCD3A; IF VACCD3B='.' THEN DELETE;
```

```
DATA QQ1CCD4A; SET Q1CCD4A; IF VACCD4B='.' THEN DELETE;
DATA QQ1CCD5A; SET Q1CCD5A; IF VACCD5B='.' THEN DELETE;
DATA QQ1CCD6A; SET Q1CCD6A; IF VACCD6B='.' THEN DELETE;
DATA QQ1CCD7A; SET Q1CCD7A; IF VACCD7B='.' THEN DELETE;
DATA QQ1BBD4A; SET Q1BBD4A; IF VABBD4B='.' THEN DELETE;
DATA QQ1BBD5A; SET Q1BBD5A; IF VABBD5B='.' THEN DELETE;
DATA QQ1BBD7A; SET Q1BBD7A; IF VABBD7B='.' THEN DELETE;
DATA QQ1FAC3A; SET Q1FAC3A; IF VAFAC3B='.' THEN DELETE;
DATA QQ1FAC4A; SET Q1FAC4A; IF VAFAC4B='.' THEN DELETE;
DATA QQ1FAC5A; SET Q1FAC5A; IF VAFAC5B='.' THEN DELETE;
DATA QQ1FAC6A; SET Q1FAC6A; IF VAFAC6B='.' THEN DELETE;
DATA QQ1FAC7A; SET Q1FAC7A; IF VAFAC7B='.' THEN DELETE;

DATA Q1CCDA; MERGE QQ1CCD2A QQ1CCD3A QQ1CCD4A QQ1CCD5A QQ1CCD6A QQ1CCD7A;
                                                                  BY RP RM;
DATA Q1BBDA; MERGE QQ1BBD4A QQ1BBD5A QQ1BBD7A;                    BY RP RM;
DATA Q1FACA; MERGE QQ1FAC3A QQ1FAC4A QQ1FAC5A QQ1FAC6A QQ1FAC7A;BY RP RM;

PROC PRINT NOOBS DATA=Q1CCDA;
TITLE6 ' V  for the CCDs under the AR Strategy';

PROC PRINT NOOBS DATA=Q1BBDA;
TITLE6 ' V  for the BBDs under the AR Strategy';

PROC PRINT NOOBS DATA=Q1FACA;
TITLE6 ' V  for the FACs under the AR Strategy';
```

Min-V | Min-B Designs in a Spherical Region

Second Order Designs and values of Design Parameters

DES	K	F	AX	R	ALP	IIFAC	NB	DELTA
CCD	2	4	4	3.00	1.41421356	1.00	2	-.000
CCD	3	8	6	3.42	1.76383421	1.00	2	0.417
CCD	4	16	8	3.24	2.05798302	1.00	2	1.384
CCD	5	32	10	2.78	2.30940108	1.00	2	3.264
CCD	6	32	12	3.48	2.51058366	1.00	2	2.549
CCD	7	64	14	2.70	2.71746488	1.00	2	5.696
BBD	4	24	.	3.00	.	0.50	3	0.000
BBD	5	40	.	4.00	.	0.40	2	0.000
BBD	7	56	.	3.00	.	0.43	2	0.000
FAC	3	27	.	1.50	.	0.67	3	0.000
FAC	4	81	.	1.50	.	0.67	3	0.000
FAC	5	243	.	1.50	.	0.67	3	0.000
FAC	6	243	.	1.50	.	0.67	3	0.000
FAC	7	729	.	1.50	.	0.67	3	0.000
SCD	3	4	6	3.88	1.54919334	1.00	2	-.214
SCD	4	8	8	6.06	2.12132034	1.00	2	-.176
SCD	5	12	10	6.04	2.34520788	1.00	2	0.148
SCD	6	16	12	6.28	2.54950976	1.00	2	0.386
SCD	7	28	14	4.75	2.69098111	1.00	2	1.562

THETA & G for Min-B and [ii] & Nc for Min-V|Min-B Designs

DESIGN	THE	G	IIOPT	NCOPT	NC
CCD2	0.50000000	0.57735027	0.2649	2.07	2
CCD3	0.45069964	0.48404414	0.2079	2.03	2
CCD4	0.38955100	0.42867040	0.1713	2.25	2
CCD5	0.31876138	0.39115913	0.1458	2.76	3
CCD6	0.32846937	0.36257003	0.1270	2.16	2
CCD7	0.25418701	0.34013415	0.1126	2.96	3
BBD4	0.37500000	0.61237244	0.1713	2.27	2
BBD5	0.38888889	0.62360956	0.1459	2.64	3
BBD7	0.27272727	0.52223297	0.1125	2.16	2
FAC3	0.30612245	0.55328334	0.2112	-0.90	0
FAC4	0.25000000	0.50000000	0.1733	-3.08	0
FAC5	0.21212121	0.46056619	0.1470	-9.28	0
FAC6	0.18461538	0.42966892	0.1278	-8.96	0
FAC7	0.16363636	0.40451992	0.1131	-25.6	0
SCD3	0.47133139	0.51696205	0.2081	1.30	1
SCD4	0.50172414	0.41935912	0.1722	1.36	1
SCD5	0.46795758	0.38529925	0.1464	1.33	1
SCD6	0.44542936	0.35851340	0.1274	1.26	1
SCD7	0.36133270	0.33988978	0.1127	1.56	2

Min-V | Min-B Designs in a Cuboidal Region

Second Order Designs and values of Design Parameters

DES	K	F	AX	R	ALP	IIFAC	NB	DELTA
CCD	2	4	4	3.00	1.41421356	1.00	2	-.000
CCD	3	8	6	3.42	1.76383421	1.00	2	0.739
CCD	4	16	8	3.24	2.05798302	1.00	2	2.952
CCD	5	32	10	2.78	2.30940108	1.00	2	8.114
CCD	6	32	12	3.48	2.51058366	1.00	2	7.223
CCD	7	64	14	2.70	2.71746488	1.00	2	18.1
BBD	4	24	.	3.00	.	0.50	3	0.000
BBD	5	40	.	4.00	.	0.40	2	0.000
BBD	7	56	.	3.00	.	0.43	2	0.000
FAC	3	27	.	1.50	.	0.67	3	0.000
FAC	4	81	.	1.50	.	0.67	3	0.000
FAC	5	243	.	1.50	.	0.67	3	0.000
FAC	6	243	.	1.50	.	0.67	3	0.000
FAC	7	729	.	1.50	.	0.67	3	0.000
SCD	3	4	6	4.92	1.67332005	1.00	2	-.703
SCD	4	8	8	5.00	2.00000000	1.00	2	0.000
SCD	5	12	10	5.30	2.25320285	1.00	2	0.709
SCD	6	16	12	5.68	2.47338777	1.00	2	1.410
SCD	7	28	14	4.75	2.69098111	1.00	2	4.962

THETA & G for Min-B and [ii] & Nc for Min-V|Min-B Designs

DESIGN	THE	G	IIOPT	NCOPT	NC
CCD2	0.70000000	0.68313005	0.3638	2.26	2
CCD3	0.79924070	0.64458494	0.3609	2.38	2
CCD4	0.83104213	0.62611320	0.3576	2.83	3
CCD5	0.79234973	0.61670728	0.3548	3.74	4
CCD6	0.93066320	0.61029630	0.3533	3.03	3
CCD7	0.80774983	0.60633412	0.3512	4.45	4
BBD4	0.80000000	0.89442719	0.3573	2.87	3
BBD5	0.96666667	0.98319208	0.3558	3.47	3
BBD7	0.86666667	0.93094934	0.3514	3.20	3
FAC3	0.54285714	0.73678840	0.3622	-0.02	0
FAC4	0.53333333	0.73029674	0.3587	-0.72	0
FAC5	0.52727273	0.72613547	0.3560	-3.06	0
FAC6	0.52307692	0.72324057	0.3538	-3.48	0
FAC7	0.52000000	0.72111026	0.3520	-11	0
SCD3	0.90057803	0.66280190	0.3635	1.60	2
SCD4	1.00000000	0.63245553	0.3598	1.79	2
SCD5	1.10142160	0.61964446	0.3570	1.82	2
SCD6	1.20531432	0.61204162	0.3548	1.81	2
SCD7	1.14823503	0.60589849	0.3523	2.27	2

Min-V | Min-B Designs in a Spherical Region

Check for a SINGULAR X"X matrix

DESIGN	SINGULAR	IIJJ	IIJJ_SIN
CCD2	No	0.087	0.017
CCD3	No	0.086	0.057
CCD4	No	0.091	0.070
CCD5	No	0.101	0.086
CCD6	No	0.094	0.084
CCD7	No	0.105	0.098
BBD4	No	0.095	0.074
BBD5	No	0.087	0.075
BBD7	No	0.102	0.094
FAC3	No	0.131	0.098
FAC4	No	0.126	0.105
FAC5	No	0.124	0.108
FAC6	No	0.122	0.109
FAC7	No	0.120	0.110
SCD3	.	0.064	.
SCD4	.	0.071	.
SCD5	.	0.074	.
SCD6	.	0.075	.
SCD7	.	0.086	.

Actual Values of [ii], Nc, N1, N2, N3A, N3C, N, V-IR

DESIGN	II	NC	N1	N2	N3A	N3C	N	VMIR
CCD2	0.2667	2	5	5	0	1	10	5.9375
CCD3	0.2083	2	9	7	0	1	16	9.9619
CCD4	0.1729	2	17	9	0	1	26	14.9484
CCD5	0.1451	3	33	11	1	2	45	20.9450
CCD6	0.1275	2	33	13	0	1	46	28.0710
CCD7	0.1125	3	65	15	1	2	81	36.0031
BBD4	0.1731	2	8	8	10	1	26	14.9259
BBD5	0.1447	3	20	20	1	2	43	21.3851
BBD7	0.1129	2	28	28	0	1	58	35.8920
FAC3	0.2041	0	9	9	9	0	27	12.5860
FAC4	0.1667	0	27	27	27	0	81	19.5000
FAC5	0.1414	0	81	81	81	0	243	27.5962
FAC6	0.1231	0	81	81	81	0	243	36.8037
FAC7	0.1091	0	243	243	243	0	729	47.0833
SCD3	0.2138	1	5	6	0	0	11	.
SCD4	0.1759	1	8	9	0	0	17	.
SCD5	0.1485	1	12	11	0	0	23	.
SCD6	0.1285	1	16	13	0	0	29	.
SCD7	0.1115	2	29	15	0	1	44	.

Min-V | Min-B Designs in a Cuboidal Region

Check for a SINGULAR X"X matrix

DESIGN	SINGULAR	IIJJ	IIJJ_SIN
CCD2	No	0.087	0.500
CCD3	No	0.086	0.600
CCD4	No	0.091	0.667
CCD5	No	0.101	0.714
CCD6	No	0.094	0.750
CCD7	No	0.105	0.778
BBD4	No	0.095	0.667
BBD5	No	0.087	0.714
BBD7	No	0.102	0.778
FAC3	No	0.131	0.600
FAC4	No	0.126	0.667
FAC5	No	0.124	0.714
FAC6	No	0.122	0.750
FAC7	No	0.120	0.778
SCD3	No	0.064	0.600
SCD4	No	0.071	0.667
SCD5	No	0.074	0.714
SCD6	No	0.075	0.750
SCD7	No	0.086	0.778

Actual Values of [ii], Nc, N1, N2, N3A, N3C, N, V-IR

DESIGN	II	NC	N1	N2	N3A	N3C	N	VMIR
CCD2	0.3733	2	5	5	0	1	10	5.8929
CCD3	0.3693	2	9	7	0	1	16	10.0035
CCD4	0.3553	3	17	9	1	2	27	14.8086
CCD5	0.3528	4	33	11	2	3	46	20.2724
CCD6	0.3535	3	33	13	1	2	47	27.9713
CCD7	0.3532	4	65	15	2	3	82	34.7214
BBD4	0.3556	3	8	8	11	2	27	14.6250
BBD5	0.3597	3	20	20	1	2	43	21.6832
BBD7	0.3525	3	28	28	1	2	59	35.1100
FAC3	0.3619	0	9	9	9	0	27	10.4176
FAC4	0.3556	0	27	27	27	0	81	15.6797
FAC5	0.3515	0	81	81	81	0	243	21.9542
FAC6	0.3487	0	81	81	81	0	243	29.2359
FAC7	0.3467	0	243	243	243	0	729	37.5222
SCD3	0.3514	2	5	7	0	1	12	.
SCD4	0.3556	2	9	9	0	1	18	.
SCD5	0.3544	2	13	11	0	1	24	.
SCD6	0.3526	2	17	13	0	1	30	.
SCD7	0.3545	2	29	15	0	1	44	.

Min-V | Min-B Designs

Spherical Region

V for the CCDs under the CR Strategy

RP	VCCCD2B	VCCCD3B	VCCCD4B	VCCCD5B	VCCCD6B	VCCCD7B
0.0	5.93750	9.9619	14.9484	20.9450	28.0710	36.0031
0.3	6.78125	11.3893	17.8306	27.4269	33.0185	48.7023
0.6	7.62500	12.8167	20.7128	33.9088	37.9660	61.4014
0.9	8.46875	14.2442	23.5949	40.3907	42.9135	74.1006

V for the BBDs under the CR Strategy

RP	VCBBD4B	VCBBD5B	VCBBD7B
0.0	14.9259	21.3851	35.8920
0.3	17.8148	27.1675	42.0948
0.6	20.7037	32.9500	48.2975
0.9	23.5926	38.7324	54.5003

V for the FACs under the CR Strategy

RP	VCFAC3B	VCFAC4B	VCFAC5B	VCFAC6B	VCFAC7B
0.0	12.5860	19.50	27.596	36.804	47.083
0.3	16.9102	37.95	92.217	98.663	251.658
0.6	21.2344	56.40	156.838	160.521	456.233
0.9	25.5586	74.85	221.460	222.380	660.808

Min-V | Min-B Designs

Cuboidal Region

V for the CCDs under the CR Strategy

RP	VCCCD2B	VCCCD3B	VCCCD4B	VCCCD5B	VCCCD6B	VCCCD7B
0.0	5.89286	10.0035	14.8086	20.2724	27.9713	34.7214
0.3	6.69643	11.2972	17.5660	26.5632	32.5741	47.0128
0.6	7.50000	12.5909	20.3235	32.8540	37.1769	59.3042
0.9	8.30357	13.8846	23.0810	39.1448	41.7797	71.5956

V for the BBDs under the CR Strategy

RP	VCBBD4B	VCBBD5B	VCBBD7B
0.0	14.6250	21.6832	35.1100
0.3	17.4375	26.8920	41.0668
0.6	20.2500	32.1009	47.0235
0.9	23.0625	37.3097	52.9802

V for the FACs under the CR Strategy

RP	VCFAC3B	VCFAC4B	VCFAC5B	VCFAC6B	VCFAC7B
0.0	10.4176	15.6797	21.954	29.236	37.522
0.3	15.3923	35.2758	88.268	93.365	244.966
0.6	20.3670	54.8719	154.582	157.494	452.409
0.9	25.3418	74.4680	220.895	221.624	659.852

Min-V | Min-B Designs in a Spherical Region

V for the CCDs under the AR Strategy

RP	RM	VACCD2B	VACCD3B	VACCD4B	VACCD5B	VACCD6B	VACCD7B
0.0	0.0	5.93750	9.96186	14.9484	20.9450	28.0710	36.0031
0.3	0.0	5.65625	9.41080	14.7331	22.8395	27.8540	41.7693
0.3	0.3	4.15625	7.04830	11.2024	18.0046	22.2584	34.5666
0.6	0.0	5.37500	8.85974	14.5178	24.7340	27.6371	47.5354
0.6	0.3	3.87500	6.49724	10.9871	19.8991	22.0414	40.3328
0.6	0.6	2.37500	4.13474	7.4563	15.0642	16.4458	33.1302
0.9	0.0	5.09375	8.30869	14.3025	26.6284	27.4201	53.3016
0.9	0.3	3.59375	5.94619	10.7718	21.7935	21.8245	46.0989
0.9	0.6	2.09375	3.58369	7.2410	16.9586	16.2288	38.8963
0.9	0.9	0.59375	1.22119	3.7102	12.1237	10.6332	31.6937

V for the BBDs under the AR Strategy

RP	RM	VABBD4B	VABBD5B	VABBD7B
0.0	0.0	14.9259	21.3851	35.8920
0.3	0.0	13.7704	21.2440	33.8244
0.3	0.3	12.2296	15.0866	25.1244
0.6	0.0	12.6148	21.1030	31.7568
0.6	0.3	11.0741	14.9455	23.0568
0.6	0.6	9.5333	8.7880	14.3568
0.9	0.0	11.4593	20.9620	29.6892
0.9	0.3	9.9185	14.8045	20.9892
0.9	0.6	8.3778	8.6470	12.2892
0.9	0.9	6.8370	2.4895	3.5892

V for the FACs under the AR Strategy

RP	RM	VAFAC3B	VAFAC4B	VAFAC5B	VAFAC6B	VAFAC7B
0.0	0.0	12.5860	19.5	27.5962	36.8037	47.083
0.3	0.0	11.8688	21.0	38.2770	45.6430	86.267
0.3	0.3	10.0688	15.6	22.0770	29.4430	37.667
0.6	0.0	11.1516	22.5	48.9577	54.4822	125.450
0.6	0.3	9.3516	17.1	32.7577	38.2822	76.850
0.6	0.6	7.5516	11.7	16.5577	22.0822	28.250
0.9	0.0	10.4344	24.0	59.6385	63.3215	164.633
0.9	0.3	8.6344	18.6	43.4385	47.1215	116.033
0.9	0.6	6.8344	13.2	27.2385	30.9215	67.433
0.9	0.9	5.0344	7.8	11.0385	14.7215	18.833

Min-V | Min-B Designs in a Cuboidal Region

V for the CCDs under the AR Strategy

RP	RM	VACCD2B	VACCD3B	VACCD4B	VACCD5B	VACCD6B	VACCD7B
0.0	0.0	5.89286	10.0035	14.8086	20.2724	27.9713	34.7214
0.3	0.0	5.62500	9.4400	14.5391	22.1623	27.5996	40.4114
0.3	0.3	4.12500	7.0775	11.2105	17.5198	22.2538	33.4894
0.6	0.0	5.35714	8.8764	14.2696	24.0523	27.2278	46.1015
0.6	0.3	3.85714	6.5139	10.9410	19.4098	21.8821	39.1795
0.6	0.6	2.35714	4.1514	7.6123	14.7672	16.5363	32.2575
0.9	0.0	5.08929	8.3129	14.0002	25.9423	26.8561	51.7915
0.9	0.3	3.58929	5.9504	10.6715	21.2997	21.5103	44.8695
0.9	0.6	2.08929	3.5879	7.3428	16.6572	16.1645	37.9475
0.9	0.9	0.58929	1.2254	4.0142	12.0146	10.8188	31.0255

V for the BBDs under the AR Strategy

RP	RM	VABBD4B	VABBD5B	VABBD7B
0.0	0.0	14.625	21.6832	35.1100
0.3	0.0	13.500	21.3317	33.1245
0.3	0.3	12.000	15.3760	24.7787
0.6	0.0	12.375	20.9802	31.1389
0.6	0.3	10.875	15.0244	22.7932
0.6	0.6	9.375	9.0687	14.4474
0.9	0.0	11.250	20.6287	29.1533
0.9	0.3	9.750	14.6729	20.8076
0.9	0.6	8.250	8.7172	12.4619
0.9	0.9	6.750	2.7614	4.1161

V for the FACs under the AR Strategy

RP	RM	VAFAC3B	VAFAC4B	VAFAC5B	VAFAC6B	VAFAC7B
0.0	0.0	10.4176	15.6797	21.9542	29.2359	37.522
0.3	0.0	10.1341	17.9437	33.7634	39.5888	78.618
0.3	0.3	8.3341	12.5437	17.5634	23.3888	30.018
0.6	0.0	9.8506	20.2078	45.5725	49.9416	119.713
0.6	0.3	8.0506	14.8078	29.3725	33.7416	71.113
0.6	0.6	6.2506	9.4078	13.1725	17.5416	22.513
0.9	0.0	9.5670	22.4719	57.3817	60.2944	160.809
0.9	0.3	7.7670	17.0719	41.1817	44.0944	112.209
0.9	0.6	5.9670	11.6719	24.9817	27.8944	63.609
0.9	0.9	4.1670	6.2719	8.7817	11.6944	15.009

Appendix R

Computer Programs for Determination of Min-V* | Min-B* Second Order Designs

```
//A0392JMD JOB 33F02,JMD,MSGLEVEL=(0,0),REGION=2048K
//*
//*
//*   This program computes the Min-V*|Min-B* values of V*
//*                                 ------------        --
//*              for various amounts of induced correlation
//*              under the IR, CR, and AR strategies.
//*
//*   The SECOND order designs examined are:
//*              k = 2,...,7 CCDs (central composite designs)
//*              k = 4,5,7   BBDs (Box-Behnken designs)
//*              k = 3,...,7 FACs (3-level factorial designs)
//*              k = 3,...,7 SCDs (small composite designs)
//*
//*   This program is not entirely applicable to the SCDs
//*              but the "optimal" values of the scaling
//*              factor (g) and the number of center runs (Nc)
//*              are computed for the SCDs using this program.
//*
//*
/*JOBPARM LINES=4
/*PRIORITY IDLE
//STEP1   EXEC SAS
//SYSIN DD *
OPTIONS CENTER NODATE NONUMBER LS=72;
TITLE1;
TITLE2 'Min-V* | Min-B*  designs';
DATA Q2;
INPUT DES$ 1-3 DESIGN$ 1-4 DES2$ 2-3 K 4 F 6-8;
*
Define  REGION=1 if Spherical or REGION=2 if Cuboidal;
REGION=1;
TITLE4 'Cuboidal Region';
```

```
TITLE4 'Spherical Region';
*
DES = CCD or BBD or FAC or SCD
DESIGN = CCD2 ... CCD7, BBD4, BBD5, BBD7, FAC3 ... FAC7, SCD3 ... SCD7
K = Number of Factors in Model
F = Number of Factorial Design Points;
*
Define  AX = Number of axial design points;
AX=0; IF DES2='CD' THEN AX=2*K;
*
Define  NB = Number of blocks into which the design partitions;
NB=2; IF DES='FAC' OR DESIGN='BBD4' THEN NB=3;
*
Define Region moments;
IF REGION=1 THEN DO;
WII=1/(K+2);
WIIJJ=WII/(K+4);
WIIII=3*WIIJJ;
WIIJJKK=WIIJJ/(K+6);
WIIIIJJ=3*WIIJJKK;
WIIIIII=15*WIIJJKK;
END;
IF REGION=2 THEN DO;
WII=1/3;
WIIJJ=1/9;
WIIII=1/5;
WIIJJKK=1/27;
WIIIIJJ=1/15;
WIIIIII=1/7;
END;
*
Compute ALPHA for the CCDs and SCDs;
N=F+AX+2;
ALP=0;
IF DES2='CD' THEN ALP=SQRT((F*N-F*F-F)/(2*F+2));
ALP22=2*ALP*ALP;
ALP24=ALP22*ALP*ALP;
FPALP22=F+ALP22;
FPALP24=F+ALP24;
*
Define  IIFAC = proportion of non-zero factorial levels of i'th point;
IIFAC=1;
IF DES='FAC' THEN IIFAC=2/3;
IF DESIGN='BBD4' THEN IIFAC=1/2;
IF DESIGN='BBD5' THEN IIFAC=2/5;
IF DESIGN='BBD7' THEN IIFAC=3/7;
*
Compute  R = [iiii] / [iijj] ;
R=(FPALP24/F)/IIFAC;
IF DES='BBD' THEN R=3;
IF DESIGN='BBD5' THEN R=4;
*
Determine the Min-B* Optimal Values of  THE=Theta  &  G=Scaling factor;
KM1=K-1;
KP2=K+2;
RM1=R-1;
RK1=R+K-1;
RK2=R+K-2;
THE=R*WII*KP2/RK1;
THE2=THE*THE;
G=SQRT(THE*FPALP22/FPALP24);
*
Determine the Min-V*|Min-B* Optimal Values of  IIOPT=[ii]  &  NCOPT=Nc;
A1=K*KM1*WII*R;
B1=THE*R*RK2;
C1=R*R*KM1;
D1=THE2*RM1*RK1;
```

```
E1=R*K*THE*RM1;
F1=(K*THE+A1)/(4*K*THE*WII);
G1=C1*E1+F1*E1*E1;
H1=(-2)*(B1*E1+D1*E1*F1);
I1=B1*D1+F1*D1*D1;
INSQRT=(H1*H1-4*G1*I1);
IIOPT=(H1*(-1)-SQRT(INSQRT))/(2*G1);
NCOPT=(IIFAC*G*G*FPALP22/IIOPT)-F-AX;
*
Determine the actual value of II=[ii] using the integer value of NC=Nc;
NC=ROUND(NCOPT);
IF DES='FAC' THEN NC=0;
N=F+AX+NC;
II=IIFAC*G*G*FPALP22/N;
II2=II*II;
*
Determine IIJJ=[iijj] and check for a singular X'X matrix;
IIJJ=THE*II/R;
IIJJ_SIN=(II/KM1)*(K*II-THE);
IF IIJJ=IIJJ_SIN THEN SINGULAR='Yes';
ELSE SINGULAR='No';
IF DES='SCD' THEN IIJJ_SIN='.';
IF DES='SCD' THEN SINGULAR='.';
*
Determine DELTA = the difference between blocks 1 & 2 sum of squares;
DELTA=0;
IF DES2='CD' THEN DELTA=G*G*(F-ALP22);
*
Determine the components of the inverted X'X matrix;
A=THE*II*RK1/R;
B=(THE*II*RK2/R)-(KM1*II2);
C=(II/R)*(R*II-THE);
D=A-K*II2;
D2=D*THE*II*RM1/R;
*
Determine the value of V* under the  IR Strategy;
VMIRS=(K/II)+((K*KM1*R*WII)/(THE*II))+(4*B*K*WII/D2);
*
Determine N1 & N2 = Number design points in blocks 1 & 2 (AR Strategy);
N1=F/NB;
N2=N1;
IF DES2='CD' THEN N1=F+1;
IF DES2='CD' THEN N2=AX+1;
*
Determine N3C & N3A = Number of design points using independent
                          streams in the  CR  &  AR Strategies;
N3C=0;
IF NC>1 THEN N3C=NC-1;
IF N3C=0 THEN SQRTN3C=0;
ELSE SQRTN3C=SQRT(N3C);
N3A=0;
IF NC>2 THEN N3A=NC-2;
IF DESIGN='BBD4' OR DES='FAC' THEN N3A=N1+NC;
IF N3A=0 THEN SQRTN3A=0;
ELSE SQRTN3A=SQRT(N3A);

FORMAT THE G ALP 10.8;
FORMAT IIOPT II 6.4;
FORMAT DELTA IIJJ IIJJ_SIN 5.3;
FORMAT NCOPT 5.2;
FORMAT IIFAC R 4.2;
CARDS;
CCD2   4
CCD3   8
CCD4  16
CCD5  32
CCD6  32
```

```
CCD7  64
BBD4  24
BBD5  40
BBD7  56
FAC3  27
FAC4  81
FAC5 243
FAC6 243
FAC7 729
SCD3   4
SCD4   8
SCD5  12
SCD6  16
SCD7  28
;
*
Print Input Data,  Optimal values of Theta, g, [ii], Nc, and
      Actual values of [ii], Nc, N1, N2, N3, N;

DATA Q2DOT;
SET Q2;
IF DES2 NE 'CD' THEN ALP='.';
IF DES2 NE 'CD' THEN AX='.';
IF DES='SCD' THEN VMIRS='.';
PROC PRINT NOOBS DATA=Q2DOT;
TITLE6 'Second Order Designs  and values of Design Parameters';
VAR DES K F AX R ALP IIFAC NB DELTA;

PROC PRINT NOOBS DATA=Q2;
TITLE6
'THETA & G  for Min-B*   and   [ii] & Nc  for Min-V*|Min-B*  Designs';
VAR DESIGN THE G IIOPT NCOPT NC;

PROC PRINT NOOBS DATA=Q2;
TITLE6 'Check for a SINGULAR X"X matrix';
VAR DESIGN SINGULAR IIJJ IIJJ_SIN;

PROC PRINT NOOBS DATA=Q2DOT;
TITLE6 'Actual Values of [ii], Nc, N1, N2, N3A, N3C, N, V*-IR';
VAR DESIGN II NC N1 N2 N3A N3C N VMIRS;
*
*
Determine the value of V* under the  CR Strategy;
DATA Q2VCR
           DATA Q2CCD2C (KEEP=RP VCCCD2S)
           DATA Q2CCD3C (KEEP=RP VCCCD3S)
           DATA Q2CCD4C (KEEP=RP VCCCD4S)
           DATA Q2CCD5C (KEEP=RP VCCCD5S)
           DATA Q2CCD6C (KEEP=RP VCCCD6S)
           DATA Q2CCD7C (KEEP=RP VCCCD7S)
           DATA Q2BBD4C (KEEP=RP VCBBD4S)
           DATA Q2BBD5C (KEEP=RP VCBBD5S)
           DATA Q2BBD7C (KEEP=RP VCBBD7S)
           DATA Q2FAC3C (KEEP=RP VCFAC3S)
           DATA Q2FAC4C (KEEP=RP VCFAC4S)
           DATA Q2FAC5C (KEEP=RP VCFAC5S)
           DATA Q2FAC6C (KEEP=RP VCFAC6S)
           DATA Q2FAC7C (KEEP=RP VCFAC7S);
SET Q2;
IF DES='SCD' THEN DELETE;
*
Define components of V* for the  CR Strategy;
R1C=(-1)*(SQRTN3C*II)/(N*D);
T1C=(N3C*II)/(N*D);
*
Compute V* for   RP=rho+   RP=(0, .3, .6, .9);
DO RP=0 TO .9 BY .3;
```

```
IF N3C=0 THEN VMCRS=(1-RP)*VMIRS;
ELSE VMCRS=((1-RP)*VMIRS)+((4*K*WII*N*RP)*(T1C*T1C+R1C*R1C));
IF DESIGN='CCD2' THEN VCCCD2S=VMCRS;
IF DESIGN='CCD3' THEN VCCCD3S=VMCRS;
IF DESIGN='CCD4' THEN VCCCD4S=VMCRS;
IF DESIGN='CCD5' THEN VCCCD5S=VMCRS;
IF DESIGN='CCD6' THEN VCCCD6S=VMCRS;
IF DESIGN='CCD7' THEN VCCCD7S=VMCRS;
IF DESIGN='BBD4' THEN VCBBD4S=VMCRS;
IF DESIGN='BBD5' THEN VCBBD5S=VMCRS;
IF DESIGN='BBD7' THEN VCBBD7S=VMCRS;
IF DESIGN='FAC3' THEN VCFAC3S=VMCRS;
IF DESIGN='FAC4' THEN VCFAC4S=VMCRS;
IF DESIGN='FAC5' THEN VCFAC5S=VMCRS;
IF DESIGN='FAC6' THEN VCFAC6S=VMCRS;
IF DESIGN='FAC7' THEN VCFAC7S=VMCRS;
OUTPUT;
END;

DATA QQ2CCD2C; SET Q2CCD2C; IF VCCCD2S='.' THEN DELETE;
DATA QQ2CCD3C; SET Q2CCD3C; IF VCCCD3S='.' THEN DELETE;
DATA QQ2CCD4C; SET Q2CCD4C; IF VCCCD4S='.' THEN DELETE;
DATA QQ2CCD5C; SET Q2CCD5C; IF VCCCD5S='.' THEN DELETE;
DATA QQ2CCD6C; SET Q2CCD6C; IF VCCCD6S='.' THEN DELETE;
DATA QQ2CCD7C; SET Q2CCD7C; IF VCCCD7S='.' THEN DELETE;
DATA QQ2BBD4C; SET Q2BBD4C; IF VCBBD4S='.' THEN DELETE;
DATA QQ2BBD5C; SET Q2BBD5C; IF VCBBD5S='.' THEN DELETE;
DATA QQ2BBD7C; SET Q2BBD7C; IF VCBBD7S='.' THEN DELETE;
DATA QQ2FAC3C; SET Q2FAC3C; IF VCFAC3S='.' THEN DELETE;
DATA QQ2FAC4C; SET Q2FAC4C; IF VCFAC4S='.' THEN DELETE;
DATA QQ2FAC5C; SET Q2FAC5C; IF VCFAC5S='.' THEN DELETE;
DATA QQ2FAC6C; SET Q2FAC6C; IF VCFAC6S='.' THEN DELETE;
DATA QQ2FAC7C; SET Q2FAC7C; IF VCFAC7S='.' THEN DELETE;

DATA Q2CCDC; MERGE QQ2CCD2C QQ2CCD3C QQ2CCD4C QQ2CCD5C QQ2CCD6C QQ2CCD7C;
                                                                     BY RP;
DATA Q2BBDC; MERGE QQ2BBD4C QQ2BBD5C QQ2BBD7C;                       BY RP;
DATA Q2FACC; MERGE QQ2FAC3C QQ2FAC4C QQ2FAC5C QQ2FAC6C QQ2FAC7C;     BY RP;

PROC PRINT NOOBS DATA=Q2CCDC;
TITLE6 'V* for the CCDs under the CR Strategy';

PROC PRINT NOOBS DATA=Q2BBDC;
TITLE6 'V* for the BBDs under the CR Strategy';

PROC PRINT NOOBS DATA=Q2FACC;
TITLE6 'V* for the FACs under the CR Strategy';
*
*
Determine the value of V* under the  AR Strategy;
DATA Q2VAR
           DATA Q2CCD2A (KEEP=RP RM VACCD2S)
           DATA Q2CCD3A (KEEP=RP RM VACCD3S)
           DATA Q2CCD4A (KEEP=RP RM VACCD4S)
           DATA Q2CCD5A (KEEP=RP RM VACCD5S)
           DATA Q2CCD6A (KEEP=RP RM VACCD6S)
           DATA Q2CCD7A (KEEP=RP RM VACCD7S)
           DATA Q2BBD4A (KEEP=RP RM VABBD4S)
           DATA Q2BBD5A (KEEP=RP RM VABBD5S)
           DATA Q2BBD7A (KEEP=RP RM VABBD7S)
           DATA Q2FAC3A (KEEP=RP RM VAFAC3S)
           DATA Q2FAC4A (KEEP=RP RM VAFAC4S)
           DATA Q2FAC5A (KEEP=RP RM VAFAC5S)
           DATA Q2FAC6A (KEEP=RP RM VAFAC6S)
           DATA Q2FAC7A (KEEP=RP RM VAFAC7S);
SET Q2;
IF DES='SCD' THEN DELETE;
```

```
*
Define components of V* for the  AR Strategy;
HFAC=2/NB;
E1A=((1-HFAC)*B/(N*D2))+((HFAC*NC*II2)/(N*N*D*D));
M1A=(N2*II-N1*II+DELTA)/(N*D);
R1A=(-1)*(SQRTN3A*II)/(N*D);
T1A=(N3A*II+N*II*HFAC-N*II)/(N*D);
*
Compute V* for   RM=rho-  < or =  RP=rho+   RP=(0, .3, .6, .9);
DO RP=0 TO .9 BY .3;
IF DES='FAC' THEN VMARS=(1-HFAC*RP)*VMIRS;
DO RM=0 TO .9 BY .3 WHILE(RM LE RP);
RPMRM=RP-RM;
RPPRM=RP+RM;
IF N3A=0 AND NB=2 THEN VMARS=(1-RP)*VMIRS;
IF N3A>0 AND NB=2 THEN VMARS=((1-RP)*VMIRS)
       +((4*K*WII)*(.5*N*T1A*T1A*RPMRM+N*R1A*R1A*RP+.5*N*M1A*M1A*RPPRM));
IF DESIGN='BBD4' THEN VMARS=((1-RP)*VMIRS)+(8*WII*N*T1A*T1A*RPMRM)
       +(RP*(4/(3*II)+(4*WII*R/(THE*II))+(16*N*WII*E1A)));
IF DESIGN='CCD2' THEN VACCD2S=VMARS;
IF DESIGN='CCD3' THEN VACCD3S=VMARS;
IF DESIGN='CCD4' THEN VACCD4S=VMARS;
IF DESIGN='CCD5' THEN VACCD5S=VMARS;
IF DESIGN='CCD6' THEN VACCD6S=VMARS;
IF DESIGN='CCD7' THEN VACCD7S=VMARS;
IF DESIGN='BBD4' THEN VABBD4S=VMARS;
IF DESIGN='BBD5' THEN VABBD5S=VMARS;
IF DESIGN='BBD7' THEN VABBD7S=VMARS;
IF DESIGN='FAC3' THEN VAFAC3S=VMARS;
IF DESIGN='FAC4' THEN VAFAC4S=VMARS;
IF DESIGN='FAC5' THEN VAFAC5S=VMARS;
IF DESIGN='FAC6' THEN VAFAC6S=VMARS;
IF DESIGN='FAC7' THEN VAFAC7S=VMARS;
IF DES NE 'FAC' THEN OUTPUT;
END;
IF DES='FAC' THEN RM='.';
IF DES='FAC' THEN OUTPUT;
END;

DATA QQ2CCD2A; SET Q2CCD2A; IF VACCD2S='.' THEN DELETE;
DATA QQ2CCD3A; SET Q2CCD3A; IF VACCD3S='.' THEN DELETE;
DATA QQ2CCD4A; SET Q2CCD4A; IF VACCD4S='.' THEN DELETE;
DATA QQ2CCD5A; SET Q2CCD5A; IF VACCD5S='.' THEN DELETE;
DATA QQ2CCD6A; SET Q2CCD6A; IF VACCD6S='.' THEN DELETE;
DATA QQ2CCD7A; SET Q2CCD7A; IF VACCD7S='.' THEN DELETE;
DATA QQ2BBD4A; SET Q2BBD4A; IF VABBD4S='.' THEN DELETE;
DATA QQ2BBD5A; SET Q2BBD5A; IF VABBD5S='.' THEN DELETE;
DATA QQ2BBD7A; SET Q2BBD7A; IF VABBD7S='.' THEN DELETE;
DATA QQ2FAC3A; SET Q2FAC3A; IF VAFAC3S='.' THEN DELETE;
DATA QQ2FAC4A; SET Q2FAC4A; IF VAFAC4S='.' THEN DELETE;
DATA QQ2FAC5A; SET Q2FAC5A; IF VAFAC5S='.' THEN DELETE;
DATA QQ2FAC6A; SET Q2FAC6A; IF VAFAC6S='.' THEN DELETE;
DATA QQ2FAC7A; SET Q2FAC7A; IF VAFAC7S='.' THEN DELETE;

DATA Q2CCDA; MERGE QQ2CCD2A QQ2CCD3A QQ2CCD4A QQ2CCD5A QQ2CCD6A QQ2CCD7A;
                                                                BY RP RM;
DATA Q2BBDA; MERGE QQ2BBD4A QQ2BBD5A QQ2BBD7A;                  BY RP RM;
DATA Q2FACA; MERGE QQ2FAC3A QQ2FAC4A QQ2FAC5A QQ2FAC6A QQ2FAC7A;BY RP;

PROC PRINT NOOBS DATA=Q2CCDA;
TITLE6 'V* for the CCDs under the AR Strategy';

PROC PRINT NOOBS DATA=Q2BBDA;
TITLE6 'V* for the BBDs under the AR Strategy';

PROC PRINT NOOBS DATA=Q2FACA;
TITLE6 'V* for the FACs under the AR Strategy';
```

Min-V* | Min-B* Designs in a Spherical Region

Second Order Designs and values of Design Parameters

DES	K	F	AX	R	ALP	IIFAC	NB	DELTA
CCD	2	4	4	3.00	1.41421356	1.00	2	-.000
CCD	3	8	6	3.42	1.76383421	1.00	2	0.583
CCD	4	16	8	3.24	2.05798302	1.00	2	1.845
CCD	5	32	10	2.78	2.30940108	1.00	2	4.197
CCD	6	32	12	3.48	2.51058366	1.00	2	3.187
CCD	7	64	14	2.70	2.71746488	1.00	2	6.961
BBD	4	24	.	3.00	.	0.50	3	0.000
BBD	5	40	.	4.00	.	0.40	2	0.000
BBD	7	56	.	3.00	.	0.43	2	0.000
FAC	3	27	.	1.50	.	0.67	3	0.000
FAC	4	81	.	1.50	.	0.67	3	0.000
FAC	5	243	.	1.50	.	0.67	3	0.000
FAC	6	243	.	1.50	.	0.67	3	0.000
FAC	7	729	.	1.50	.	0.67	3	0.000
SCD	3	4	6	4.92	1.67332005	1.00	2	-.555
SCD	4	8	8	5.00	2.00000000	1.00	2	0.000
SCD	5	12	10	5.30	2.25320285	1.00	2	0.367
SCD	6	16	12	5.68	2.47338777	1.00	2	0.622
SCD	7	28	14	4.75	2.69098111	1.00	2	1.909

THETA & G for Min-B* and [ii] & Nc for Min-V*|Min-B* Designs

DESIGN	THE	G	IIOPT	NCOPT	NC
CCD2	0.75000000	0.70710678	0.3629	3.02	3
CCD3	0.63097950	0.57272875	0.2688	3.35	3
CCD4	0.51940133	0.49498595	0.2134	4.09	4
CCD5	0.40983607	0.44353276	0.1769	5.44	5
CCD6	0.41058671	0.40536561	0.1511	4.52	5
CCD7	0.31067301	0.37603246	0.1317	6.58	7
BBD4	0.50000000	0.70710678	0.2134	4.12	4
BBD5	0.50000000	0.70710678	0.1772	5.15	5
BBD7	0.33333333	0.57735027	0.1317	4.76	5
FAC3	0.42857143	0.65465367	0.2768	0.87	0
FAC4	0.33333333	0.57735027	0.2183	1.45	0
FAC5	0.27272727	0.52223297	0.1800	2.42	0
FAC6	0.23076923	0.48038446	0.1531	1.23	0
FAC7	0.20000000	0.44721360	0.1331	1.36	0
SCD3	0.71098266	0.58891494	0.2707	2.30	2
SCD4	0.62500000	0.50000000	0.2147	2.63	3
SCD5	0.56970083	0.44564517	0.1779	2.73	3
SCD6	0.53175632	0.40652488	0.1518	2.74	3
SCD7	0.44162886	0.37576229	0.1320	3.45	3

Min-V* | Min-B* Designs in a Cuboidal Region

Second Order Designs and values of Design Parameters

DES	K	F	AX	R	ALP	IIFAC	NB	DELTA
CCD	2	4	4	3.00	1.41421356	1.00	2	-.000
CCD	3	8	6	3.42	1.76383421	1.00	2	0.972
CCD	4	16	8	3.24	2.05798302	1.00	2	3.690
CCD	5	32	10	2.78	2.30940108	1.00	2	9.792
CCD	6	32	12	3.48	2.51058366	1.00	2	8.498
CCD	7	64	14	2.70	2.71746488	1.00	2	20.88
BBD	4	24	.	3.00	.	0.50	3	0.000
BBD	5	40	.	4.00	.	0.40	2	0.000
BBD	7	56	.	3.00	.	0.43	2	0.000
FAC	3	27	.	1.50	.	0.67	3	0.000
FAC	4	81	.	1.50	.	0.67	3	0.000
FAC	5	243	.	1.50	.	0.67	3	0.000
FAC	6	243	.	1.50	.	0.67	3	0.000
FAC	7	729	.	1.50	.	0.67	3	0.000
SCD	3	4	6	4.92	1.67332005	1.00	2	-.925
SCD	4	8	8	5.00	2.00000000	1.00	2	0.000
SCD	5	12	10	5.30	2.25320285	1.00	2	0.856
SCD	6	16	12	5.68	2.47338777	1.00	2	1.659
SCD	7	28	14	4.75	2.69098111	1.00	2	5.726

THETA & G for Min-B* and [ii] & Nc for Min-V*|Min-B* Designs

DESIGN	THE	G	IIOPT	NCOPT	NC
CCD2	1.00000000	0.81649658	0.4838	3.02	3
CCD3	1.05163250	0.73938964	0.4480	3.35	3
CCD4	1.03880266	0.70001584	0.4269	4.09	4
CCD5	0.95628415	0.67750749	0.4129	5.44	5
CCD6	1.09489789	0.66195928	0.4028	4.52	5
CCD7	0.93201903	0.65130732	0.3950	6.58	7
BBD4	1.00000000	1.00000000	0.4268	4.12	4
BBD5	1.16666667	1.08012345	0.4135	5.15	5
BBD7	1.00000000	1.00000000	0.3950	4.76	5
FAC3	0.71428571	0.84515425	0.4614	0.87	0
FAC4	0.66666667	0.81649658	0.4366	1.45	0
FAC5	0.63636364	0.79772404	0.4201	2.42	0
FAC6	0.61538462	0.78446454	0.4082	1.23	0
FAC7	0.60000000	0.77459667	0.3993	1.36	0
SCD3	1.18497110	0.76028592	0.4511	2.30	2
SCD4	1.25000000	0.70710678	0.4294	2.63	3
SCD5	1.32930193	0.68073424	0.4151	2.73	3
SCD6	1.41801685	0.66385234	0.4048	2.74	3
SCD7	1.32488658	0.65083938	0.3959	3.45	3

Min-V* | Min-B* Designs in a Spherical Region

Check for a SINGULAR X"X matrix

DESIGN	SINGULAR	IIJJ	IIJJ_SIN
CCD2	No	0.091	-.008
CCD3	No	0.051	0.026
CCD4	No	0.034	0.024
CCD5	No	0.026	0.022
CCD6	No	0.018	0.015
CCD7	No	0.015	0.013
BBD4	No	0.036	0.026
BBD5	No	0.022	0.017
BBD7	No	0.015	0.013
FAC3	No	0.082	0.061
FAC4	No	0.049	0.041
FAC5	No	0.033	0.029
FAC6	No	0.024	0.021
FAC7	No	0.018	0.016
SCD3	.	0.040	.
SCD4	.	0.026	.
SCD5	.	0.019	.
SCD6	.	0.014	.
SCD7	.	0.012	.

Actual Values of [ii], Nc, N1, N2, N3A, N3C, N, V*-IR

DESIGN	II	NC	N1	N2	N3A	N3C	N	VMIRS
CCD2	0.3636	3	5	5	1	2	11	26.583
CCD3	0.2744	3	9	7	1	2	17	64.186
CCD4	0.2141	4	17	9	2	3	28	124.685
CCD5	0.1786	5	33	11	3	4	47	215.113
CCD6	0.1496	5	33	13	3	4	49	342.486
CCD7	0.1310	7	65	15	5	6	85	508.340
BBD4	0.2143	4	8	8	12	3	28	124.444
BBD5	0.1778	5	20	20	3	4	45	219.911
BBD7	0.1311	5	28	28	3	4	61	506.469
FAC3	0.2857	0	9	9	9	0	27	84.000
FAC4	0.2222	0	27	27	27	0	81	166.500
FAC5	0.1818	0	81	81	81	0	243	286.786
FAC6	0.1538	0	81	81	81	0	243	450.938
FAC7	0.1333	0	243	243	243	0	729	665.000
SCD3	0.2775	2	5	7	0	1	12	.
SCD4	0.2105	3	9	9	1	2	19	.
SCD5	0.1760	3	13	11	1	2	25	.
SCD6	0.1505	3	17	13	1	2	31	.
SCD7	0.1333	3	29	15	1	2	45	.

Min-V* | Min-B* Designs in a Cuboidal Region

Check for a SINGULAR X"X matrix

DESIGN	SINGULAR	IIJJ	IIJJ_SIN
CCD2	No	0.162	-.015
CCD3	No	0.141	0.073
CCD4	No	0.137	0.096
CCD5	No	0.143	0.117
CCD6	No	0.125	0.104
CCD7	No	0.135	0.119
BBD4	No	0.143	0.102
BBD5	No	0.121	0.094
BBD7	No	0.131	0.115
FAC3	No	0.227	0.170
FAC4	No	0.198	0.165
FAC5	No	0.180	0.157
FAC6	No	0.168	0.151
FAC7	No	0.160	0.147
SCD3	.	0.111	.
SCD4	.	0.105	.
SCD5	.	0.103	.
SCD6	.	0.100	.
SCD7	.	0.112	.

Actual Values of [ii], Nc, N1, N2, N3A, N3C, N, V*-IR

DESIGN	II	NC	N1	N2	N3A	N3C	N	VMIRS
CCD2	0.4848	3	5	5	1	2	11	19.938
CCD3	0.4574	3	9	7	1	2	17	38.511
CCD4	0.4283	4	17	9	2	3	28	62.342
CCD5	0.4167	5	33	11	3	4	47	92.191
CCD6	0.3989	5	33	13	3	4	49	128.432
CCD7	0.3931	7	65	15	5	6	85	169.447
BBD4	0.4286	4	8	8	12	3	28	62.222
BBD5	0.4148	5	20	20	3	4	45	94.247
BBD7	0.3934	5	28	28	3	4	61	168.823
FAC3	0.4762	0	9	9	9	0	27	50.400
FAC4	0.4444	0	27	27	27	0	81	83.250
FAC5	0.4242	0	81	81	81	0	243	122.908
FAC6	0.4103	0	81	81	81	0	243	169.102
FAC7	0.4000	0	243	243	243	0	729	221.667
SCD3	0.4624	2	5	7	0	1	12	.
SCD4	0.4211	3	9	9	1	2	19	.
SCD5	0.4106	3	13	11	1	2	25	.
SCD6	0.4014	3	17	13	1	2	31	.
SCD7	0.3999	3	29	15	1	2	45	.

Min-V* | Min-B* Designs

Spherical Region

V* for the CCDs under the CR Strategy

RP	VCCCD2S	VCCCD3S	VCCCD4S	VCCCD5S	VCCCD6S	VCCCD7S
0.0	26.5833	64.1856	124.685	215.113	342.486	508.340
0.3	23.0083	53.0652	103.931	182.389	274.717	423.182
0.6	19.4333	41.9448	83.178	149.665	206.947	338.025
0.9	15.8583	30.8244	62.424	116.941	139.178	252.867

V* for the BBDs under the CR Strategy

RP	VCBBD4S	VCBBD5S	VCBBD7S
0.0	124.444	219.911	506.469
0.3	103.911	184.795	400.075
0.6	83.378	149.679	293.681
0.9	62.844	114.562	187.287

V* for the FACs under the CR Strategy

RP	VCFAC3S	VCFAC4S	VCFAC5S	VCFAC6S	VCFAC7S
0.0	84.0	166.50	286.786	450.938	665.0
0.3	58.8	116.55	200.750	315.656	465.5
0.6	33.6	66.60	114.714	180.375	266.0
0.9	8.4	16.65	28.679	45.094	66.5

Min-V* | Min-B* Designs

Cuboidal Region

V* for the CCDs under the CR Strategy

RP	VCCCD2S	VCCCD3S	VCCCD4S	VCCCD5S	VCCCD6S	VCCCD7S
0.0	19.9375	38.5113	62.3423	92.1915	128.432	169.447
0.3	17.2562	31.8391	51.9656	78.1669	103.019	141.061
0.6	14.5750	25.1669	41.5889	64.1423	77.605	112.675
0.9	11.8938	18.4946	31.2122	50.1178	52.192	84.289

V* for the BBDs under the CR Strategy

RP	VCBBD4S	VCBBD5S	VCBBD7S
0.0	62.2222	94.2474	168.823
0.3	51.9556	79.1977	133.358
0.6	41.6889	64.1480	97.894
0.9	31.4222	49.0982	62.429

V* for the FACs under the CR Strategy

RP	VCFAC3S	VCFAC4S	VCFAC5S	VCFAC6S	VCFAC7S
0.0	50.40	83.250	122.908	169.102	221.667
0.3	35.28	58.275	86.036	118.371	155.167
0.6	20.16	33.300	49.163	67.641	88.667
0.9	5.04	8.325	12.291	16.910	22.167

Min-V* | Min-B* Designs in a Spherical Region

V* for the CCDs under the AR Strategy

RP	RM	VACCD2S	VACCD3S	VACCD4S	VACCD5S	VACCD6S	VACCD7S
0.0	0.0	26.5833	64.1856	124.685	215.113	342.486	508.340
0.3	0.0	19.7083	46.9743	93.093	164.298	254.344	391.727
0.3	0.3	19.3417	46.3070	90.580	158.930	247.962	379.514
0.6	0.0	12.8333	29.7631	61.501	113.482	166.202	275.115
0.6	0.3	12.4667	29.0957	58.988	108.114	159.820	262.901
0.6	0.6	12.1000	28.4284	56.476	102.746	153.438	250.687
0.9	0.0	5.9583	12.5518	29.909	62.666	78.060	158.502
0.9	0.3	5.5917	11.8844	27.396	57.298	71.678	146.289
0.9	0.6	5.2250	11.2171	24.884	51.930	65.296	134.075
0.9	0.9	4.8583	10.5498	22.371	46.562	58.914	121.861

V* for the BBDs under the AR Strategy

RP	RM	VABBD4S	VABBD5S	VABBD7S
0.0	0.0	124.444	219.911	506.469
0.3	0.0	108.267	165.509	371.609
0.3	0.3	103.289	158.566	361.361
0.6	0.0	92.089	111.107	236.748
0.6	0.3	87.111	104.164	226.500
0.6	0.6	82.133	97.221	216.252
0.9	0.0	75.911	56.705	101.887
0.9	0.3	70.933	49.762	91.639
0.9	0.6	65.956	42.820	81.391
0.9	0.9	60.978	35.877	71.143

V* for the FACs under the AR Strategy

RP	RM	VAFAC3S	VAFAC4S	VAFAC5S	VAFAC6S	VAFAC7S
0.0	.	84.0	166.5	286.786	450.938	665
0.3	.	67.2	133.2	229.429	360.750	532
0.6	.	50.4	99.9	172.071	270.563	399
0.9	.	33.6	66.6	114.714	180.375	266

Min-V* | Min-B* Designs in a Cuboidal Region

V* for the CCDs under the AR Strategy

RP	RM	VACCD2S	VACCD3S	VACCD4S	VACCD5S	VACCD6S	VACCD7S
0.0	0.0	19.9375	38.5113	62.3423	92.1915	128.432	169.447
0.3	0.0	14.7812	28.1846	46.5463	70.4132	95.379	130.576
0.3	0.3	14.5063	27.7842	45.2900	68.1127	92.986	126.505
0.6	0.0	9.6250	17.8578	30.7503	48.6349	62.326	91.705
0.6	0.3	9.3500	17.4574	29.4941	46.3344	59.933	87.634
0.6	0.6	9.0750	17.0570	28.2378	44.0339	57.539	83.562
0.9	0.0	4.4688	7.5311	14.9544	26.8567	29.273	52.834
0.9	0.3	4.1938	7.1307	13.6981	24.5561	26.879	48.763
0.9	0.6	3.9188	6.7303	12.4418	22.2556	24.486	44.692
0.9	0.9	3.6438	6.3299	11.1855	19.9550	22.093	40.620

V* for the BBDs under the AR Strategy

RP	RM	VABBD4S	VABBD5S	VABBD7S
0.0	0.0	62.2222	94.2474	168.823
0.3	0.0	54.1333	70.9324	123.870
0.3	0.3	51.6444	67.9569	120.454
0.6	0.0	46.0444	47.6173	78.916
0.6	0.3	43.5556	44.6418	75.500
0.6	0.6	41.0667	41.6663	72.084
0.9	0.0	37.9556	24.3023	33.962
0.9	0.3	35.4667	21.3268	30.546
0.9	0.6	32.9778	18.3513	27.130
0.9	0.9	30.4889	15.3758	23.714

V* for the FACs under the AR Strategy

RP	RM	VAFAC3S	VAFAC4S	VAFAC5S	VAFAC6S	VAFAC7S
0.0	.	50.40	83.25	122.908	169.102	221.667
0.3	.	40.32	66.60	98.327	135.281	177.333
0.6	.	30.24	49.95	73.745	101.461	133.000
0.9	.	20.16	33.30	49.163	67.641	88.667

Appendix S

Computer Programs for the Small Composite Designs

```
//A0392JMD JOB 33F02,JMD,TIME=(1,0),MSGLEVEL=(0,0),REGION=2048K
//*
//*
//*    This program computes the Min-V|Min-B values of  V
//*                                  -----------        -
//*                      and the Min-V*|Min-B* values of  V*
//*                                  -------------       --
//*              for various amounts of induced correlation
//*              under the IR, CR, and AR strategies.
//*
//*
//*    The SCDs (small composite designs) examined are:
//*              k = 3  SCD  (Hartley)
//*              k = 4  SCD  (Hartley)
//*              k = 5  SCD  (Draper)
//*              k = 6  SCD  (Hartley)
//*              k = 7  SCD  (Draper)
//*
//*
//*    This program uses the PROC MATRIX procedure of SAS to compute
//*              the values of  V  and  V*  using matrix algebra on
//*              the design matrix of each SCD.
//*
//*
/*JOBPARM LINES=4
/*PRIORITY IDLE
//STEP1   EXEC SAS
//SYSIN DD *
OPTIONS NOCENTER NODATE NONUMBER LS=120;
TITLE1;
```

```
* Input the Design Matrix  for the  "SCD3" design
  (k=3  Hartley  Small Composite Design);

DATA HAR3;

TITLE2 'K=3   SCD-Hartley';
INPUT BC BA X1 X2 X3;
X11=X1*X1; X22=X2*X2; X33=X3*X3;
X12=X1*X2; X13=X1*X3;
X23=X2*X3;
X111=X1*X1*X1; X122=X1*X2*X2; X133=X1*X3*X3;
X222=X2*X2*X2; X112=X1*X1*X2; X233=X2*X3*X3;
X333=X3*X3*X3; X113=X1*X1*X3; X223=X2*X2*X3;
X123=X1*X2*X3;

*
Input  BC=1 and BA=1   if a    Common   random number stream is used
Input           BA=-1  if an  Antithetic random number stream is used
Input  BC=0 and BA=0   if an Independent random number stream is used

Input  X1, X2, X3 =  1   or -1    for the values of the factorial points
Input  X1, X2, X3 =  1.0 or -1.0  for the values of the   axial   points
Input  X1, X2, X3 =      0        for the values of the  center   points

BC BA    X1          X2          X3;

CARDS;
1  1     1           1           1
1  1    -1          -1           1
1  1    -1           1          -1
1  1     1          -1          -1
1  -1   -1.0         0           0
1  -1    1.0         0           0
1  -1    0          -1.0         0
1  -1    0           1.0         0
1  -1    0           0          -1.0
1  -1    0           0           1.0
1  1     0           0           0
0  -1    0           0           0
;
PROC MATRIX FUZZ FW=15;
*
Define values of g, Nc, & Alpha for the J and J* criteria (both regions);
K=3;
F=4;
* For Sphere Box (BS);
GBS=0.516962000;
NCBS=1;*(factorial);
ALPBS=1.54919334;
* For Cube box (BC);
GBC=0.66280190;
NCBC=2;
ALP=1.67332005;
* For Sphere Star (SS);
GSS=0.58891494;
NCS=2;
* For Cube star (SC);
GSC=0.76028592;
FETCH D12 DATA=HAR3 COLNAME=CALL;
```

```
* Input the Design Matrix  for the  "SCD4" design
   (k=4  Hartley  Small Composite Design);

 DATA HAR4;

 TITLE2 'K=4   SCD-Hartley';
 INPUT BC B X1 X2 X3 X4;
 X11=X1*X1; X22=X2*X2; X33=X3*X3; X44=X4*X4;
 X12=X1*X2; X13=X1*X3; X14=X1*X4;
 X23=X2*X3; X24=X2*X4;
 X34=X3*X4;
 X111=X1*X1*X1; X122=X1*X2*X2; X133=X1*X3*X3; X144=X1*X4*X4;
 X222=X2*X2*X2; X112=X1*X1*X2; X233=X2*X3*X3; X244=X2*X4*X4;
 X333=X3*X3*X3; X113=X1*X1*X3; X223=X2*X2*X3; X344=X3*X4*X4;
 X444=X4*X4*X4; X114=X1*X1*X4; X224=X2*X2*X4; X334=X3*X3*X4;
 X123=X1*X2*X3; X124=X1*X2*X4; X134=X1*X3*X4; X234=X2*X3*X4;
 CARDS;
 1  1    1           -1          -1          -1
 1  1   -1            1          -1          -1
 1  1   -1           -1           1          -1
 1  1    1            1           1          -1
 1  1    1           -1          -1           1
 1  1   -1            1          -1           1
 1  1   -1           -1           1           1
 1  1    1            1           1           1
 1  -1  -1.0          0           0           0
 1  -1   1.0          0           0           0
 1  -1   0           -1.0         0           0
 1  -1   0            1.0         0           0
 1  -1   0            0          -1.0         0
 1  -1   0            0           1.0         0
 1  -1   0            0           0          -1.0
 1  -1   0            0           0           1.0
 1  -1   0            0           0           0
 0   1   0            0           0           0
 0   0   0            0           0           0
 ;
 PROC MATRIX FUZZ;
 *
 Define values of g, Nc, & Alpha for the J and J* criteria (both regions);
 K=4;
 F=8;
 * For Sphere Box (BS);
 GBS=0.41935912;
 NCBS=1;*(axial);
 ALPBS=2.12132034;
 * For Cube box (BC);
 GBC=0.632455532;
 NCBC=2;
 ALP=2;
 * For Sphere Star (SS);
 GSS=0.5;
 NCS=3;
 * For Cube star (SC);
 GSC=0.707106781;
 FETCH D12 DATA=HAR4 COLNAME=CALL;
```

```
* Input the Design Matrix  for the  "SCD5" design
  (k=5  Draper  Small Composite Design);

DATA DRA5;

TITLE2 'K=5   SCD-Draper';
INPUT BC B X1 X2 X3 X4 X5;
X11=X1*X1; X22=X2*X2; X33=X3*X3; X44=X4*X4; X55=X5*X5;
X12=X1*X2; X13=X1*X3; X14=X1*X4; X15=X1*X5;
X23=X2*X3; X24=X2*X4; X25=X2*X5;
X34=X3*X4; X35=X3*X5;
X45=X4*X5;
X111=X1*X1*X1; X122=X1*X2*X2; X133=X1*X3*X3; X144=X1*X4*X4; X155=X1*X5*X5;
X222=X2*X2*X2; X112=X1*X1*X2; X233=X2*X3*X3; X244=X2*X4*X4; X255=X2*X5*X5;
X333=X3*X3*X3; X113=X1*X1*X3; X223=X2*X2*X3; X344=X3*X4*X4; X355=X3*X5*X5;
X444=X4*X4*X4; X114=X1*X1*X4; X224=X2*X2*X4; X334=X3*X3*X4; X455=X4*X5*X5;
X555=X5*X5*X5; X115=X1*X1*X5; X225=X2*X2*X5; X335=X3*X3*X5; X445=X4*X4*X5;
X123=X1*X2*X3; X124=X1*X2*X4; X125=X1*X2*X5;
X134=X1*X3*X4; X135=X1*X3*X5; X145=X1*X4*X5;
X234=X2*X3*X4; X235=X2*X3*X5; X245=X2*X4*X5;
X345=X3*X4*X5;
CARDS;
1  1   -1          -1          -1          -1          -1
1  1    1           1           1           1           1
1  1   -1          -1           1           1           1
1  1   -1           1           1          -1           1
1  1    1          -1          -1           1           1
1  1    1           1          -1           1          -1
1  1    1           1           1          -1          -1
1  1    1          -1           1          -1          -1
1  1    1          -1          -1          -1           1
1  1   -1           1          -1           1          -1
1  1   -1           1          -1          -1           1
1  1   -1          -1           1           1          -1
1  -1  -1.0         0           0           0           0
1  -1   1.0         0           0           0           0
1  -1   0          -1.0         0           0           0
1  -1   0           1.0         0           0           0
1  -1   0           0          -1.0         0           0
1  -1   0           0           1.0         0           0
1  -1   0           0           0          -1.0         0
1  -1   0           0           0           1.0         0
1  -1   0           0           0           0          -1.0
1  -1   0           0           0           0           1.0
1  -1   0           0           0           0           0
0  1    0           0           0           0           0
0  0    0           0           0           0           0
;
PROC MATRIX FUZZ;
*
Define values of g, No, & Alpha for the J and J* criteria (both regions);
K=5;
F=12;
* For Sphere Box (BS);
GBS=0.38529925;
NCBS=1;*(axial);
ALPBS=2.34520788;
* For Cube box (BC);
GBC=0.61964446;
NCBC=2;
ALP=2.25320285;
* For Sphere Star (SS);
GSS=0.44564517;
NCS=3;
* For Cube star (SC);
GSC=0.68073424;
FETCH D12 DATA=DRA5 COLNAME=CALL;
```

```
* Input the Design Matrix  for the  "SCD6" design
  (k=6  Hartley  Small Composite Design);

DATA HAR6;

TITLE2 'K=6   SCD-Hartley';
INPUT BC B X1 X2 X3 X4 X5 X6;
X11=X1*X1; X22=X2*X2; X33=X3*X3; X44=X4*X4; X55=X5*X5; X66=X6*X6;
X12=X1*X2; X13=X1*X3; X14=X1*X4; X15=X1*X5; X16=X1*X6;
X23=X2*X3; X24=X2*X4; X25=X2*X5; X26=X2*X6;
X34=X3*X4; X35=X3*X5; X36=X3*X6;
X45=X4*X5; X46=X4*X6;
X56=X5*X6;
X111=X1*X1*X1;
X122=X1*X2*X2;X133=X1*X3*X3;X144=X1*X4*X4;X155=X1*X5*X5;X166=X1*X6*X6;
X222=X2*X2*X2;
X112=X1*X1*X2;X233=X2*X3*X3;X244=X2*X4*X4;X255=X2*X5*X5;X266=X2*X6*X6;
X333=X3*X3*X3;
X113=X1*X1*X3;X223=X2*X2*X3;X344=X3*X4*X4;X355=X3*X5*X5;X366=X3*X6*X6;
X444=X4*X4*X4;
X114=X1*X1*X4;X224=X2*X2*X4;X334=X3*X3*X4;X455=X4*X5*X5;X466=X4*X6*X6;
X555=X5*X5*X5;
X115=X1*X1*X5;X225=X2*X2*X5;X335=X3*X3*X5;X445=X4*X4*X5;X566=X5*X6*X6;
X666=X6*X6*X6;
X116=X1*X1*X6;X226=X2*X2*X6;X336=X3*X3*X6;X446=X4*X4*X6;X556=X5*X5*X6;
X123=X1*X2*X3; X124=X1*X2*X4; X125=X1*X2*X5; X126=X1*X2*X6;
X134=X1*X3*X4; X135=X1*X3*X5; X136=X1*X3*X6;
X145=X1*X4*X5; X146=X1*X4*X6;
X156=X1*X5*X6;
X234=X2*X3*X4; X235=X2*X3*X5; X236=X2*X3*X6;
X245=X2*X4*X5; X246=X2*X4*X6;
X256=X2*X5*X6;
X345=X3*X4*X5; X346=X3*X4*X6; X356=X3*X5*X6;
X456=X4*X5*X6;
CARDS;
1  1    1        -1        -1         1        -1      -1
1  1   -1         1        -1         1        -1      -1
1  1   -1        -1         1         1        -1      -1
1  1    1         1         1         1        -1      -1
1  1    1        -1        -1        -1         1      -1
1  1   -1         1        -1        -1         1      -1
1  1   -1        -1         1        -1         1      -1
1  1    1         1         1        -1         1      -1
1  1    1        -1        -1        -1        -1       1
1  1   -1         1        -1        -1        -1       1
1  1   -1        -1         1        -1        -1       1
1  1    1         1         1        -1        -1       1
1  1    1        -1        -1         1         1       1
1  1   -1         1        -1         1         1       1
1  1   -1        -1         1         1         1       1
1  1    1         1         1         1         1       1
1 -1   -1.0       0         0         0         0       0
1 -1    1.0       0         0         0         0       0
1 -1    0        -1.0       0         0         0       0
1 -1    0         1.0       0         0         0       0
1 -1    0         0        -1.0       0         0       0
1 -1    0         0         1.0       0         0       0
1 -1    0         0         0        -1.0       0       0
1 -1    0         0         0         1.0       0       0
1 -1    0         0         0         0        -1.0     0
1 -1    0         0         0         0         1.0     0
1 -1    0         0         0         0         0      -1.0
1 -1    0         0         0         0         0       1.0
1 -1    0         0         0         0         0       0
0  1    0         0         0         0         0       0
0  0    0         0         0         0         0       0
;

PROC MATRIX FUZZ;
*
Define values of g, No. & Alpha for the J and J* criteria (both regions);
K=6;
F=16;
* For Sphere Box (BS);
GBS=0.3585134;
NCBS=1;*(axial);
ALPBS=2.54950976;
* For Cube box (BC);
GBC=0.61204162;
NCBC=2;
ALP=2.47338777;
* For Sphere Star (SS);
GSS=0.40652488;
NCS=3;
* For Cube star (SC);
GSC=0.66385234;
FETCH D12 DATA=HAR6 COLNAME=CALL;
```

```
* Input the Design Matrix  for the  "SCD7" design
  (k=7  Draper. Small Composite Design);

DATA DRA7;

TITLE2 'K=7   SCD-Draper';
INPUT BC B X1 X2 X3 X4 X5 X6 X7;
X11=X1*X1; X22=X2*X2; X33=X3*X3; X44=X4*X4; X55=X5*X5; X66=X6*X6; X77=X7*X7;
X12=X1*X2; X13=X1*X3; X14=X1*X4; X15=X1*X5; X16=X1*X6; X17=X1*X7;
X23=X2*X3; X24=X2*X4; X25=X2*X5; X26=X2*X6; X27=X2*X7;
X34=X3*X4; X35=X3*X5; X36=X3*X6; X37=X3*X7;
X45=X4*X5; X46=X4*X6; X47=X4*X7;
X56=X5*X6; X57=X5*X7;
X67=X6*X7;
X111=X1*X1*X1;X122=X1*X2*X2;X133=X1*X3*X3;X144=X1*X4*X4;X155=X1*X5*X5;X166=X1*X6*X6;X177=X1*X7*X7;
X222=X2*X2*X2;X112=X1*X1*X2;X233=X2*X3*X3;X244=X2*X4*X4;X255=X2*X5*X5;X266=X2*X6*X6;X277=X2*X7*X7;
X333=X3*X3*X3;X113=X1*X1*X3;X223=X2*X2*X3;X344=X3*X4*X4;X355=X3*X5*X5;X366=X3*X6*X6;X377=X3*X7*X7;
X444=X4*X4*X4;X114=X1*X1*X4;X224=X2*X2*X4;X334=X3*X3*X4;X455=X4*X5*X5;X466=X4*X6*X6;X477=X4*X7*X7;
X555=X5*X5*X5;X115=X1*X1*X5;X225=X2*X2*X5;X335=X3*X3*X5;X445=X4*X4*X5;X566=X5*X6*X6;X577=X5*X7*X7;
X666=X6*X6*X6;X116=X1*X1*X6;X226=X2*X2*X6;X336=X3*X3*X6;X446=X4*X4*X6;X556=X5*X5*X6;X677=X6*X7*X7;
X777=X7*X7*X7;X117=X1*X1*X7;X227=X2*X2*X7;X337=X3*X3*X7;X447=X4*X4*X7;X557=X5*X5*X7;X667=X6*X6*X7;
X123=X1*X2*X3;  X124=X1*X2*X4;  X125=X1*X2*X5;  X126=X1*X2*X6;X127=X1*X2*X7;
X134=X1*X3*X4;  X135=X1*X3*X5;  X136=X1*X3*X6;  X137=X1*X3*X7;
X145=X1*X4*X5;  X146=X1*X4*X6;  X147=X1*X4*X7;
X156=X1*X5*X6;  X157=X1*X5*X7;
X167=X1*X6*X7;
X234=X2*X3*X4;  X235=X2*X3*X5;  X236=X2*X3*X6;  X237=X2*X3*X7;
X245=X2*X4*X5;  X246=X2*X4*X6;  X247=X2*X4*X7;
X256=X2*X5*X6;  X257=X2*X5*X7;
X267=X2*X6*X7;
X345=X3*X4*X5;  X346=X3*X4*X6;  X347=X3*X4*X7;
X356=X3*X5*X6;  X357=X3*X5*X7;
X367=X3*X6*X7;
X456=X4*X5*X6;  X457=X4*X5*X7;
X467=X4*X6*X7;
X567=X5*X6*X7;
CARDS;
1  1    1     1       -1       1       1     -1     1
1  1   -1     1        1      -1       1      1    -1
1  1    1    -1        1       1      -1      1     1
1  1    1     1       -1       1       1     -1     1
1  1    1     1        1      -1       1      1    -1
1  1    1     1        1       1      -1      1     1
1  1   -1     1        1       1       1     -1     1
1  1   -1    -1        1       1       1      1    -1
1  1   -1    -1       -1       1       1      1     1
1  1   -1    -1       -1      -1       1      1     1
1  1    1    -1       -1      -1      -1      1     1
1  1   -1     1       -1      -1      -1     -1     1
1  1   -1    -1        1      -1      -1     -1    -1
1  1   -1    -1       -1       1      -1     -1    -1
1  1    1    -1       -1      -1       1     -1    -1
1  1   -1     1       -1      -1      -1      1    -1
1  1   -1    -1        1      -1      -1     -1     1
1  1    1    -1       -1       1      -1     -1    -1
1  1    1     1       -1      -1       1     -1    -1
1  1    1     1        1      -1      -1      1    -1
1  1   -1     1        1       1      -1     -1     1
1  1    1    -1        1       1       1     -1    -1
1  1   -1     1       -1       1       1      1    -1
1  1    1    -1        1      -1       1      1     1
1  1    1     1       -1       1      -1      1     1
1  1   -1     1        1      -1       1     -1     1
1  1    1    -1        1       1      -1      1    -1
1  1   -1    -1       -1      -1      -1     -1    -1
1 -1   -1.0   0        0       0       0      0     0
1 -1    1.0   0        0       0       0      0     0
1 -1   0     -1.0      0       0       0      0     0
1 -1   0      1.0      0       0       0      0     0
1 -1   0      0       -1.0     0       0      0     0
1 -1   0      0        1.0     0       0      0     0
1 -1   0      0        0      -1.0     0      0     0
1 -1   0      0        0       1.0     0      0     0
1 -1   0      0        0       0      -1.0    0     0
1 -1   0      0        0       0       1.0    0     0
1 -1   0      0        0       0       0     -1.0   0
1 -1   0      0        0       0       0      1.0   0
1 -1   0      0        0       0       0      0    -1.0
1 -1   0      0        0       0       0      0     1.0
1 -1   0      0        0       0       0      0     0
0  1   0      0        0       0       0      0     0
0  0   0      0        0       0       0      0     0
;
PROC MATRIX FUZZ;
K=7;
F=28;
* For Sphere Box (BS);
GBS=.33988978;
NCBS=2;
ALPBS=2.69098111;
* For Cube box (BC);
GBC=.60589849;
NCBC=2;
ALP=2.69098111;
* For Sphere Star (SS);
GSS=.37576229;
NCS=3;
* For Cube star (SC);
GSC=.65083938;
FETCH D12 DATA=DRA7 COLNAME=CALL;
```

```
*
This program is run following EACH of the Data files for the SCDs
---- ------- -- --- --------- ------- --- ---- ----- --- --- ----

Input '1' if Spherical  &  '2' if Cuboidal;
REGION=2;
TITLE4 'Spherical Region';
TITLE4 'Cuboidal Region';
*
Compute the region moments;
IF REGION=1 THEN DO;
NCB=NCBS;
GB=GBS;
GS=GSS;
ALPB=ALPBS;
WII=1#/(K+2);
WIIJJ=WII#/(K+4);
WIIII=3#WIIJJ;
END;
IF REGION=2 THEN DO;
NCB=NCBC;
GB=GBC;
GS=GSC;
ALPB=ALP;
WII=1#/3;
WIIJJ=1#/9;
WIIII=1#/5;
END;
*
Compute the number of fitted & unfitted model parameters;
P1=(K+1)#(K+2)#/2;      P2=K#(K+1)#(K+2)#/6;
IK=I(K);                K2=K#(K-1)#/2;              IK2=I(K2);
AX=2#K;                 NB=F+AX+NCB;                NS=F+AX+NCS;
M=NCOL(D12);            IDB=I(NB);                  IDS=I(NS);
*
Construct the region moment matrices  u11 & u*11;
Z1K=J(1,K,0);           Z1K2=J(1,K2,0);             ZK1=Z1K';
ZKK=J(K,K,0);           ZKK2=J(K,K2,0);             ZK2K2=J(K2,K2,0);
ZK21=Z1K2';             ZK2K=ZKK2';
UU1KK=J(1,K,WII);       UUKK=WII#IK;                UUKK1=UU1KK';
UUK2K2=WIIJJ#IK2;       UUSKKK=4#WII#IK;            UUSK2K2=2#WII#IK2;
UUKKK=J(K,K,WIIJJ)+IK#(WIIII-WIIJJ);
U11=(1||Z1K||UU1KK||Z1K2)//(ZK1||UUKK||ZKK||ZKK2)
    //(UUKK1||ZKK||UUKKK||ZKK2)//(ZK21||ZK2K||ZK2K||UUK2K2);
US11=(0||Z1K||Z1K||Z1K2)//(ZK1||IK||ZKK||ZKK2)
    //(ZK1||ZKK||UUSKKK||ZKK2)//(ZK21||ZK2K||ZK2K||UUSK2K2);
*
Compute the values of the scaled design points for the J & J* criteria;
GB2=GB#GB;              GS2=GS#GS;
GB3=GB2#GB;             GS3=GS2#GS;
ALPB2=ALPB#ALPB;        ALP2=ALP#ALP;
X0B=J(NB,1,1);          X0S=J(NS,1,1);
X1AB=(D12(1:F,3:K+2)#GB)//(D12(F+1:NB,3:K+2)#ALPB#GB);
X1AS=(D12(1:F,3:K+2)#GS)//(D12(F+1:NS,3:K+2)#ALP#GS);
X1BB=(D12(1:F,K+3:P1+1)#GB2)//(D12(F+1:NB,K+3:P1+1)#ALPB2#GB2);
X1BS=(D12(1:F,K+3:P1+1)#GS2)//(D12(F+1:NS,K+3:P1+1)#ALP2#GS2);
X1B=X0B||X1AB||X1BB;  X1S=X0S||X1AS||X1BS;
X2B=(D12(1:F,P1+2:M)#GB3)//J(NB-F,P2,0);
X2S=(D12(1:F,P1+2:M)#GS3)//J(NS-F,P2,0);
XPXB=X1B'*X1B;          IXPXB=INV(XPXB);
XPXS=X1S'*X1S;          IXPXS=INV(XPXS);
ZB=X1B*IXPXB;           ZBP=ZB';
ZS=X1S*IXPXS;           ZSP=ZS';
*
Retrieve the u  & v  vectors  for the CR & AR strategies;
UBC=D12(1:NB,1);        USC=D12(1:NS,1);
UBMC=DIAG(UBC);         USMC=DIAG(USC);
```

```
VB=D12(1:NB,2);         VS=D12(1:NS,2);
UB=VB##2;               US=VS##2;
UBM=DIAG(UB);           USM=DIAG(US);
*
Compute  Var(y), Var(b), V, & V*  under the  IR  Strategy;
VIRB=IDB;
VIRS=IDS;
VARBIRB=ZBP*VIRB*ZB;
VARBIRS=ZSP*VIRS*ZS;
VBOXIR=NB#(TRACE(U11*VARBIRB));
VSTARIR=NS#(TRACE(US11*VARBIRS));
OUTIR=VBOXIR||VSTARIR;
*
Compute  Var(y), Var(b), V, & V*  under the  CR  Strategy;
COUNTRP=0;              OUTCR=J(4,3);
COUNT=0;                OUTAR=J(10,4);
DO RP=0 TO .9 BY .3;
COUNTRP=1+COUNTRP;
VCRB=IDB+RP#(UBC*UBC')-RP#UBMC;
VCRS=IDS+RP#(USC*USC')-RP#USMC;
VARBCRB=ZBP*VCRB*ZB;
VARBCRS=ZSP*VCRS*ZS;
VBOXCR=NB#(TRACE(U11*VARBCRB));
VSTARCR=NS#(TRACE(US11*VARBCRS));
OUTCR(COUNTRP,1)=RP;
OUTCR(COUNTRP,2)=VBOXCR;
OUTCR(COUNTRP,3)=VSTARCR;
*
Compute  Var(y), Var(b), V, & V*  under the  AR  Strategy;
DO RM=0 TO .9 BY .3 WHILE(RM LE RP);
COUNT=1+COUNT;
VARB=IDB+.5#(RP-RM)#(UB*UB')+.5#(RP+RM)#(VB*VB')-RP#UBM;
VARS=IDS+.5#(RP-RM)#(US*US')+.5#(RP+RM)#(VS*VS')-RP#USM;
VARBARB=ZBP*VARB*ZB;  VARBARS=ZSP*VARS*ZS;
VBOXAR=NB#(TRACE(U11*VARBARB));
VSTARAR=NS#(TRACE(US11*VARBARS));
OUTAR(COUNT,1)=RP;
OUTAR(COUNT,2)=RM;
OUTAR(COUNT,3)=VBOXAR;
OUTAR(COUNT,4)=VSTARAR;
END;*RP;
END;*RM;
*
Print Output values of  V & V*;
CIR='IRbox' 'IRstar';
CCR='RP' 'CRbox' 'CRstar';
CAR='RP' 'RM' 'ARbox' 'ARstar';
RALL='1' '2' '3' '4' '5' '6' '7' '8' '9' '10';
PRINT OUTIR COLNAME=CIR ROWNAME=RALL;
PRINT OUTCR COLNAME=CCR ROWNAME=RALL;
PRINT OUTAR COLNAME=CAR ROWNAME=RALL;
```

Note: The computer output on the next four pages was compiled from ten runs of the above computer program (5 SCDs in 2 regions). The form is similar to the computer output for the CCDs, BBDs, and FACs shown in Appendices Q and R.

Min-V | Min-B Designs

Spherical Region

V for the SCDs under the CR Strategy

RP	VCSCD3B	VCSCD4B	VCSCD5B	VCSCD6B	VCSCD7B
0.0	15.2269	23.1583	48.6213	49.0201	102.220
0.3	13.9588	21.3108	40.9349	43.0141	84.401
0.6	12.6908	19.4633	33.2485	37.0080	66.583
0.9	11.4227	17.6158	25.5621	31.0020	48.765

V for the SCDs under the AR Strategy

RP	RM	VASCD3B	VASCD4B	VASCD5B	VASCD6B	VASCD7B
0.0	0.0	15.2269	23.1583	48.6213	49.0201	102.220
0.3	0.0	12.3225	18.7696	37.4914	38.7106	84.574
0.3	0.3	10.6861	16.2284	34.0479	34.4072	84.394
0.6	0.0	9.4180	14.3810	26.3616	28.4011	66.928
0.6	0.3	7.7817	11.8398	22.9181	24.0977	66.749
0.6	0.6	6.1453	9.2986	19.4746	19.7942	66.569
0.9	0.0	6.5136	9.9923	15.2317	18.0917	49.283
0.9	0.3	4.8772	7.4511	11.7882	13.7882	49.103
0.9	0.6	3.2409	4.9099	8.3447	9.4848	48.923
0.9	0.9	1.6045	2.3688	4.9013	5.1813	48.744

Min-V | Min-B Designs

Cuboidal Region

V for the SCDs under the CR Strategy

RP	VCSCD3B	VCSCD4B	VCSCD5B	VCSCD6B	VCSCD7B
0.0	16.6699	24.00	51.8174	50.9878	107.381
0.3	14.8900	21.75	42.9756	44.1620	87.665
0.6	13.1101	19.50	34.1339	37.3363	67.949
0.9	11.3302	17.25	25.2921	30.5105	48.233

V for the SCDs under the AR Strategy

RP	RM	VASCD3B	VASCD4B	VASCD5B	VASCD6B	VASCD7B
0.0	0.0	16.6699	24.0	51.8174	50.9878	107.381
0.3	0.0	13.5190	19.5	39.8972	40.2715	88.668
0.3	0.3	11.7690	16.8	36.3222	35.8515	88.969
0.6	0.0	10.3680	15.0	27.9770	29.5551	69.955
0.6	0.3	8.6180	12.3	24.4020	25.1351	70.256
0.6	0.6	6.8680	9.6	20.8270	20.7151	70.557
0.9	0.0	7.2170	10.5	16.0567	18.8388	51.242
0.9	0.3	5.4670	7.8	12.4817	14.4188	51.543
0.9	0.6	3.7170	5.1	8.9067	9.9988	51.844
0.9	0.9	1.9670	2.4	5.3317	5.5788	52.145

Min-V* | Min-B* Designs

Spherical Region

V* for the SCDs under the CR Strategy

RP	VCSCD3S	VCSCD4S	VCSCD5S	VCSCD6S	VCSCD7S
0.0	97.2231	186.306	500.031	569.370	1432.79
0.3	72.3266	140.547	364.295	417.127	1030.68
0.6	47.4302	94.789	228.560	264.884	628.57
0.9	22.5338	49.031	92.824	112.641	226.46

V* for the SCDs under the AR Strategy

RP	RM	VASCD3S	VASCD4S	VASCD5S	VASCD6S	VASCD7S
0.0	0.0	97.2231	186.306	500.031	569.370	1432.79
0.3	0.0	68.0562	132.947	353.598	403.228	1098.68
0.3	0.3	68.0562	132.103	352.417	401.709	1185.17
0.6	0.0	38.8892	79.589	207.166	237.087	764.58
0.6	0.3	38.8892	78.744	205.985	235.567	851.07
0.6	0.6	38.8892	77.900	204.803	234.047	937.56
0.9	0.0	9.7223	26.231	60.733	70.945	430.48
0.9	0.3	9.7223	25.386	59.552	69.426	516.97
0.9	0.6	9.7223	24.542	58.371	67.906	603.46
0.9	0.9	9.7223	23.697	57.189	66.386	689.95

Min-V* | Min-B* Designs

Cuboidal Region

V* for the SCDs under the CR Strategy

RP	VCSCD3S	VCSCD4S	VCSCD5S	VCSCD6S	VCSCD7S
0.0	58.3338	93.1528	214.299	213.514	477.595
0.3	43.3960	70.2736	156.127	156.423	343.559
0.6	28.4581	47.3944	97.954	99.331	209.523
0.9	13.5203	24.5153	39.782	42.240	75.487

V* for the SCDs under the AR Strategy

RP	RM	VASCD3S	VASCD4S	VASCD5S	VASCD6S	VASCD7S
0.0	0.0	58.3338	93.1528	214.299	213.514	477.595
0.3	0.0	40.8337	66.4736	151.542	151.211	366.228
0.3	0.3	40.8337	66.0514	151.036	150.641	395.058
0.6	0.0	23.3335	39.7944	88.785	88.908	254.861
0.6	0.3	23.3335	39.3722	88.279	88.338	283.691
0.6	0.6	23.3335	38.9500	87.773	87.768	312.521
0.9	0.0	5.8334	13.1153	26.028	26.604	143.493
0.9	0.3	5.8334	12.6931	25.522	26.035	172.324
0.9	0.6	5.8334	12.2708	25.016	25.465	201.154
0.9	0.9	5.8334	11.8486	24.510	24.895	229.984

The vita has been removed from the scanned document